

# jero

2  
—  
2020

*Journal for  
Educational Research  
Online*

---

*Journal für Bildungsforschung Online*

WAXMANN

**JERO**  
**Journal for Educational Research Online**  
**Journal für Bildungsforschung Online**

*Executive Editors:*

Nele McElvany, Institut für Schulentwicklungsforschung, Technische Universität Dortmund

Cornelia Gräsel, Institut für Bildungsforschung in der School of Education, Bergische Universität Wuppertal

*Editorial Office:*

Rahim Schaufelberger, Institut für Schulentwicklungsforschung, Technische Universität Dortmund

Michael Rochnia, Institut für Bildungsforschung in der School of Education, Bergische Universität Wuppertal

E-mail: editorialoffice@j-e-r-o.com

*Editors:*

Rolf Becker (Bern), Oliver Böhm-Kasper (Bielefeld), Wolfgang Böttcher (Münster), Martin Bonsen (Münster), Wilfried Bos (Dortmund), Hartmut Ditton (München), Cornelia Gräsel (Wuppertal), Stephan Huber (Zug), Olaf Köller (Kiel), Harm Kuper (Berlin), Katharina Maag Merki (Zürich), Romain Martin (Luxemburg), Nele McElvany (Dortmund), Florian H. Müller (Klagenfurt), R. Daniel Muijs (Southampton), Jost Reinecke (Bielefeld), Kerstin Schneider (Wuppertal), Claudia Schuchart (Wuppertal), Wolfram Schulz (Camberwell), Birgit Spinath (Heidelberg), Petra Stanat (Berlin), Jan Van Damme (Leuven), Stefanie van Ophuysen (Münster), Andreas Voss (Hamburg), Horst Weishaupt (Frankfurt), Ludger Wößmann (München), Olga Zlatkin-Troitschanskaia (Mainz)

Please subscribe to the e-mail alert at [www.waxmann.com/jero](http://www.waxmann.com/jero) if you want to be informed about the publication of each new issue. There is a spring edition and an autumn edition each year. The subscription of the journal is free of charge.

Open access at [www.waxmann.com/jero](http://www.waxmann.com/jero)

ISSN 1866-6671

© Waxmann Verlag GmbH 2020  
Steinfurter Straße 555, 48159 Münster, Germany  
Internet: [www.waxmann.com](http://www.waxmann.com), E-mail: [info@waxmann.com](mailto:info@waxmann.com)

Setting: Stoddart Satz- und Layoutservice, Münster  
All rights reserved

# Inhalt

<i>Paul Fabian &amp; Nele McElvany</i>	
Bildungsprozesse, Bildungsentscheidungen und Kompetenzentwicklungen in verschiedenen Lebensphasen – Aktuelle Forschungsbefunde aus NEPS.....	3
<i>Anja Linberg, Manja Attig &amp; Sabine Weinert</i>	
Soziale Disparitäten im kindlichen Wortschatz von zweijährigen Kindern und ihre Vermittlung durch sprachanregendes Interaktionsverhalten.....	12
<i>Michaela Sixt &amp; Christian Aßmann</i>	
Der Einfluss regionaler Bildungsinfrastruktur und Arbeitsmarktbedingungen beim Übergangsprozess zu weiterführenden Schulen in Deutschland.....	36
<i>Florian Wohlkinger &amp; Michael Bayer</i>	
Typologie des Schülerhabitus in der Sekundarstufe I: Ein quantitativer Operationalisierungsansatz mit Daten des Nationalen Bildungspanels .....	67
<i>Jennifer Lorenz &amp; Tobias C. Stubbe</i>	
Nachhilfe als Mittel zum sozialen Statuserhalt .....	89
<i>Ruben Kleinkorres, Justine Stang &amp; Nele McElvany</i>	
Eine Längsschnittanalyse der wechselseitigen Beziehungen zwischen schulischem Wohlbefinden und akademischer Leistung .....	114
<i>Karin Gehrer &amp; Lena Nusser</i>	
Binnendifferenzierender Deutschunterricht und dessen Einfluss auf die Lesekompetenzentwicklung in der Sekundarstufe I .....	166
<i>Divan Mouton, Hui Zhang &amp; Bernhard Ertl</i>	
Gründe für den Studienabbruch deutscher Universitätsstudierender: Analyse latenter Klassen von Studienabbrechern .....	190

## **Contents**

<i>Paul Fabian &amp; Nele McElvany</i> Educational processes, educational decisions and competence development across the lifespan – recent research results from the NEPS .....	3
<i>Anja Linberg, Manja Attig &amp; Sabine Weinert</i> Social disparities in the vocabulary of 2-year-old children and the mediating effect of language-stimulating interaction behavior .....	12
<i>Michaela Sixt &amp; Christian Aßmann</i> The influence of regional school infrastructure and labor market conditions on the transition process to secondary schooling in Germany.....	36
<i>Florian Wohlkinger &amp; Michael Bayer</i> Typology of students' academic habitus in lower secondary education: A quantitative operationalization approach with data from the National Educational Panel Study .....	67
<i>Jennifer Lorenz &amp; Tobias C. Stubbe</i> Private tutoring as a means for maintaining social status.....	89
<i>Ruben Kleinkorres, Justine Stang &amp; Nele McElvany</i> A longitudinal analysis of reciprocal relations between students' well-being and academic achievement.....	114
<i>Karin Gehrer &amp; Lena Nusser</i> Adaptive instruction in German classes and its effect on reading competence development in secondary school .....	166
<i>Divan Mouton, Hui Zhang &amp; Bernhard Ertl</i> German university student's reasons for dropout: Identifying latent classes .....	190

Paul Fabian & Nele McElvany

## Bildungsprozesse, Bildungsentscheidungen und Kompetenzentwicklungen in verschiedenen Lebensphasen – Aktuelle Forschungsbefunde aus NEPS

### 1. Einleitung

Die Perspektive des Lebenslangen Lernens hat im Rahmen der Bildungsforschung stark an Bedeutung gewonnen. Mit Globalisierungsprozessen, der Entwicklung von Technologien sowie den Anforderungen für ein selbstbestimmtes Leben und gesellschaftliche Teilhabe gehen neue Herausforderungen für das Individuum einher. Eine Flexibilisierung und Weiterentwicklung der staatlichen und gesellschaftlich konstituierten Bildungseinrichtungen und -möglichkeiten zur Unterstützung von Individuen bei ihrem lebenslangen Lernprozess ist dabei von zentraler Bedeutung für das Gelingen individueller Lebensverläufe und den Erfolg gesellschaftlichen Zusammenlebens. Die empirische Bildungsforschung hat dabei maßgeblich die Aufgabe, Stellschrauben zu identifizieren und diese mit Bildungspraxis, Bildungsadministration und Bildungspolitik fruchtbar zu kommunizieren (McElvany, Gebauer & Gräsel, 2019) sowie Optimierungsoptionen aufzuzeigen.

### 2. Lebenslanges Lernen

Im Rahmen der Analyse Lebenslangen Lernens gibt es verschiedene Ansätze, die – teilweise disziplinär geprägt – komplementäre, aber auch distinkte Herangehensweisen und Theorien favorisieren, insbesondere da die Disziplinen teilweise auf unterschiedliche Outcomes fokussieren (Mayer, 2003). Mayer (2009) führt aus, dass es bis dato keine interdisziplinäre Theorie der Entwicklung von zentralen, über die Lebensspanne relevanten Outcomes gibt, sondern dass diese immer disziplinär geprägt – wenngleich auch teilweise integrierend – sind. So gibt es beispielsweise ökonomische Theorien zur Erklärung verschiedener Prozesse und

---

Dr. Paul Fabian (corresponding author) · Prof. Dr. Nele McElvany, Institut für Schulentwicklungs- und Bildungsforschung (IFS), Technische Universität Dortmund, Vogelpothsweg 78, 44227 Dortmund

E-Mail: paul.fabian@tu-dortmund.de  
nele.mcelvany@tu-dortmund.de

Outcomes im Rahmen von Investment/Return-Überlegungen und Humankapital („human capital theory“, Becker, 1964; Behrmann, 2003). Psychologisch geprägte Forschung stützt sich im Rahmen der Lebensverlaufsforschung unter anderem auf entwicklungspsychologische Forschungsansätze („selection, optimization, and compensation“, Baltes, 2003; „assimilation and accommodation“, Block, 1982; „stage-environment-fit“, Eccles & Migley, 1989).

Alle theoretischen Ansätze, die eine Entwicklungsperspektive im Lebensverlauf fokussieren, haben jedoch bezogen auf die methodischen Voraussetzungen und Notwendigkeiten zur empirischen Überprüfung einige zentrale Gemeinsamkeiten. Elder, Johnson und Crosnoe (2003) beschreiben fünf theoretische Prinzipien, die im Kontext der Erforschung des Lebenslangen Lernens von zentraler Bedeutung sind (siehe auch Blossfeld & von Maurice, 2019).

## **2.1 Das Prinzip der Entwicklung im Lebensverlauf „The principle of lifespan development“)**

Zur Eingrenzung des Begriffs des Lebenslangen Lernens steht das Prinzip der Entwicklung im Lebensverlauf im Vordergrund. Das bedeutet einerseits, dass für das Lebenslange Lernen die zentrale Erkenntnis steht, dass Prozesse in langen Zeiträumen betrachtet werden müssen, da in modernen Wissensgesellschaften bereits im frühkindlichen Bereich zentrale Bildungsprozesse – sowohl institutionell als auch non-formal – beginnen. Gleichzeitig endet die Bildung von Individuen nicht mit dem Austritt aus dem formalisierten Bildungswesen. Gerade in der sich immer schneller wandelnden Welt ist die konstante Auseinandersetzung mit neuem Wissen und Anforderungen von zentraler Bedeutung für die Teilhabe an gesellschaftlichen Prozessen (z.B. Campbell, 2006), für die Arbeitsumwelt (Becker, 1964) und die langfristige Lebenszufriedenheit (Diener, 1984). Im Kontext der bildungswissenschaftlichen Forschung besteht die zentrale Herausforderung in der Veränderung der Gewichtung formeller, informeller und non-formaler Lernumwelten für die Entwicklung von Fähigkeiten und Kompetenzen in unterschiedlichen Lebensphasen.

## **2.2 Das Prinzip interdependenten Lebens und geteilter Beziehungen („The principle of linked lives“)**

Das zweite grundlegende Prinzip betont die Wichtigkeit der Berücksichtigung von sich gegenseitig beeinflussenden Individuen (vgl. Bronfenbrenner, 1981, 1986). Vor allem im Kontext des Lebenslangen Lernens stehen hier einerseits die Familien bzw. deren Mitglieder (Großeltern, Eltern, Geschwister, weitere Verwandte) im Fokus. Doch auch sich über die Zeit verändernde Gruppen stellen zentrale Entwicklungskontexte des Individuums dar – innerhalb von Institutionen als

Mitglied von Gruppen im Kindergarten, als Mitlerner\*innen sowie außerhalb von Institutionen in Form von Freundschaftsbeziehungen.

## **2.3 Das Prinzip der Fähigkeit zum Handeln („The principle of agency“)**

Jahrzehnte soziologischer, psychologischer und pädagogischer Forschung haben gezeigt, von welch großer Bedeutung frühe Prozesse für die *Befähigung* zu lebenslangem Lernen sind. Die Ausbildung eines Selbstkonzepts und die Erfahrung von Selbstwirksamkeit, die bereits im frühkindlichen Bereich beeinflusst werden, stellen hier zentrale psychologische Konstrukte dar (Bandura 1982). Daneben stehen auch soziologische Theorien (prominent die der „rational choice theory/rational action theory“, Breen & Goldthorpe, 1997; Erikson & Jonsson, 1996; Esser, 1999), die auf Entscheidungen fokussieren, die den weiteren Bildungsverlauf massiv beeinflussen. Die Befähigung, diese Entscheidungen informiert zu treffen, gilt im Allgemeinen als Ziel von Bildung und Lebenslangem Lernen.

## **2.4 Das Prinzip des Zeitpunktes („The principle of timing“)**

Der Zeitpunkt einer Entscheidung oder des Wechsels eines Kontextes ist für den weiteren Verlauf einer Bildungskarriere und damit des Lebenslangen Lernens von großer Bedeutung. So kann, bezugnehmend auf die Wert-Erwartungs-Theorie (Wigfield & Eccles, 2000), davon ausgegangen werden, dass eine Bildungsentscheidung (wie bspw. der Übergang auf eine Schulform im Kontext des deutschen, hierarchisch gegliederten Schulsystems, Maaz, Hausen, McElvany & Baumert, 2006), von zentralen, nicht zeitstabilen Faktoren abhängt. Diese Faktoren können – je nachdem wann sie zum Tragen kommen, unterschiedliche Ergebnisse erzeugen. Demnach ist es von großer Bedeutung, *wann* eine Entscheidung zu treffen ist.

## **2.5 Das Prinzip des Zeitgeschehens („The principle of time and place“)**

Das fünfte Prinzip ist das des Zeitgeschehens. Gerade im Kontext des Erscheinungsjahres dieses Sonderbandes, das von der weltweiten Corona-Pandemie geprägt ist, wird die Wichtigkeit dieses Prinzips noch einmal besonders deutlich: Kritische Lebensereignisse (Filipp & Aymanns, 2010) beeinflussen das Individuum, aber auch die Systeme, in denen es sich bewegt. Derzeit ist nicht abzusehen, wie sich sowohl die Individuen, als auch deren Familien, die Institutionen und Arbeitsumfelder entwickeln und welchen Einfluss die einzelnen Bereiche ent-

falten, was die Wichtigkeit dieses Prinzips für die Analyse von Bildungsprozessen, Bildungsentscheidungen und Kompetenzentwicklung noch einmal unterstreicht.

Mayer (2009) führt ein zusätzliches sechstes Prinzip auf, das die Relevanz der Analyse von Lebenslangem Lernen gesamtgesellschaftlich rahmt: Die Analyse von Lern-, Bildungs- und Entscheidungsprozessen ist demnach essentielle Grundlage für politische Maßnahmen. Damit sollte ein Paradigmenwechsel von kurativer Reaktion hin zu präventiver Intervention vollzogen werden.

### **3. Bildungsprozesse, Bildungsentscheidungen und Kompetenzentwicklung in verschiedenen Lebensphasen**

Im Rahmen dieses Themenheftes fokussieren wir auf drei zentrale Lebens- bzw. Bildungsphasen: Frühkindliche Bildung, Sekundarstufe I und Hochschule.

#### **3.1 Frühkindliche Bildung**

Im Kontext des Lebenslangen Lernens wird der frühkindlichen Bildung ein großer Stellenwert eingeräumt, da gerade in dieser Lebensphase die Grundlage für weitere Lernprozesse gelegt wird (Cunha & Heckman, 2007; Lehrl, Ebert, Blaurock, Rossbach & Weinert, 2020).

#### **3.2 Sekundarstufe I**

Die Sekundarstufe I bildet einen zentralen Entwicklungszeitraum im Leben von Kindern und jungen Erwachsenen. Sie verbringen einen Großteil ihrer Zeit in Schulen, die – neben Familie und Peers – eine zentrale Sozialisationsinstanz darstellt (z.B. Baumert, Stanat & Watermann 2006) und alle folgenden Entscheidungs- und Entwicklungsprozesse maßgeblich beeinflusst. Gleichzeitig werden in der Sekundarstufe I die entwicklungspsychologisch hochrelevanten Phasen der Pubertät und Adoleszenz durchlebt, die ebenfalls Einfluss auf Outcomes haben können, die bei der Erforschung Lebenslangen Lernens von Bedeutung sind.

#### **3.3 Hochschule**

Vor dem Hintergrund der Hochschulpolitik der letzten Dekaden (z.B. Mayer, 2008), die maßgeblich auf Öffnung der Hochschulen, der Erhöhung des Anteils akademischer Abschlüsse vor allem im MINT-Bereich sowie der Verringerung sozialer Ungleichheiten beim Zugang zur tertiären Bildung fokussierte, bilden

Analysen dieses Abschnittes gesellschaftlich und bildungspolitisch sehr relevante Forschungsbereiche.

## 4. Das NEPS

Eine sehr gute Grundlage, um die genannten Lebens- und Bildungsphasen in Bezug auf Bildungsprozesse, Bildungsentscheidungen und Kompetenzerwerb untersuchen zu können, bildet die National Educational Panel Study (NEPS; Blossfeld & Roßbach, 2019). Aus den dargestellten theoretischen Prinzipien geht hervor, dass das Lebenslange Lernen als ein sehr komplexes Phänomen angesehen werden muss, da es von unzähligen individuellen Dispositionen, strukturellen Faktoren und pädagogischen Prozessen innerhalb und außerhalb von Bildungseinrichtungen beeinflusst wird. Das komplexe Gefüge der Faktoren, die das individuelle Lernen beeinflussen, zu erfassen, zu modellieren und mit anerkannten soziologischen, psychologischen, ökonomischen und pädagogischen Theorien verknüpfbar zu machen, ist ein zentrales Anliegen des NEPS. Ein weiteres Ziel ist es, bestehende Längsschnittstudien methodisch weiterzuentwickeln, indem beispielsweise zentrale, für das Lebenslange Lernen, die Teilhabe an der Gesellschaft, den Arbeitsmarkterfolg oder die langfristige Lebenszufriedenheit relevante Kompetenzen in die Erhebungen der NEPS-Studie einfließen. Das NEPS begreift sich dabei als Infrastrukturprojekt mit dem Ziel, längsschnittliche Datensätze zu generieren, die vielfältige Forschungen ermöglichen.

Um den sechs dargestellten Prinzipien Rechnung zu tragen, ist das NEPS einerseits so konzipiert, dass es in sechs Säulen Themengebiete – interdisziplinär – bündelt. Diese Themengebiete bilden Inhaltsbereiche, deren Themen in unterschiedlicher Gewichtung über die gesamte Lebensspanne von großer Relevanz sind (Blossfeld, von Maurice & Schneider, 2019). Dazu quer liegend finden sich die sogenannten Etappen – Bildungsabschnitte, die sich maßgeblich durch den Ein-, Über- und Austritt in bzw. aus formalisierten Bildungskontexten konstituieren – stets anerkennend, dass neben der formalisierten Bildung immer auch non-formale und informelle Bildungsprozesse stattfinden (ebd.).

Das NEPS und daraus hervorgehende Daten bilden die Grundlage dieses Themenschwerpunktes. Dabei fokussieren wir auf zwei übergeordnete Ziele:

- a) Die Darstellung aktueller, altersgruppenspezifischer Fragestellungen und Forschungsergebnisse aus drei zentralen Phasen des Lebenslangen Lernens im Kontext der Bildungsinstitutionen Kindergarten, Schule und Hochschule sowie
- b) einen Einblick in die interdisziplinäre, multiperspektivische Forschung mit verschiedenen Datensätzen des NEPS.

## 5. Übersicht über die Beiträge

Im frühkindlichen Bereich befasst sich die Autorinnengruppe von *Anja Linberg, Manja Attig und Sabine Weinert* anhand einer Teilstichprobe der NEPS-Startkohorte 1 mit einem entwicklungspsychologischen Themenfeld. Sie bauen dabei auf Befunden auf, die eine durch einen unterschiedlichen sozioökonomischen Hintergrund inzidierte Variation in den Sprachkompetenzen von etwa dreijährigen Kindern berichten. Ein wichtiges Ziel ihres Beitrags ist es herauszuarbeiten, ob ein sprachanregendes Interaktionsverhalten der Eltern diesen Einfluss medierte. Mit Hilfe von Pfadanalysen können sie zeigen, dass mit einer höheren Bildung der Mütter auch ein höheres Interaktionsverhalten einhergeht, wenngleich das Interaktionsverhalten nur 9 % der Effekte des mütterlichen Bildungsniveaus auf den kindlichen Wortschatzerwerb vermittelt.

Für den Zeitraum der Sekundarstufe I liegen mehrere Beiträge mit wichtigen Themenschwerpunkten vor. Diese Beiträge nutzen allesamt Daten der Startkohorte 3.

*Michaela Sixt und Christian Aßmann* befassen sich damit, ob die regionale Schulinfrastruktur sowie die lokalen Arbeitsmarktbedingungen den Übergang auf ein Gymnasium beeinflussen. Vor dem Hintergrund der immer stärkeren Urbanisierung und der vorliegenden Befunde, dass Kinder in ländlichen Gebieten beim Übergang in vermeintlich höherwertige Schulformen systematisch gegenüber Kindern in städtischen Gebieten benachteiligt sind, bildet dieses Forschungsthema einen relevanten Beitrag für die Genese bildungspolitischer Entscheidungen. Sixt und Aßmann nutzen methodische Möglichkeiten des NEPS, die es in vorherigen Forschungsarbeiten nicht gab. So kommen zentrale Kontrollvariablen von Schüler\*innen (wie Kompetenzen, Schulnoten, Laufbahnempfehlungen und Aspirationen) zum Einsatz. Sie verknüpfen diese mit Regionaldaten und spezifizieren Analysen mittels hierarchischer logistischer Regressionen. So können sie herausarbeiten, dass das Vorhandensein vieler Gymnasien in der Region den Übergang auf ein Gymnasium begünstigt – unabhängig von individuellen Faktoren der Lernenden. Gleichzeitig zeigt die Arbeitsmarktsituation keinen Effekt.

*Florian Wohlkinger und Michael Bayer* stellen in ihrem Beitrag eine soziologische Forschungsfrage in den Vordergrund, indem sie eine quantitative Operationalisierung des Schülerhabitus vornehmen. Sie stützen sich hierbei auf qualitative Vorarbeiten von Kramer, Helsper, Thiersch und Ziems (2009). Grundlage ihrer Analysen sind Daten der Startkohorte 3, fokussiert auf Schüler\*innen der Jahrgangsstufen 5 und 6, die sich an Regelschulen der Sekundarstufe I befinden. Mittels latenter Klassenanalysen können sie vier abgrenzbare Klassen finden, von denen sich die Ausprägungen dreier der Klassen klar voneinander unterscheiden lassen. Ihnen ist es damit gelungen, eine Brücke zu den detailreichen, qualitativ gewonnenen Typenbeschreibungen hin zu einer in die Breite gehenden Typisierung mit der Möglichkeit zur Beschreibung von Verteilungen herzustellen.

*Jennifer Lorenz und Tobias C. Stubbe* greifen eine weitere relevante soziologische Fragestellung auf. Im Rahmen von Bildungsverläufen und -entscheidungen steht regelmäßig die „rational choice theory“ (bspw. Esser, 1999) als ein Erklärungsansatz im Fokus. Lorenz und Stubbe stellen sich die Frage, ob die soziale Herkunft und insbesondere das Motiv des Statuserhalts dazu führt, dass Kinder Nachhilfeunterricht in Anspruch nehmen. Die Analysen zeigen, dass die Inanspruchnahme von Nachhilfe nicht mit dem ökonomischen oder kulturellen Kapital in Verbindung steht, wohl aber mit dem sozialen Kapital. Gleichzeitig bilden die Schulform und die schulischen Leistungen zentrale Einflussgrößen. Das Motiv des Statuserhalts wiederum spielt keine Rolle, wohl aber der Wunsch der Eltern, den Status ihrer Kinder ihrem eigenen Status gegenüber zu erhöhen.

*Ruben Kleinkorres, Justine Stang und Nele McElvany* fokussieren mit einer psychologischen Perspektive das Thema des Wohlbefindens von Schüler\*innen und dessen Verbindung mit schulischen Leistungen und stellen die Frage, ob reziproke Effekte zwischen körperlichem, kognitiven und emotionalem Wohlbefinden und akademischer Leistung bestehen. Sie analysieren dazu ebenfalls Daten der NEPS-Startkohorte 3. Die Autor\*innen decken dabei hauptsächlich positive Zusammenhänge zwischen verschiedenen, in cross-lagged panel-Modellen integrierten Facetten von Wohlbefinden und akademischer Leistung auf.

*Karin Gehrer und Lena Nusser* legen einen Fokus auf Unterrichtsprozesse und gehen in ihrem Beitrag auf den Einfluss von Binnendifferenzierung auf die Lesekompetenzentwicklung in der Sekundarstufe I ein. Sie zeigen dazu mit Daten der NEPS-Startkohorte 3 zunächst auf, inwieweit binnendifferenzierter Unterricht umgesetzt wird und schließen die Frage an, ob binnendifferenzierende Unterrichtsmethoden Einfluss auf die Kompetenzentwicklung haben. Zuletzt beantworten sie die Frage, ob Heterogenität – die auf verschiedene Weise, bspw. in Form von Leistungs- oder kultureller Heterogenität vorliegen kann – einen Einfluss auf die Anwendung von Binnendifferenzierung und ihre Wirkung hat. Sie kommen zu dem Schluss, dass die Binnendifferenzierung im Verlauf der Sekundarstufe I abnimmt. Einen (positiven) Einfluss auf die Leistungsentwicklung können sie nur für den Beginn der Sekundarstufe I finden. Das Vorhandensein einer heterogenen Schülerschaft hat wiederum keinen Einfluss auf den Einsatz von Binnendifferenzierung.

Das Schwerpunkttheft schließt mit einem Beitrag aus dem Hochschulbildungskontext von *Divan Mouton, Hui Zhang und Bernhard Ertl*, der Daten der NEPS-Startkohorte 5 nutzt. Die Autorengruppe befasst sich mit dem wichtigen Thema des Studienabbruchs. Zwar gibt es einige Forschungsarbeiten, die das Vorliegen einer großen Zahl an Faktoren, die einen Studienabbruch begünstigen, aufzeigen. Die Frage danach, wie diese Faktoren jedoch zusammenspielen, wird bisher nur in wenigen Arbeiten aufgenommen. Der von Mouton und Kollegen analysierte Datensatz aus der NEPS-Startkohorte 5 bietet die Möglichkeit, auf Basis der herausgearbeiteten Faktoren Typisierungen von Abbrecher\*innen vorzunehmen und so empirisch aufzuzeigen, welche Faktorkombinationen besonders häufig auftreten. Sie nutzen dafür Latent-Class-Analysen, mittels derer sie 12 spezifische Profile herausarbeiten

können. Durch das Herausarbeiten dieser Profile besteht die Möglichkeit, perspektivisch zielgerichtete Interventionsstrategien zu entwickeln.

Mit diesen sieben Beiträge gibt das Themenheft „Bildungsprozesse, Bildungsentscheidungen und Kompetenzentwicklungen in verschiedenen Lebensphasen – Aktuelle Forschungsbefunde aus NEPS“ des Journal for Educational Research Online Einblick in aktuelle Forschungsarbeiten im Bereich der *frühkindlichen Bildung*, der *Sekundarstufe I* und der *Hochschule*, die im Kontext von NEPS entstanden sind.

## Literatur

- Baltes P. B. (2003). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundation of developmental theory. In U. M. Staudinger & U. Lindenberger (Hrsg.), *Understanding Human Development* (S. 17–43). Boston, MA: Springer.
- Bandura, A. (1982). *Self-efficacy mechanism in human agency*. *American Psychologist*, 37, 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: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*. Wiesbaden: VS.
- Becker, G. S. (1964). *Human capital*. New York, NY: National Bureau of Economic Research.
- Behrmann, J. R. (2003). An economic perspective on selection, optimization, and compensation (SOC). In U.M. Staudinger & U. Lindenberger (Hrsg.), *Understanding Human Development* (S. 125–155). Boston, MA: Springer.
- Block, J. (1982). Assimilation, accomodation, and the dynamics of personality development. *Child Development*, 53(2), 281–295.
- Blossfeld, H.-P. & Maurice, J. von (2019). Education as a Lifelong Process. In H.-P. Blossfeld & H.-G. Roßbach (Hrsg.), *Education as a lifelong process: The German National Educational Panel Study (NEPS)* (2. Aufl., S. 1–16). Wiesbaden: VS.
- Blossfeld, H.-P., Maurice, J. von & Schneider, T. (2019). The National Educational Panel Study: Need, main features, and research potential. In H.-P. Blossfeld & H.-G. Roßbach (Hrsg.), *Education as a lifelong process: The German National Educational Panel Study (NEPS)* (2. Aufl., S. 17–33). Wiesbaden: VS.
- Blossfeld, H.-P. & Roßbach, H.-G. (Eds.). (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS)* (2. Aufl.). Wiesbaden: VS.
- Breen, R. & Goldthorpe, J. H. (1997). Explaining educational differentials. Towards a formal rational action theory. *Rationality and Society*, 9, 275–305.
- Bronfenbrenner, U. (1981). *Die Ökologie der menschlichen Entwicklung: Natürliche und geplante Experimente*. Stuttgart: Klett-Cotta.
- Bronfenbrenner, U. (1986). Recent advances in research on the ecology of human development. In R.-K. Silbereisen, K. Eyferth & G. Rudinger (Hrsg.), *Development as action in context – Problem behavior and normal youth development* (S. 287–310). Berlin: Springer.
- Campbell, D. E. (2006). What is education's impact on civic and social engagement? In R. Desjardins & T. Schuller (Hrsg.). *Measuring the effects of education on health and civic engagement. Proceedings of the Copenhagen Symposium* (S. 25–126).

- Cunha, F. & Heckman, J. J. (2007). The technology of skill formation. *American Economic Review*, 97(2), 31–47.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95, 542–575.
- Eccles, J. S. & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. In C. Ames & R. Ames (Hrsg.), *Research on Motivation in Education. Goals and Cognitions (Volume 3)*. San Diego: Academic Press.
- Elder, G. H. Jr., Johnson, M. K. & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer & M. J. Shanahan (Hrsg.), *Handbook of the life course* (pp. 3–19). New York, NY: Springer.
- Erikson, R. & Jonsson, J. O. (1996). Explaining class inequality in education: The Swedish case. In R. Erikson & J. O. Jonsson (Hrsg.), *Can education be equalized? The Swedish case in comparative perspective* (pp. 1–63). Oxford, England: Westview Press.
- Esser, H. (1999). *Soziologie: Spezielle Grundlagen. Band 1: Situationslogik und Handeln*. Frankfurt a. M.: Campus.
- Filipp, S.-H. & Aymanns, P. (2010). *Kritische Lebensereignisse und Lebenskrisen: Vom Umgang mit den Schattenseiten des Lebens*. Stuttgart: Kohlhammer.
- Kramer, R.-T., Helsper, W., Thiersch, S. & Ziems, C. (2009). *Selektion und Schulkarriere: Kindliche Orientierungsrahmen beim Übergang in die Sekundarstufe I*. Wiesbaden: VS.
- Lehrl, S., Ebert, S., Blaurock, S., Rossbach, H.-G. & Weinert, S. (2020). Long-term and domain-specific relations between the early years home learning environment and students' academic outcomes in secondary school. *School Effectiveness and School Improvement*, 31(1), 102–124.
- Maaz, K., Hausen, C., McElvany, N. & Baumert, J. (2006). Stichwort: Übergänge im Bildungssystem. *Zeitschrift für Erziehungswissenschaften*, 9, 299–327.
- Mayer, K. U. (2003). The sociology of the life course and lifespan psychology: Diverging or converging pathways?. In U. M. Staudinger & U. Lindenberger (Hrsg.), *Understanding Human Development* (463–481). Boston, MA: Springer.
- Mayer, K. U. (2008). Das Hochschulwesen. In K. S. Cortina, J. Baumert, A. Leschinsky, L. Trommer & K. U. Mayer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland: Strukturen und Entwicklungen im Überblick* (S. 599–645). Reinbek: Rowohlt.
- Mayer, K. U. (2009). New directions in life course research. *Annual Review of Sociology*, 35, 413–433.
- McElvany, N., Gebauer, M. M. & Gräsel, C. (2019). Empirische Bildungsforschung – eine Standortbestimmung. Special Issue Editorial. *Journal for Educational Research Online*, 11(1), 5–8.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81.

Anja Linberg, Manja Attig & Sabine Weinert

## **Social disparities in the vocabulary of 2-year-old children and the mediating effect of language-stimulating interaction behavior**

### **Abstract**

*It is well documented that children's language skills already vary by socioeconomic status by the time they are about 2-3 years old. In addition, study results demonstrate that the frequency of language-stimulating interaction behavior – as an important aspect of the proximal familial learning environment – significantly predicts children's later language skills. However, it is less clear how early social disparities in children's language emerge and to what extent they are mediated by maternal language-stimulating interaction behavior. Therefore, we investigated disparities in children's vocabulary at 26 months of age and analyzed whether and to what extent the effect of maternal education on vocabulary acquisition is mediated through language-related episodes of joint attention, as a facet of language-stimulating interaction behavior. We used data from N = 778 children of the Infant Cohort Study of the German National Educational Panel Study.*

*The results show early disparities in children's vocabulary development as well as in language stimulating interaction behavior in mother-child interaction according to maternal education; however, only 9 % of the effect of maternal education on early vocabulary development is mediated through the assessed language stimulating interaction behavior when controlling for child language skills at 14-17 months of age.*

### **Keywords**

*Early childhood; Social disparities; ELFRA; Joint attention; Vocabulary*

---

Dr. Anja Linberg (corresponding author), German Youth Institute, Nockherstraße 2, 81541 Munich, Germany  
e-mail: linberg@dji.de

Dr. Manja Attig, Leibniz Institute for Educational Trajectories, Wilhelmsplatz 3, 96047 Bamberg, Germany  
e-mail: manja.attig@lifbi.de

Prof. Dr. Sabine Weinert, Department Psychology I – Developmental Psychology, University of Bamberg, Markusplatz 3, 96050 Bamberg, Germany  
e-mail: sabine.weinert@uni-bamberg.de

# Soziale Disparitäten im kindlichen Wortschatz von zweijährigen Kindern und ihre Vermittlung durch sprachanregendes Interaktionsverhalten

## Zusammenfassung

*Es ist breit dokumentiert, dass Sprachkompetenzen von Kindern bereits im Alter von etwa 3 Jahren in Abhängigkeit vom sozioökonomischen Hintergrund variieren. Gleichzeitig zeigen Befunde, dass die Häufigkeit des sprachanregenden Interaktionsverhaltens als Aspekt der proximalen familiären Lernumgebung bedeutsam mit dem Spracherwerb der Kinder zusammenhängt. Weniger klar ist jedoch, wie entsprechende Unterschiede in den frühen sprachlichen Fähigkeiten und Fertigkeiten von Kindern entstehen und ob bzw. inwieweit sie durch mütterliches sprachanregendes Interaktionsverhalten vermittelt werden. Daher untersuchen wir Unterschiede im kindlichen Wortschatz im Alter von 26 Monaten und analysieren, inwieweit Effekte der mütterlichen Bildung auf den Wortschatz ihrer Kinder durch sprachrelevante Episoden geteilter Aufmerksamkeit (joint attention), als eine Facette sprachfördernden Interaktionsverhaltens, vermittelt werden. Hierzu wurden Daten von N = 778 Kindern der Säuglingskohorte des Nationalen Bildungspanels genutzt.*

*Die Ergebnisse unserer Analysen zeigen sowohl frühe Unterschiede im kindlichen Wortschatzerwerb als auch im sprachstimulierenden Interaktionsverhalten im Zusammenhang mit der mütterlichen Bildung; dabei werden aber nur 9 % der Effekte der mütterlichen Bildung auf den kindlichen Wortschatzerwerb durch das erfassste sprachstimulierende Interaktionsverhalten vermittelt, wenn der kindliche Sprachstand im Alter von 14 resp. 17 Monaten kontrolliert wird.*

## Schlagworte

*Frühe Kindheit; Soziale Disparitäten; ELFRA; Geteilte Aufmerksamkeit; Wortschatz*

## 1. Introduction

Indisputably, the early years of a child's life are important for their later educational development (Belsky et al., 2007). Following the results of previous research, language development plays a key role for later development as well as for school readiness, reading skills, and school success (NICHD Early Child Care Research Network, 2005; Rowe, Raudenbush, & Goldin-Meadow, 2012; Bates, Dale, & Thal, 1996; Weinert & Ebert, 2017). It has even been suggested that language skills in early childhood are the best predictor of school readiness and later school success (Hoff, 2013). Having said this, it seems even more challenging that children's expressive and receptive language varies by socioeconomic background so early in life (Hart & Risley, 1995; Linberg & Wenz, 2017; Fernald, Marchman, & Weisleder,

2013; Hoff, 2006; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010). Already at the age of 2 years, children's productive vocabulary was found to be associated with the socioeconomic status (SES) of the family (Attig & Weinert, 2019; Fernald et al., 2013). Moreover, these disparities seem to be rather stable across preschool age (Weinert & Ebert, 2013).

In various developmental and educational models, such as the bio-ecological model (Bronfenbrenner & Morris, 2006) or the model of educational quality (Roßbach, 2005), these SES-related disparities are explained by differences in educational processes within learning environments associated with the different socioeconomic and educational characteristics of the family. In particular, these models suggest that socioeconomic and cultural resources may influence the prospects of providing stimulating educational processes for the child, which in turn impacts the development of child competencies. In these models, educational processes are seen as a primary engine of competence development, which serve as a mediator between socioeconomic characteristics and child outcomes (Bronfenbrenner & Morris, 2006; Roßbach, 2005). During the first years of life, it is suggested that the interactions between parent and child are particularly important for educational processes within these learning environments (Bornstein, 2002). In line with that assumption, research has shown that, for example, the global quality of interaction behavior (e.g. sensitivity) as well as joint activities are associated with language development in childhood (Olson, Bayles, & Bates, 1986; Tamis-LeMonda, Bornstein, Baumwell, & Damast, 1996; Tamis-LeMonda, Bornstein, Kahana-Kalman, Baumwell, & Cyphers, 1998). Further, these educational processes vary according to the socioeconomic and educational characteristics of the family (Gudmundson, 2012; Weinert & Ebert, 2017).

Although the interrelation between social disparities, language development, and the role of the quality of interaction behavior is well documented, the mediation, which is assumed in developmental and educational models, has rarely been tested explicitly (Tudge, Mokrova, Hatfield, & Karnik, 2009). Thus, it is still an open question to what extent early social disparities in the language development of children (i.e. disparities in child language that are associated with differences in socioeconomic family background) are mediated by the quality of interaction behavior. However, and despite some general interrelations, from a theoretical as well as from a practical point of view, it is important to specify the general concept of quality of interaction behavior. In fact, which facet of interaction behavior promotes language acquisition may differ according to the phase of language acquisition as well as the respective language component under study (e.g. Vallotton, Mastergeorge, Foster, Decker, & Ayoub, 2017). Hence, the aim of present paper is to investigate the mediating role of a specific facet of early language-stimulating interaction behavior which is suggested to be especially relevant to vocabulary acquisition in the second year of life (Baldwin, 1995; Farrant & Zubrick, 2012), namely the frequency of language-related joint attention episodes.

## 2. Social disparities and language development

Differences in socioeconomic background are already shown in the language skills of 3-year-old children (Hart & Risley, 1999; Weinert & Ebert, 2013). Hoff (2013) stated that children from families with low socioeconomic status showed lower levels of oral language skills than children from more advantaged families (see also Fernald et al., 2013; Hoff, 2006; Huttenlocher et al., 2010; Law, Clegg, Rush, Roulstone, & Peters, 2019). Within the language domain, vocabulary in particular is suggested to be related to family background (for a brief overview: Vasilyeva & Waterfall, 2011). Study results by Hart and Risley (1995) point out that children growing up in families with a low SES have smaller vocabularies and more restricted language skills compared to their peers from higher SES families. Already at the age of 18 months, Fernald and colleagues (2013) documented differences in productive vocabulary and language processing related to social background. At 2 years of age children from families differing in SES already showed a 6-month gap in language processing skills important to language development (Fernald et al., 2013). However, the data patterns are somewhat controversial as Peyre et al. (2014) found that parental education predicts the language skills of their children at the age of 3 years but not the child's vocabulary at 2 years of age.

Yet, as Bornstein and Bradley (2010), amongst others, pointed out, socioeconomic background is a multi-dimensional construct and, for instance, occupation, income and education may have differential influences. Especially in the first years of a child's life, as mothers often withdraw from the labor market, occupation and income are extremely volatile compared to the rather stable dimension of education (Bornstein, Hahn, Suwalsky, & Haynes, 2010). Study results focusing on the latter dimension demonstrate an association of maternal education and child language. For example, using the data of the infant cohort study of the German National Educational Panel Study (NEPS), Attig and Weinert (2019) showed that maternal education already predicted the language level (a combination of productive vocabulary and grammar) at the age of two. Further, Law, Rush, Anandan, Cox, and Wood (2012) also found that maternal education predicted the productive vocabulary of 5 year olds as well as the change in vocabulary skills from 3-5 years.

## 3. Language-stimulating behavior in mother-child interaction and language development

Sylva, Melhuish, Sammons, Siraj-Blatchford, and Taggart (2008) claimed that, in terms of child development, what parents are doing with their children in their daily life is more important than their socioeconomic background. In line with this assumption, studies suggest that social disparities in children's language development (i.e. disparities in language related to social background) mainly reflect differences in experiences and educational processes, not just individual differences in

basic abilities or capacities – at least in typically developing children (Dale et al., 1998; Hoff, 2003; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002).

Studies focusing on the global quality of educational processes demonstrate that the quality of interaction behavior is associated with the language skills of children (Baumwell, Tamis-LeMonda, & Bornstein, 1997; Nozadi et al., 2013; Pungello, Irufka, Dotterer, Mills-Koonce, & Reznik, 2009). In addition, there is evidence that sensitive parenting behavior as indicated by a prompt, contingent, adequate and warm reaction to the child's needs and signals predicts language outcome (Tamis-LeMonda, Briggs, McClowry, & Snow, 2009; Tamis-LeMonda et al., 1996; Tamis-LeMonda et al., 1998). More specifically, it was shown that children of more sensitive mothers begin to talk earlier and reach the milestone of a 50-word productive vocabulary at a comparatively younger age than children of less responsive mothers (Tamis-LeMonda et al., 1996; Tamis-LeMonda et al., 1998). Attig and Weinert (2019) added that mothers' sensitive and stimulating behavior as well as the frequency of early joint picture book reading predicted language skills in 2 year olds with all three predictors explaining separated amounts of variance.

However, via an animated discussion it is suggested that indicators aiming to measure the global quality of educational processes might underestimate the associations with specific domains, such as language (Anders et al., 2012; Lehrl, Ebert, Blaurock, Roßbach, & Weinert, 2019; Mashburn et al., 2008). In addition, within global measures of interaction quality various specific behaviors that may or may not be relevant are included. Therefore, it seems important to focus on more specific processes that are suggested to stimulate the development of specific domains such as language. Yet, as already mentioned, which facet of interaction behavior is most relevant might depend on the phase of language development and the language component under study (e.g. Weinert & Grimm, 2018 for an overview). In fact, with respect to language development from the early years onwards, various interactional variables have been shown to impact developmental progress, ranging from stimulating verbal interactions including questioning and verbal distancing as well as language teaching strategies such as repeating, expanding, and reformulating the child's utterances (Hoff-Ginsberg & Shatz, 1982) and the quantity and quality of picture book exposure to specific features of speech input such as the complexity of adult talk to children (Anderka, 2018; Huttenlocher et al., 2002) or the diversity of word used (Weinert & Lockl, 2008 for an overview) with partially differentiated effects on child vocabulary (Lehrl, Ebert, Roßbach, & Weinert, 2012; Vasilyeva & Waterfall, 2011). With respect to the second year of life, it is suggested that language-related joint attention episodes, amongst others, may be particularly relevant to a child's language development and early word learning (e.g. Baldwin, 1995; Tomasello & Farrar, 1986; Bruner, 1978), as within those episodes children are more easily able to associate words with specific objects or scenes and thus to unravel the meaning of new words. In line with this assumption, study results show that vocabulary growth is predicted by the time mother-child dyads spend in joint attention (Carpenter, Nagell, & Tomasello, 1998; Laakso, Poikkeus, Katajamaki, & Lyytinen, 1999) as well as by maternal input that follows the child's

attentional focus (Akhtar, Dunham, & Dunham, 1991; Tomasello & Farrar, 1986). Further, language skills develop more rapidly if children are advanced in achieving joint attentional states (Carpenter et al., 1998; Mundy & Gomes, 1998).

#### 4. Social disparities and language-stimulating interaction behavior

In bio-ecological and educational quality models the home learning environment is seen as a multidimensional construct. According to the models, the SES of the family influence educational processes which in turn impacts the development of the child. Hence, the models suspect a mediation of the effect of SES on child outcome via educational processes (Bronfenbrenner & Morris, 2006; Roßbach, 2005). Following this idea, the above mentioned SES-related disparities in language acquisition might be due to related differences in interaction quality. Research not only shows differences in child development according to SES but also with respect to the interaction quality and the language use of parents (Hoff, Laursen, & Tardif, 2002). Thus, associated with the socioeconomic status of the family and particularly the level of education of the mother, various differences have been documented in the literature. For instance, the global quality of the home learning environment (HLE; as, for example, measured by the Home Observation for Measurement of the Environment Inventory; Bradley & Caldwell, 1979) has been shown to be associated with the education level of the family (Bradley, Corwyn, McAdoo, & Coll, 2001; Lugo-Gil & Tamis-LeMonda, 2008; Magnuson, Sexton, Davis-Kean, & Huston, 2009). Further, compared to mothers with lower education, higher educated mothers showed a higher global quality of interaction behavior (Bradley et al., 2001; Gudmundson, 2012; Attig & Weinert, 2018). In addition, language-stimulating interaction behavior has been shown to differ according to the families' education level (Hoff, 2003). In particular, the amount of talking as well as the qualitative and interactive aspects of mothers' speech differ according to her education level. Higher educated mothers talk more to their children and their language input to their children is suggested to be more supportive for language development compared to less educated mothers (Hart & Risley, 1995; Hoff, 2003, 2013; Huttenlocher, Vasilyeva, Waterfall, Vevea, & Hedges, 2007). Rowe (2012) found that more highly educated parents used more word tokens and word types as well as more rare words and more decontextualized utterances at 18 months, 30 months and 42 months of age. Even more, mothers also differ in the purpose of language use depending on their socioeconomic status (e.g., in directing children's behavior or eliciting and maintaining conversation; Hoff, 2006; Hoff et al., 2002).

However the findings are at least partially controversial. While Hart and Risley (1992, 1995) extrapolated their findings to the famous statement that, at the age of 3 years, children from low SES families hear 30 million words less compared to children from high SES families, others could not replicate this finding (e.g. Sperry,

Sperry, & Miller, 2018; see Golinkoff, Hoff, Rowe, Tamis-LeMonda, & Hirsh-Pasek, 2019; Sperry, Sperry, & Miller, 2019 for a critical discussion of the findings). Further, SES-related disparities in grammar seem to be mediated by partially different facets of interactive communications in contrast to disparities in vocabulary (Anderka, 2018).

With respect to SES indicators, as already mentioned, it is claimed that maternal education is especially relevant for stimulating interactions with very young children (Duncan & Magnuson, 2010) as, amongst others, mothers with higher education are found to gather more information on parenting and child development, which might impact their interaction behavior (Bornstein, Cote, Haynes, Hahn, & Park, 2010; Fagan, 2017).

## 5. Present study

To sum up, language-stimulating processes have been shown to impact vocabulary development, which unfold on the background of the socioeconomic resources of the family. In addition, during the first years of a child's life, maternal education seems to be a key dimension of socioeconomic background.

In bio-ecological theories (Bronfenbrenner & Morris, 2006) or theories of educational quality (Roßbach, 2005) it is often assumed that structural characteristics, such as socioeconomic resources, exert an influence on child development through educational processes. Although empirical results show that maternal education is associated with the language status of 2-year-old children (Attig & Weinert, 2019), the mediation is rarely tested (Tudge et al., 2009). In addition, this paper takes a domain-specific developmental approach, assuming that specific facets of stimulation might be particularly relevant to domain-specific developmental tasks.

Thus, we aim to investigate whether domain-specific language-stimulating interaction behavior during mother-child interaction and particularly the frequency of language-related joint attention episodes (assessed by a combined time- and event-sampling coding approach), might mediate the effect of maternal education on the development of early productive vocabulary.

### Research Questions:

- 1) To what extent is the early development of vocabulary (namely productive vocabulary at age 2 when controlling for earlier language skills) related to maternal education?
- 2) To what extent is the frequency of language-related joint attention episodes as language-stimulating interaction behavior related to maternal education in the second year of life (when controlling for the child's language status which might affect mothers' behavior)?
- 3) Does the assessed early language-stimulating interaction behavior mediate the effect of maternal education on vocabulary development in 2 year old children?

## 6. Methods and data

### 6.1 Sample

We used data of the Starting Cohort 1 – Newborns of the NEPS (NEPS; Blossfeld & Roßbach, 2019; Weinert, Linberg, Attig, Freund, & Linberg, 2016) of the first three waves to address these research questions. In wave 1 children were 7 months old; in wave 2 parent interviews took place when children were 14 months and the assessment of interaction behavior when they were 17 months old; in wave 3 children were 26 months of age. Starting Cohort 1 is a representatively drawn sample of around 3,500 children born between February and June 2012 and their families, who were followed up longitudinally (Weinert et al., 2016).

In the second wave – by design – only half of the sample was assessed by video-observational measures, including the measure for language-stimulating interaction behavior. In our analyses we used all cases in which valid information on child's vocabulary at wave 3 (26 months) and valid information on language stimulating interaction behavior in wave 2 (17 months) were available, but excluded cases in which another language than German was exclusively the interaction language in the household, which resulted in a total sample size of 778 children.

### 6.2 Vocabulary at 26 months

Child's *vocabulary at 26 months* was assessed by the ELFRA-2, a standardized parental report on child language including a vocabulary check-list of 260 words and phrases the child actively uses (ELFRA; Grimm & Doil, 2006). This instrument is comparable to the internationally well-known and extensively validated "MacArthur-Bates Communicative Development Inventories (Toddler Form) – CDI" (Fenson et al., 1993; for the validity of the ELFRA see Sachse, Anke, & von Suchodoletz, 2007).

### 6.3 Language-stimulating interaction behavior at 17 months

As an indicator for language stimulating interaction behavior we used information on language-related joint attention episodes at the age of 17 months, which was based on videos of a semi-standardized interaction situation between mother and child playing in the home-setting with a standardized toy set for 10 minutes in German (see Linberg et al., 2019 for detailed information on design and set-up). Mothers were asked to play with the child as she would normally do when having time and being alone with the child. These videos were coded with a combined

time- and event-sampling approach using 15 second intervals (see Linberg, 2018).<sup>1</sup> For each of the 15 second time intervals the raters coded whether the mother was engaging in a joint attention episode [0: no; 1: yes] with the child, indicated by attention of both interaction partners on the same object / event, and whether any language was used by the mother (excluding fillers such as “mh”) [0: no; 1: yes] (Linberg, 2018).

If both aspects were coded with yes, the interval was coded as an interval in which language-related joint attention episodes were demonstrated as an indicator of language-stimulating behavior (interrater-agreement = 77 %). In our analyses, we used the percentage of all possible 40 intervals in which the respective indicator occur.

## 6.4 Maternal education

As an SES characteristic we considered maternal education at wave 2 (at the age of 14 months) as this is a rather stable SES characteristic (compared to income which, due to maternal leave regulations, is rather volatile in the first years of a child’s life; (Bornstein & Bradley, 2010) and considered to particularly relevant to the research question under study (e.g. Law et al., 2012). We used the CASMIN-classification (König, Lüttinger, & Müller, 1988), an indicator for maternal education, containing information on school and vocational training certificates, ranging from 1-8 with [1] indicating general elementary education, [4] secondary school leaving certificate with vocational training and [8] higher tertiary education (university degree).

## 6.5 Controls

In order to control for children’s earlier language skills, we included *verbal expressions* of the child as coded based on the videotaped mother-child interaction situation at 17 months of age as well as a rough proxy for the child’s early *productive vocabulary* assessed in the parent interview at 14 months of age when parents were asked how many people or things the child names correctly, ranging from [1] none to [5] more than 20. For child’s *verbal expressions*, raters indicated (off-line coding) whether the child has spoken one- or two-word sentences during a time interval (interrater-agreement = 79%) and we used the percentage of all intervals coded ‘yes’ as an indicator.

We also included two additional SES indicators from wave 2 in our analyses: Weighted *household income* (OECD, 2013) as an indicator of financial resources

---

<sup>1</sup> These codings were conducted within the ViVA-project (Video-based Validity Analyses of Measures of Early Childhood Competencies and Home Learning Environment) funded by the German Research Foundation (DFG; grant to S. Weinert and H.-G. Roßbach) within the priority program 1646.

and the variable *single parent*, which is defined by a partner living in household [0] no [1] yes, as an indicator of social resources, as for both indicators associations to child language as well as to the home learning environment have been demonstrated (e.g. Linberg & Wenz, 2017; Sylva et al., 2008; Huttenlocher et al., 2010).

Furthermore, we controlled for child's *age* at wave 2 and 3, and whether the *child is a boy* [0] no [1] yes and the language spoken in the household assessed in wave 1 (1 = *only German*, 2 = *mostly German*, 3 = *mostly another language*). Descriptives for all variables are shown in Table 1.

**Table 1:** Descriptives

	N	M / %	SD	Min	Max
Child's Vocabulary <sub>26</sub>	778	147.12	62.06	2	260
Child's verbal expressions <sub>17</sub>	778	21.71	18.58	0	90
Child's vocabulary <sub>14</sub>	770	2.16	0.85	1	5
Maternal education	769	6.17	1.85	0	8
Household income	753	1750.00	796.21	100	6250
Single parent (0 = no 1 = yes)	770	4.55 %		0	1
Language-stimulating interaction behavior	778	67.06	16.38	0	97.5
Child's age <sub>17</sub> (in months)	778	17.03	0.67	16	20
Child's age <sub>26</sub> (in months)	778	26.38	1.14	24	32
Child is a boy (0=no 1=yes)	778	51.93 %		0	1
Household language (1 = German; 2 = sometimes other; 3 = mostly other)	778	1.15	0.42	1	3

## 6.6 Analytic strategy

For analysing disparities according to maternal education (a) in child's vocabulary at the age of 26 months as well as (b) in the frequency of joint attention episodes as an indicator for the language-stimulating interaction behavior the child experienced at the age of 17 months we used path models (with correlated covariates) and controlled for child's age in the wave of interest, child's gender, household language, household income, and single parent. Moreover, as language-stimulating interaction behavior (in wave 2) might be a reaction to the child's actual language skills, we additionally controlled for children's language skills in the second year of life by including the child's verbal expressions in the interaction situation and vocabulary at wave 2 (see correlations in the appendix between the child's verbal expressions and the language-stimulating interaction behavior at 17 months of age:  $r = .24$ ,  $p < .001$ ). Note, that by doing so, we more clearly address language development instead of just effects on language status (correlations between variables (see appendix) reveal some stability in individual differences in early language development ( $r = .42$ ,  $p < .001$ ;  $r = .29$ ,  $p < .001$ ) in the second year of life). For test-

ing the mediation of the effect of maternal education through language-stimulating interaction behavior, we used path models, in which direct, indirect and total effects for each predictor were estimated and indicated by the estimated Bentler-Raykov  $R^2$ . Missing variables (max. 3 %) were treated using full-information-maximum-likelihood estimation (Acock, 2013) using all valid information of all observations with Stata 15.

## 7. Results

### 7.1 Disparities in child's vocabulary

Regarding SES-related disparities in children's vocabulary, our bivariate analyses (see Table A1) demonstrate no or only small significant associations between maternal education and our proxies for child language skills and vocabulary at 17 and 14 months respectively. However, significant bivariate associations are observable at 26 months (see Table A1), which are even apparent in path models when controlling for earlier language skills. This association remains significant when other SES characteristics are additionally included in the model (Table 2). In fact, the effect of maternal education stayed rather stable when the model controlled for income and family status (single parent). However, only 2 % of variance in child's vocabulary development is explained by maternal education.

**Table 2:** Separate prediction of child's vocabulary at 26 months by maternal education with control variables

	Vocabulary <sub>26</sub>	Vocabulary <sub>26</sub>	Vocabulary <sub>26</sub>
Maternal education		.14***	.11**
<i>Control variables</i>			
Child's age <sub>26</sub>	.21***	.20***	.20***
Child is a boy	-.11***	-.11***	-.10**
Child's verbal expressions <sub>17</sub>	.33***	.32***	.32***
Child's vocabulary <sub>14</sub>	.20***	.21***	.21***
Household language	-.10**	-.08**	-.07*
Household income			.07*
Single parent			-.03
<i>R</i> <sup>2</sup>	.28	.30	.30

*Notes.* N = 778. Standardized coefficients.

\*p < .10. \*p < .05. \*\*p < .01. \*\*\*p < .001.

## 7.2 Disparities in language-stimulating interaction behavior

In the next step, SES-related disparities in the frequency of language-related joint attention episodes as a dimension of language-stimulating interaction behavior during the semi-standardized play situation at 17 months of age was analyzed. Again, the observed disparities in language-stimulating interaction behavior according to maternal education are significant in bivariate analyses ( $r = .16$ ,  $p < .001$ ; see Table A1) and also remain significant when controlling for concurrent child language and other child and SES characteristics (Table 3).

**Table 3:** Separate predictions of language-stimulating interaction behavior at the age of 17 months with control variables

	Language-stimulating interaction behavior <sub>17</sub>	Language-stimulating interaction behavior <sub>17</sub>	Language-stimulating interaction behavior <sub>17</sub>
Maternal education		.13***	.12***
<i>Control variables</i>			
Child's age <sub>17</sub> (in months)	.06	.04	.05
Child is a boy	-.12**	-.11**	-.11**
Child's verbal expressions <sub>17</sub>	.22***	.21***	.21***
Child's vocabulary <sub>14</sub>	-.01	-.01	-.01
Household language	-.09*	-.07*	-.07*
Household income			.03
Single parent			-.01
<i>R</i> <sup>2</sup>	.08	.10	.10

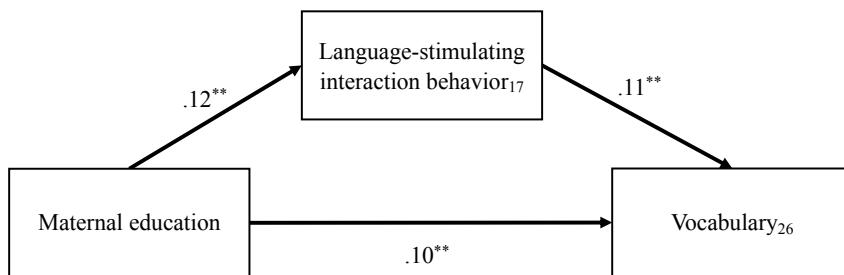
Notes.  $N = 778$ . Standardized coefficients.

\* $p < .10$ . \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

## 7.3 Mediation

In a last step we tested to what extent the effect of maternal education on child's vocabulary at wave 3 was mediated by language-stimulating interaction behavior. Using a path model, we estimated the effect of maternal education on the frequency of language-stimulating interaction behavior and child's vocabulary, as well as the effect of language-stimulating interaction behavior on child's vocabulary at wave 3 simultaneously, while controlling for child's earlier language skills and further child and SES characteristics (see Table 4). Of the total effect ( $\beta = 0.11$ ) only ( $\beta = 0.01$ ) is indirect, meaning 9 % of the effect of maternal education is mediated through language-related joint attention episodes (Table 5).

Figure 1: Simplified model of the relation of maternal education, language-stimulating interaction behavior at the age of 17 months and vocabulary at the age of 26 months



Notes. N = 778. Standardized coefficients.  $R^2_{\text{Language-stimulating interaction behavior}_{17}} = .10$ ;  $R^2_{\text{Vocabulary}_{26}} = .31$ ; controlled for child's verbal expression<sub>17</sub>, vocabulary<sub>17</sub>, age<sub>26</sub>; gender, as well as for household language and household income.

<sup>\*</sup>p < .10. <sup>\*</sup>p < .05. <sup>\*\*</sup>p < .01. <sup>\*\*\*</sup>p < .001.

Table 4: Relation of maternal education, language-stimulating interaction behavior at the age of 17 months and vocabulary at the age of 26 months

	Language-stimulating interaction behavior <sub>17</sub>	Vocabulary <sub>26</sub>
Maternal education	.12**	.10**
Language-stimulating interaction behavior <sub>17</sub>		.11**
<i>Control variables</i>		
Child's age <sub>17</sub>	.05	
Child's age <sub>26</sub>		.19***
Child is a boy	-.11**	-.09**
Child's verbal expressions <sub>17</sub>	.21***	.30***
Child's vocabulary <sub>14</sub>	-.01	.21***
Household language	-.07*	-.07*
Household income	.03	.07*
Single parent	-.01	-.03
R <sup>2</sup>	.10	.31

Notes. N = 778. Standardized coefficients.

<sup>\*</sup>p < .10. <sup>\*</sup>p < .05. <sup>\*\*</sup>p < .01. <sup>\*\*\*</sup>p < .001.

**Table 5:** Standardized effects of maternal education and language stimulating interaction behavior at the age of 17 months with direct and indirect effects

Outcome	Direct effect	Indirect effect	Total effect
Maternal education → Language-stimulating interaction behavior <sub>17</sub>	.12**	-	.12**
Maternal education → Vocabulary <sub>26</sub>	.10**	.01*	.11**
Language-stimulating interaction behavior <sub>17</sub> → Vocabulary <sub>26</sub>	.11**	-	.11**

Notes. N = 778. Standardized coefficients.

\*p < .10. \*p < .05. \*\*p < .01. \*\*\*p < .001.

## 8. Discussion

Language competencies already differ early in a child's life by socioeconomic background (e.g. Hart & Risley, 1995; Huttenlocher et al., 2010). This is rather concerning as early language skills predict later language competencies and are important for academic success and participation in society (e.g. NICHD Early Child Care Research Network, 2005; Rowe et al., 2012; Weinert & Ebert, 2017). Different theories point to the important role of educational processes, e.g. in interactions between the mothers and their children, for the "transmission" of these disparities and the emergence of SES-related differences in child development (Bronfenbrenner & Morris, 2006; Roßbach, 2005). However, it is still rather unclear to what extent domain-specific educational processes, such as language stimulating interaction behavior, serve as a mediator of the effects of socioeconomic status in early childhood.

The present paper therefore addressed three main questions: (1) whether disparities in vocabulary development at the age of 26 months by maternal education are still observable when controlling for earlier language skills and other child and SES characteristics; (2) whether SES related disparities in the frequency of episodes of joint attention as a facet of language-stimulating interaction behavior in mother-child interactions are visible even when controlling for early child language and other characteristics and (3) whether the disparities in child's vocabulary can be explained by differences in the respective language-stimulating interaction behavior, namely whether the effect of maternal education on a child's vocabulary development is mediated by the frequency of language-related joint attention episodes taking a proxy for earlier language skills into account. While the first question addresses the specificity of the association between maternal education and vocabulary development between 17 and 26 months, the second focusses on the relation between maternal education and a potentially functional interaction variable that is suggested to be particularly relevant to vocabulary acquisition in the second year of life. Both questions are the basis for the mediation model which analyzed whether these educational processes mediate the effect of maternal education on

vocabulary development in the second half of the second year of life when productive vocabulary increases substantially.

Our results demonstrate that – already at the age of 26 months – the children's productive vocabulary is significantly associated with maternal education. This association has also been reported by Attig and Weinert (2019) and is in line with previous study results that found a relationship between children's productive vocabulary and their socioeconomic family background already at two years of age (Fernald et al., 2013). The present study adds to these previous results that the association still holds true if proxies for the child's earlier vocabulary and language skills are taken into account. This result hints to the relevance of maternal education for early vocabulary development. In addition, compared to other characteristics of socioeconomic background, maternal education shows the strongest association with child's vocabulary development; however, the association is only mediocre in nature. Yet, studies focusing on older children point to an increase of SES-related disparities over time (Lehrl et al., 2012; Linberg & Wenz, 2017). For instance, Weinert and Ebert (2013) showed that the social background accounted for 6-12 % of the differences in the language skills of 3-year-old children. Overall, the data pattern of the present study hints to an early emergence of SES-related disparities in children's vocabulary during the second year of life, as – on a bivariate level – facets of SES were even less associated with the assessed proxies of early child language and vocabulary at 17 and 14 months respectively.

With respect to the second main question, our data shows a significant association between maternal education and the frequency of mothers' engagement in language-stimulating behavior in an interaction situation. A higher number of language-related joint attention episodes was found in families with higher educated mothers compared to lower educated mothers. Yet, by showing this association between maternal education and language relevant episodes of joint attention, our study results contribute to and expand existing findings on the association between SES (particularly maternal education) and various measures of general as well as domain-specific quality of educational processes. Thus, for instance, study findings by Hart and Risley (1995) also point to the assumption that children in low SES families experience significantly less language stimulation (herein the number and diversity of words) compared to children from higher SES families. But, as already mentioned, a recent study by Sperry and colleagues (2018) could not replicate this finding (Hart & Risley, 1992, 1995). Besides showing hardly any SES-related disparities in the number of words children heard, they documented a great variation in the language learning environments within the different socioeconomic stratum. Further, their results suggest that not only direct speech from the primary caregiver plays a role, but also the language input from other caregivers as well as bystander talk (see Golinkoff et al., 2019; Sperry et al., 2019 for a critical discussion of the findings). The association of maternal education and language stimulating behavior might be traced back to differences in gathering information on supportive interaction behavior (Bornstein et al., 2010; Fagan, 2017) which may impact the behavior in interaction situations. However, further studies are needed to ad-

dress this possible link explicitly. Likewise, higher education may lead to cognitive and socio-cognitive resources and competences that may allow on the one hand, to perceive and correctly interpret the child's signals comparatively better, which might be important in establishing joint attention episodes (Carpenter et al., 1998); on the other hand, these resources or competencies may help to learn more efficiently from everyday interactions with the child, to adapt to the child's needs and signals and to derive stimulating behaviors from them. Of course, we cannot rule out the possibility that personal or other factors that we did not control for may have contributed to the observed interrelation.

Further, our results support the assumption that the frequency of joint attention episodes may be relevant to vocabulary acquisition. First, we find a predictive association between the frequency of these episodes and later vocabulary; second, this association was still observable when controlling for the child's earlier language skills suggesting an effect on vocabulary development and thus a relationship that goes beyond a pure predictive association with later vocabulary status. Note that joint attention episodes may also be initiated by the child (Baldwin, 1995); in addition, the mother may react to characteristics of the child. Thus, it is particularly relevant that the associations with later child language remained significant even when controlling for various child characteristics including a proxy of early language status. Yet, controlling for early child language may at the same time lead to a rather conservative estimation of the effect of mothers' language stimulating behavior. If this behavior is rather stable across the second year of life and impacts child language already at 17 months of age, controlling for early child language leads to an underestimation of the effect which may be more pronounced from a developmental point of view that highlights cumulative effects.

Focusing on the assumed mediation of disparities, our study showed that only a very small proportion of the disparities in the children's vocabulary could be traced back to differences in the assessed indicator of language stimulating experiences. In fact, only 9 % of the effect of maternal education is mediated through the frequency of language-related episodes of joint attention. There are many aspects that could possibly account for this finding which we, however, cannot directly address and disentangle in our study. First, we measured only one specific facet of language stimulating behavior and second, we did so by drawing on a ten-minute interaction situation which may reduce differences. Third, mother's education may be associated with many factors over and above the specific interactional measure we used as well as beyond direct mother-child interactions. In fact, various variables have been shown to contribute to early child language development such as the quantity and diversity of words (Hart & Risley, 1995), but also quality as indicated by use of shared symbols and conversational fluency; (Hirsh-Pasek et al., 2015), maternal guiding language (Dave, Mastergeorge, & Olswang, 2018) as well as verbal stimulation and responsiveness (Olson et al., 1986). Thus, many aspects of language stimulation behavior seem to play a role for the developing language skills of the children and may also mediate the effect of maternal education on these skills.

Further, SES and maternal education are associated with decisions on early extrafamilial childcare, which has also been shown to be associated with children's language development (e.g., NICHD Early Child Care Research Network, 2005). Although we did not find an association of childcare attendance and child's language skills at this age in our data, we could not exclude the possible impact of the actual (language fostering) experiences within childcare on child language. Investigating the impact of the different learning environments and their possible interactions with a special focus on domain-specific processes could be part of further studies. In addition, other limitations must be mentioned.

First of all the videotaped interaction only lasted 10 minutes and therefore represents only a short insight into everyday language stimulating behavior. However, ecologic validity of the observed maternal interaction behavior seems to be high, as results from a validation study demonstrate that the interaction behavior assessed in the semi-structured play situation is comparable to interaction behavior in other situations such as natural feeding or diapering (Weinert et al., 2016; Vogel, 2020). Additionally, the measure of maternal language stimulating interaction behavior when children were 17 months old focusses on any language provided in a joint attention episode, which complies with results that show maternal input that follows the child's attentional focus to be especially relevant for language development (e.g., Akhtar et al., 1991; Tomasello & Farrar, 1986). However, we could not differentiate the exact quality of language stimulation (e.g., how or if the mother scaffolds the child's language as described by Bruner (1983) within the concept of a Language Acquisition Support System). Our coding system did not differentiate specific language structures or interactional features which have been shown to be relevant at least for children at the age of two to three years or beyond. This should be addressed in further research. It is also necessary to draw attention to the measurement of child's vocabulary, which has been obtained through parental report (extensive check-list). Obviously, report measures can potentially be influenced by differences in monitoring the child and the child's progress or even by a social desirability response bias. However, the vocabulary list is an established instrument, which shows high external validity, demonstrated by high correlations with established standardized language tests (Sachse et al., 2007). Similar vocabulary lists are used in many studies (Law et al., 2019; Morgan, Farkas, Hillemeiner, Hammer, & Maczuga, 2015; Nozadi et al., 2013) due to the difficulty in testing these young children by standardized measures in large-scale assessments. Specifically, at the age of two, children are already very active and at the same time still restricted in following verbal instructions. As a last point, the sample size should be mentioned, as only a part of the sample could be included in our analyses. Although most of the dropout of the sample is due to design, a bias in the sample could not be excluded.

The present study points out that, even in a well-educated country such as Germany, at the age of 26 months disparities in vocabulary development by maternal education are already visible and that – associated with maternal education – children experience different frequencies of language-related joint attention episodes as a facet of language stimulating interaction behavior of their mothers.

Although, this facet of language stimulating interaction behavior contributes to the child's vocabulary development, these SES-related disparities in a child's vocabulary can hardly be traced back only to differences in those experiences. Overall, our results underline the importance of reducing early disparities in a child's vocabulary from early on, potentially by intensifying parental education programs in order to strengthen the early HLE a child experiences.

## Acknowledgments

This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Newborns, doi:10.5157/NEPS: SC1:6.o.o. From 2008 to 2013, NEPS data was collected as part of the Framework Programme for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

This work was supported partially by the Deutsche Forschungsgemeinschaft (DFG) under Grant RO 820/14-2/AOBJ:615658 and WE1478/7-2/AOBJ: 615659.

## References

- Acock, A. C. (2013). *Discovering structural equation modeling using Stata*. College Station, TX: Stata Press.
- Akhtar, N., Dunham, F., & Dunham, P. J. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language*, 18, 41–50.
- Anderka, A. (2018). *Elterliches Sprachangebot und vorschulischer Spracherwerb. Eine empirische Analyse zu Zusammenhängen und sozialen Disparitäten*. Münster: Waxmann.
- Anders, Y., Rossbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., & Maurice, J. von (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly*, 27(2), 231–244.
- Attig, M., & Weinert, S. (2018). Soziale Disparitäten im Kontext von Mutter-Kind-Interaktionen und frühen Entwicklungsmaßen von Kindern. *Frühe Bildung*, 7(1), 22–31.
- Attig, M., & Weinert, S. (2019). Häusliche Lernumwelt und Spracherwerb in den ersten Lebensjahren. *Sprache · Stimme · Gehör*, 43(2), 86–92.
- Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, NJ: Lawrence Erlbaum.
- Bates, E., Dale, P. S., & Thal, D. (1996). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 96–151) Oxford: Blackwell.
- Baumwell, L., Tamis-LeMonda, C. S., & Bornstein, M. H. (1997). Maternal verbal sensitivity and child language comprehension. *Infant Behavior & Development*, 20(2), 247–258.

- Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., Owen, M. T., & The NICHD Early Child Care Research Network. (2007). Are there long-term effects of early child care? *Child Development*, 78(2), 681–701.
- Blossfeld H.-P., & Roßbach H.-G. (2019). *Education as a lifelong process. The German National Educational Panel Study (NEPS)* (2nd and revised edition). Wiesbaden, Germany: VS.
- Bornstein, M. H. (2002). Parenting infants. In M. H. Bornstein (Ed.), *Handbook of parenting. Volume 1: Children and parenting* (pp. 3–43). Mahwah, NJ: Lawrence Erlbaum.
- Bornstein, M. H., & Bradley, R. H. (Eds.). (2010). *Socioeconomic status, parenting, and child development*. Mahwah, NJ: Lawrence Erlbaum.
- Bornstein, M. H., Cote, L. R., Haynes, O. M., Hahn, C.-S., & Park, Y. (2010). Parenting knowledge: Experiential and sociodemographic factors in European American mothers of young children. *Developmental Psychology*, 46(6), 1677–1693.
- Bornstein, M. H., Hahn, C.-S., Suwalsky, J. T. D., & Haynes, O. M. (2010). Socioeconomic status, parenting, and child development. The Hollingshead Four-Factor Index of Social Status and the Socioeconomic Index of Occupations. In M. H. Bornstein & R. H. Bradley (Eds.), *Socioeconomic status, parenting, and child development* (pp. 29–82). Mahwah, NJ: Lawrence Erlbaum.
- Bradley, R. H., & Caldwell, B. M. (1979). Home observation for measurement of the environment: A revision of the preschool scale. *American Journal of Mental Deficiency*, 84(3), 235–244.
- Bradley, R. H., Corwyn, R., McAdoo, H., & Coll, C. (2001). The home environments of children in the United States. Part I: Variations by age, ethnicity, and poverty status. *Child Development*, 72(6), 1844–1867.
- Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner (Ed.), *Handbook of child psychology. Volume 1: Theoretical models of human development* (pp. 793–828). Hoboken, NJ: John Wiley.
- Bruner, J. S. (1978). The role of dialogue in language acquisition. In A. Sinclair, R. Jarvella & W. M. Levelt (Eds.), *The child's conception of language* (pp. 241–256). Berlin: Springer.
- Bruner, J. S. (1983). *Child's talk: Learning to use language*. New York, NY: Norton.
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4), i–174.
- Dale, P. S., Simonoff, E., Bishop, D. V M., Eley, T. C., Oliver, B., Price, T. S., Purcell, S., Stevenson, J., & Plomin, R. (1998). Genetic influence on language delay in two-year-old children. *Nature Neuroscience*, 1, 324–328.
- Dave, S., Mastergeorge, A. M., & Olswang, L. B. (2018). Motherese, affect, and vocabulary development: dyadic communicative interactions in infants and toddlers. *Journal of Child Language*, 45(4), 917–938.
- Duncan, G. J., & Magnuson, K. A. (2010). Off with Hollingshead: Socioeconomic resources, parenting, and child development. In M. H. Bornstein & R. H. Bradley (Eds.), *Socioeconomic status, parenting, and child development* (pp. 83–106). Mahwah, NJ: Lawrence Erlbaum.
- Fagan, J. (2017). Income and cognitive stimulation as moderators of the association between family structure and preschoolers' emerging literacy and math. *Journal of Family Issues*, 38(17), 2400–2424.
- Farrant, B. M., & Zubrick, S. R. (2012). Early vocabulary development: The importance of joint attention and parent-child book reading. *First Language*, 32(3), 343–364.
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., ... Reilly, J. S. (1993). *Technical manual for the MacArthur Communicative Development Inventories*. San Diego, CA: Singular Publishing Group.

- Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, 16(2), 234–248.
- Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. *Child Development*, 90(3), 985–992.
- Grimm, H., & Doil, H. (2006). *ELFRA-Elternfragebögen für die Früherkennung von Risikokindern* (2. überarb. Aufl.). Göttingen, Germany: Hogrefe.
- Gudmundson, J. A. (2012). *Links between maternal education and parenting quality during children's first three years: The moderating role of income and partner status*. (Doctoral dissertation), University of North Carolina, Chapel Hill.
- Hart, B., & Risley, T.R. (1992). American parenting of language-learning children: persisting differences in family child interactions observed in natural home environments. *Developmental Psychology*, 28(6), 1096–105.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes.
- Hart, B., & Risley, T. R. (Eds.). (1999). *The social world of children learning to talk*. Baltimore, MD: Paul H. Brookes.
- Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., ... Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, 26(7), 1071–1083.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74(5), 1368–1378.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55–88.
- Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Developmental Psychology*, 49(1), 4–14.
- Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In M. H. Bornstein (Ed.), *Handbook of parenting. Volume 2: Biology and ecology of parenting* (pp. 231–252). Mahwah, NJ: Lawrence Erlbaum.
- Hoff-Ginsberg, E., & Shatz, M. (1982). Linguistic input and the child's acquisition of language. *Psychological Bulletin*, 92(1), 3–26.
- Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, 45(3), 337–374.
- Huttenlocher, J., Vasilyeva, M., Waterfall, H. R., Vevea, J. L., & Hedges, L. V. (2007). The varieties of speech to children. *Developmental Psychology*, 43, 1062–1083.
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, 61(4), 343–365.
- König, W., Lüttinger, P., & Müller, W. (1988). *A comparative analysis of the development and structure of educational systems. Methodological foundations and the construction of a comparative educational scale*. Mannheim, Germany: Institut für Sozialwissenschaften.
- Laakso, M.-L., Poikkeus, A.-M., Katajamaki, J., & Lyytinen, P. (1999). Early intentional communication as a predictor of language development in young toddlers. *First Language*, 19(56), 207–231.
- Law, J., Clegg, J., Rush, R., Roulstone, S., & Peters, T. J. (2019). Association of proximal elements of social disadvantage with children's language development at 2 years: an analysis of data from the Children in Focus (CiF) sample from the ALSPAC birth cohort. *International Journal of Language & Communication Disorders*, 54(3), 362–376.

- Law, J., Rush, R., Anandan, C., Cox, M., & Wood, R. (2012). Predicting language change between 3 and 5 years and its implications for early identification. *Pediatrics*, 130(1), e132–e137.
- Lehrl, S., Ebert, S., Blaurock, S., Rossbach, H.-G., & Weinert, S. (2019). Long-term and domain-specific relations between the early years home learning environment and students' academic outcomes in secondary school. *School Effectiveness and School Improvement*, 24(2), 1–23.
- Lehrl, S., Ebert, S., Roßbach, H.-G., & Weinert, S. (2012). Die Bedeutung der familiären Lernumwelt für Vorläufer schriftsprachlicher Kompetenzen im Vorschulalter. *Zeitschrift für Familienforschung*, 24(2), 115–133.
- Linberg, A. (2018). *Interaktion zwischen Mutter und Kind. Dimensionen, Bedingungen und Effekte*. Münster, Germany: Waxmann.
- Linberg, A., Mann, D., Attig, M., Vogel, F., Weinert, S., & Roßbach, H.-G. (2019). *Assessment of interactions with the macro-analytic rating system of parent-child-interactions in the NEPS at the child's age of 7, 17, and 26 months*. Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Linberg, T., & Wenz, S. E. (2017). Ausmaß und Verteilung sozioökonomischer und migrationsspezifischer Ungleichheit im Sprachstand fünfjähriger Kindergartenkinder. *Journal for Educational Research Online*, 9(1), 77–98.
- Lugo-Gil, J., & Tamis-LeMonda, C. S. (2008). Family resources and parenting quality: Links to children's cognitive development across the first 3 years. *Child Development*, 79(4), 1065–1085.
- Magnuson, K. A., Sexton, H. R., Davis-Kean, P. E., & Huston, A. C. (2009). Increases in maternal education and young children's language skills. *Merrill-Palmer Quarterly*, 55(3), 319–350.
- Mashburn, A., Pianta, R. C., Hamre, B., Downer, J. T., Barbarin, O., Bryant, D., Burchinal, M., & Early, D. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79(3), 732–749.
- Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development*, 86(5), 1351–1370.
- Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development*, 21, 469–482.
- NICHD ECCRN – National Institute of Child Health and Human Development Early Child Care Research Network. (2005). Early child care and children's development in the primary grades. Follow-up results from the NICHD Study of Early Child Care. *American Educational Research Journal*, 42(3), 537–570.
- Nozadi, S. S., Spinrad, T. L., Eisenberg, N., Bolnick, R., Eggum-Wilkens, N. D., Smith, C. L., ... Sallquist, J. (2013). Prediction of toddlers' expressive language from maternal sensitivity and toddlers' anger expressions: A developmental perspective. *Infant Behavior and Development*, 36(4), 650–661.
- OECD – Organisation for Economic Co-operation and Development. (2013). *Framework of statistics on the distribution of household income, consumption and wealth*: OECD.
- Olson, S. L., Bayles, K., & Bates, J. E. (1986). Mother-child interaction and children's speech progress: A longitudinal study of the first two years. *Merrill-Palmer Quarterly*, 32(1), 1–20.
- Peyre, H., Bernard, J. Y., Forhan, A., Charles, M. A., De Agostini, M., Heude, B., & Ramus, F. (2014). Predicting changes in language skills between 2 and 3 years in the EDEN mother-child cohort. *PeerJ*, 2, e335.

- Pungello, E. P., Iruka, I. U., Dotterer, A. M., Mills-Koonce, R., & Reznick, J. S. (2009). The effects of socioeconomic status, race, and parenting on language development in early childhood. *Developmental Psychology, 45*(2), 544.
- Roßbach, H.-G. (2005). Effekte qualitativ guter Betreuung, Bildung und Erziehung im frühen Kindesalter auf Kinder und ihre Familien. In Sachverständigenkommission Zwölfter Kinder- und Jugendbericht (Ed.), *Bildung, Betreuung und Erziehung von Kindern unter sechs Jahren* (Materialien zum Zwölften Kinder- und Jugendbericht, pp. 55–173). München, Germany: DJI.
- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*(5), 1762–1774.
- Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child development, 83*(2), 508–525.
- Sachse, S., Anke, B., & von Suchodoletz, W. (2007). Früherkennung von Sprachentwicklungsstörungen – ein Methodenvergleich. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie, 35*(5), 323–331.
- Sperry, D. E., Sperry, L. L., & Miller, P. J. (2018). Reexamining the verbal environments of children from different socioeconomic backgrounds. *Child Development, 90*(4), 1303–1318.
- Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019). Language does matter: But there is more to language than vocabulary and directed speech. *Child Development, 90*(3), 993–997.
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). *Final report from the primary phase: Pre-school, school and family influences on children's development during key stage 2* (age 7–11). University of London, Institute of Education.
- Tamis-LeMonda, C. S., Bornstein, M. H., Baumwell, L., & Damast, A. M. (1996). Responsive parenting in the second year: Specific influences on children's language and play. *Early Development and Parenting, 5*(4), 173–183.
- Tamis-LeMonda, C. S., Bornstein, M. H., Kahana-Kalman, R., Baumwell, L., & Cyphers, L. (1998). Predicting variation in the timing of language milestones in the second year: An events history approach. *Journal of Child Language, 25*(3), 675–700.
- Tamis-LeMonda, C. S., Briggs, R. D., McClowry, S. G., & Snow, D. L. (2009). Maternal control and sensitivity, child gender, and maternal education in relation to children's behavioral outcomes in African American families. *Journal of Applied Developmental Psychology, 30*(3), 321–331.
- Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development, 57*, 1454–1463.
- Tudge, J. R. H., Mokrova, I., Hatfield, B. E., & Karnik, R. B. (2009). Uses and misuses of Bronfenbrenner's bioecological theory of human development. *Journal of Family Theory & Review, 1*(4), 198–210.
- Vallotton, C. D., Masturgeorge, A., Foster, T., Decker, K. B., & Ayoub, C. (2017). Parenting supports for early vocabulary development: Specific effects of sensitivity and stimulation through infancy. *Infancy, 22*, 78–107.
- Vasilyeva, M., & Waterfall, H. (2011). Variability in language development: Relation to socioeconomic status and environmental input. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (Vol. 3, pp. 36–48). New York, NY: Guilford.
- Vogel, F. (2020). *Elterliche und kindliche Einflussfaktoren auf Interaktionsverhalten und Bindungssicherheit – Auswirkungen von Temperament und Persönlichkeit auf das dyadische Interaktionsverhalten und die kindliche Bindungssicherheit in Belskys transaktionalem Prozessmodell des Erziehungsverhaltens*. Dissertation. Otto-Friedrich Universität Bamberg. Bamberg, Germany.

- Weinert, S., & Ebert, S. (2013). Spracherwerb im Vorschulalter. Soziale Disparitäten und Einflussvariablen auf den Grammatikerwerb. *Zeitschrift für Erziehungswissenschaft*, 16(2), 303–332.
- Weinert, S., & Ebert, S. (2017). Verlaufsmerkmale und Wirkfaktoren der frühen kognitiv-sprachlichen Entwicklung – Ergebnisse aus der BiKS-3-10 Studie. In V. Mall, F. Voigt, & N. H. Jung (Eds.), *Entwicklungsstörungen und chronische Erkrankungen: Diagnose, Behandlungsplanung und Familienbegleitung* (pp. 13–33). Lübeck, Germany: Schmid-Römhild Verlag.
- Weinert, S., & Grimm, H. (2018). Sprachentwicklung. In W. Schneider & U. Lindenberger (Eds.), *Entwicklungspsychologie* (pp. 445–470). Weinheim, Germany: Beltz.
- Weinert, S., Linberg, A., Attig, M., Freund, J.-D., & Linberg, T. (2016). Analyzing early child development, influential conditions, and future impacts: Prospects of a German newborn cohort study. *International Journal of Child Care and Education Policy*, 10(1), 1–20.
- Weinert, S., & Lockl, K. (2008). Sprachförderung. In F. Petermann & W. Schneider (Eds.), *Angewandte Entwicklungspsychologie (Enzyklopädie der Psychologie C/V/7)* (pp. 91-134). Göttingen: Hogrefe.

## Appendix

**Table A1:** Intercorrelations among variables

	1	2	3	4	5	6	7	8	9	10
1 Child's vocabulary <sub>26</sub>										
2 Child's verbal expressions <sub>14</sub>		.42***								
3 Child's vocabulary <sub>17</sub>	.29***		.27***							
4 Maternal education	.18***	.06+	-.05							
5 Household income	.15***	.03	.02	.32***						
6 Single parent (0=no 1=yes)	-.10**	-.08*	-.01	-.08*	-.19***					
7 Language-stimulating interaction behavior <sub>17</sub>	.24***	.24***	.05	.16***	.09*	-.05				
8 Child's age <sub>17</sub> (in months)	.01	.19***	.03	.10**	.03	-.02	.10**			
9 Child's age <sub>26</sub> (in months)	.23***	.04	.01	.09*	.09*	-.05	.02	.16***		
10 Child is a boy (0=no 1=yes)	-.17***	-.14***	-.07+	-.04	-.03	.04	-.15***	-.04	-.02	
11 Household language (1=German; 2=sometimes other; 3=mostly other)	-.10**	-.05	.03	-.12**	-.11**	.03	-.10**	-.04	-.01	-.02

*Note.* N = 778.<sup>+</sup>p < .10. \*p < .05. \*\*p < .01. \*\*\*p < .001.

Michaela Sixt & Christian Aßmann

## The influence of regional school infrastructure and labor market conditions on the transition process to secondary schooling in Germany

### Abstract

*After the systematic disadvantage of educational opportunities for children from rural areas in comparison to children from urban areas was a central research topic in the 1960s and 1970s, the role of regional disparities for educational opportunities at the transition to secondary level of the education system in Germany is now only irregularly discussed. There is still a relative lack of educational infrastructure in rural areas and regional labor market structures have hardly been considered as possible further determinants of regional educational disparities. Moreover, in the few current studies, important individual factors such as competencies, grades, school recommendations and educational aspirations could not be taken into account or they refer to individual regions. This paper picks up the question and tries to overcome these restrictions via combining data from the National Educational Panel Study Starting Cohort 3 (NEPS SC3) and regional information at the individual level. Once enriched with information on regional infrastructure, this data set not only contains necessary information on educational decisions, but also on important control variables at the individual level as mentioned above. Using multiple imputation to deal with missing values, hierarchical logistic regressions allow for investigating the impact of regional infrastructure on the probability of attending a high school in the school year 2010/2011 in Germany. The results show that school infrastructures are still significant factors in school choice even when controlling for relevant individual factors. The analyses show that a higher proportion of high schools to all secondary schools in the county a student lives in increases the probability that a high school is chosen at the transition from elementary to secondary school. For regional labor market structures no effects are revealed in this study.*

---

Dr. Michaela Sixt (corresponding author), Leibniz Institute for Educational Trajectories, Wilhelmsplatz 3, 96047 Bamberg, Germany  
e-mail: michaela.sixt@lifbi.de

Prof. Dr. Christian Aßmann, Leibniz Institute for Educational Trajectories, Wilhelmsplatz 3, 96047 Bamberg, Germany & Chair of Survey Statistics and Data Analysis, Otto-Friedrich-Universität Bamberg, Germany  
e-mail: christian.assmann@lifbi.de

## **Keywords**

*School choice; Regional infrastructure; Hierarchical binary logistic regression, Missing values*

# **Der Einfluss regionaler Bildungsinfrastruktur und Arbeitsmarktbedingungen beim Übergangsprozess zu weiterführenden Schulen in Deutschland**

## **Zusammenfassung**

*Nachdem die Bildungsbenachteiligung von Kinder aus ländlichen Gebieten im Vergleich zu Kindern aus städtischen in den 1960er und 1970er Jahren noch ein zentrales Forschungsthema war, wird die Rolle regionaler Disparitäten in Bezug auf Bildungschancen am Übergang zur Sekundarstufe des Bildungssystems in Deutschland zwischenzeitlich nur mehr unregelmäßig diskutiert. Dabei ist nach wie vor ein relativer Mangel an Bildungsinfrastruktur in ländlichen Gebieten zu verzeichnen. Zudem wurden regionale Arbeitsmarktstrukturen als mögliche weitere Determinanten von regionalen Bildungsdisparitäten bislang noch kaum in Betracht gezogen. Auch konnten in den wenigen, aktuellen Studien wichtige individuelle Faktoren wie Kompetenzen, Noten, Schulempfehlungen und Bildungsaspirationen nicht berücksichtigt werden oder sie beziehen sich auf einzelne Regionen. Daher greift das vorliegende Papier die Frage nach regionalen Bildungsdisparitäten erneut auf und versucht, diese Restriktionen mit Daten der Startkohorte 3 (NEPS SC3) des Nationalen Bildungspanels zu überwinden. Nach der Anreicherung mit Informationen über die regionale Infrastruktur enthält dieser Datensatz nicht nur notwendige Informationen über Bildungsentscheidungen, sondern auch über die oben erwähnten Kontrollvariablen auf der individuellen Ebene. Hierarchische logistische Regressionen ermöglichen unter Rückgriff auf Multiple Imputation zwecks Handhabung fehlender Werte die Untersuchung des Einflusses der regionalen Infrastruktur auf die Wahrscheinlichkeit im Schuljahr 2010/2011 in Deutschland ein Gymnasium zu besuchen. Die Ergebnisse zeigen, dass regionale Schulinfrastrukturen weiterhin signifikante Faktoren bei der Schulwahl sind, selbst nach Kontrolle relevanter individueller Faktoren. Die Analysen zeigen, dass ein höherer Anteil der Gymnasien an allen weiterführenden Schulen des Kreises, in dem ein Schüler lebt, die Wahrscheinlichkeit erhöht, dass ein Gymnasium beim Übergang von der Grundschule zur weiterführenden Schule gewählt wird. Für regionale Arbeitsmarktstrukturen können keine Effekte festgestellt werden.*

## **Schlagworte**

*Schulwahl; Regionale Infrastruktur; Mehrebenenanalyse; Imputation*

## 1. Introduction

Peisert (1967) coined the metaphor of the catholic working class girl in rural areas. It refers to the four especially disadvantaged groups regarding educational opportunities in the 1960ies in Germany. This metaphor was, and is still present in current educational research, since sex (e.g. Stanat, Bergmann, & Tarazow, 2016) as well as social background (e.g. Henz & Maas, 1995; Shavit & Blossfeld, 1993; Georg, 2006; Becker & Lauterbach, 2016; Schindler, 2015, 2017; Buchholz & Pratter, 2017; Blossfeld, Blossfeld, & Blossfeld, 2019) are either central objects of investigation or serve as default control variables. Next to sex and social background, research has addressed the effects of being affiliated to a certain religion, but there are doubts that there ever was, and is an effect of religion per se if social origin is controlled for (Helbig & Schneider 2014).

However, the effects of regional origin and infrastructure on educational opportunities have not received the same level of attention. There are few studies (see review of literature) dealing with this question, most of which utilize indicators for urbanity, instead of measuring infrastructure directly. Sixt (2010, 2013) measures the educational offering based on the number of schools per county ("Kreise"). Furthermore, the regional perspective was widened not only to include educational opportunity structures, but also to encompass labor market conditions, which could also be relevant for educational decisions. The analyses are based on data for birth cohorts ranging from 1980 to 1996 in West Germany as provided by the GSOEP. Unfortunately with these data, relevant factors for educational decisions such as competencies, grades, school recommendations or educational aspirations cannot be controlled for. Also for labor market structures the operationalization of the appropriate context measurements is still an open question. Further, these former analyses refer to educational decisions between 1990 and 2006 and are hence not informative about developments in the school infrastructure over the last few years, such as for example the school closures related to decreasing numbers of students as discussed in Weishaupt (2006). After educational expansion and in times of overall increased educational aspirations (Gehrman, 2019 p. 104) one could hence ask if regional infrastructures still play a role at all.

The fact that less attention has been set on regional disparities is astonishing since on the aggregate level school infrastructure as well as educational participation rates in Germany are often focused. According to the National Report on Education, the sizes of catchment areas in general have been increasing from metropolis to sparsely populated districts for all school tracks, see Tab. D1-7web Autorengruppe Bildungsberichterstattung (2016). For schools where pupils can pursue a high school degree ("Abitur"), the average catchment area varies from 9.5 sq. km for a metropolis, to 52.2 sq. km for urban counties, and 139.9 sq. km and 236.1 sq. km in rural counties with some agglomerations and sparsely populated counties respectively. The report also documents regional educational disparities between districts in terms of the proportion of students attending high schools in

the seventh grade in the school year 2014/2015, with this figure varying between 74 and 12 percent (Autorengruppe Bildungsberichterstattung, 2016, p. 78). This evidence suggests that regional disparities prevail. The question is then whether they have an impact at the individual level of educational decisions.

In line with Sixt (2010, 2013) this paper addresses the question of the influence of regional infrastructure in Germany on educational decisions at the transition from elementary to secondary education. The paper tries to overcome some of the restrictions of the former studies and reassesses the results of former analyses. The transition itself has always been a sensible phase for later educational success (Blossfeld, 1988). Students typically enter secondary education after grade four in most federal states. The individual choice is between different tracks or types of school, where a rough categorization is based on the school types “Hauptschule”, “Realschule”, “Gesamtschule”, “Schulen mit mehreren Bildungsgängen”, and “Gymnasium”. While in general “Hauptschule” and “Realschule” prepare for a vocational education, the “Gymnasium” can be seen as a kind of high or grammar school preparing for academic education and university. “Gesamtschule” or “Schulen mit mehreren Bildungsgängen” are kinds of comprehensive schools where students learn both vocational and academic skills, but only in some cases they provide a high school degree (“Hochschulreife”/“Abitur”) that enables university entrance directly. Throughout secondary education, students can move upward or downward between school tracks. The “downward mobility” of students from the academically oriented school type to the lower or middle secondary school types and the comprehensive school type is, however, much higher than “upward mobility” (Baumert, Trautwein, & Artelt, 2003; Rösner, 2005; Hillmert & Jacob, 2005; Ditton, Elsäßer, Gölz, Stahn, & Wohlkinger, 2017; Zielonka, 2017).

This article focuses in particular on the choice to attend high school (“Gymnasium”) in comparison to other types of secondary schooling. High schools are much more spatially segregated than other non-academic school types, and the choice of a high school is the clearest expression of the educational desire for academic education. It opens up a direct path to tertiary education, and offers the most options for revising the path taken. The de-differentiation of the German secondary school system in recent years, due to demographic change (Bartl & Sackmann, 2014), tending towards a two-track system is also in line with a dichotomous view on the decision. Consequently, the corresponding regional infrastructure is operationalized as the proportion of high schools in a county in comparison to all other schools in the secondary school system. In addition, the paper at hand delves into the question of the influence that regional infrastructures related to labor market conditions exhibit on the decision for high school attendance. In doing so, the paper largely follows the analyses of Sixt (2010, 2013) and replicates these results in a broader sense on the basis of more adequate and recent data.

After the recapitulation of current research in section 2, section 3 provides theoretical considerations regarding the educational decision studied following the cost-return approach. Regional characteristics are defined as determining factors of relative costs and returns, rating different benefits of educational alternatives.

Section 4 gives an overview of the data used and section 5 provides a description of the statistical methods. Section 6 provides the empirical results, and the paper ends with a short conclusion and discussion of the results in section 7.

## 2. Review of literature

Baur (1972) provided an analysis addressing the effects of residence size on the educational decision to enter different types of secondary schooling. The transition rates calculated based on data from parent surveys in Baden-Württemberg in the years 1967, 1968, and 1969 differ according to the size of residential areas: the bigger the community, the higher the high school attendance rates. Based on the Mikrozensus data from 1972, Trommer-Krug (1980) compares the proportion of students in the age range between ten and fifteen in the different school tracks and uses as indicators for the regional infrastructure the length of the journey to school and the size of residence. Her results show, that in communities under 20,000 inhabitants, only 11 percent of the observed students attend high school, whilst this proportion reaches 20 percent in communities with more than 20,000 inhabitants. Henz and Maas (1995) study the differences regarding educational opportunities between rural and urban regions in Germany via multilevel regression models. Using data from the The German Life History Study on western German cohorts born between 1929 and 1961, their results point to an – although for younger cohorts diminished – still persistent negative effect of rural residence on the probability of attending high school when separating between sex and controlling for social origin. Taking a look at school offer and demand in Hamburg in 1973, Bartels (1975) analyses data from a parent questionnaire of grade four students. While the intention prevailed to send children to a certain school type (“Gesamtschule”), the children actually attended other types of schools. One reason was that the intended school type was too far away. Nowey (1983) has published similar results from a study in Bavaria combining official data from the Bavarian State Office for Statistics and data from a second official source (“Strukturdatenbestand in der Fachdatenbank des Bayerischen Staatsministeriums für Landesentwicklung und Umweltfragen”). He concluded that the choice probability of a certain school decreases with the distance to that school.

Hansen (1993) has looked at the motivations of school choice for parents in Dortmund using a full census of all parents of grade four students at elementary schools in the school year 1986/1987. As the most important motives for selecting a school, more than half of the parents named “no long journeys” and “school nearby”. In the group of parents who had not chosen a high school, the proportion of parents providing these answers was even higher. Only for the group of parents of future high school students the intended degree was a more important determinant of school choice than the physical distance to the school (Hansen, 1993, p. 142). This is in line with the results of Fickermann (1997) investigating the pre-

ferred school type of elementary school parents and the regional school structure in Thuringia in 1991/1992. He reported that the aspiration to attend a high school declined with the distance to the next high school: 46 percent of parents wanted their children to go to high school if there was a high school at their place of residence, 42 percent if the high school was up to five kilometers away. This proportion became much smaller if in order to attend a high school the children had to travel five to ten kilometers or even further, with 27 percent and 15 percent of parents respectively. Via path analysis, Fickermann (1997) confirms the development of regional disparities in terms of the density of schools as a determining factor of school choice. Even nowadays the distance to school seems to be a very important factor for school choice. This has been confirmed by Clausen (2006, 2008) using sample data of parents and students in grade four from Mannheim and Heidelberg. Riedel, Schneider, Schuchart, and Weishaupt (2010) investigating elementary school choice in Wuppertal highlight that schools nearby are chosen more often than those in line with the administrative school districts.

Sixt (2010, 2013) has investigated the correlation between the proportion of high schools to all secondary schools in a county and the probability to attend a high school via logistic multilevel regressions with data from the German Socio-Economic Panel (GSOEP) to obtain a better approximation for school infrastructure. The results show a significant and positive relationship between the proportion of high schools in a county and the probability of starting at high school for first-born children of the birth cohorts ranging from 1980 to 1996 in West Germany. Furthermore, the analysis focused on the interaction of regional and social origin and the question whether the regional infrastructure is more important for children with a less advantageous social background. The study documents an interaction effect but not as expected, as especially children with parents with an academic background profit from a higher proportion of high schools to all secondary schools in a county. As a possible explanation a so-called "Composition" effect was discussed, as the result could be influenced by the effect that families with higher amount of positive resources for education (measured by income, educational or occupational status) could be found in regions with a better (school) infrastructure more often. However, when interpreting the results of this study, one should take into consideration that the results could not be controlled for important individual variables, since e.g. competencies, grades, school recommendation or educational aspirations are not incorporated in the GSOEP data. As these variables are central for the selection of the school track (see theoretical considerations below) the results also could be biased. Sixt (2018) tries to overcome these restrictions by taking a closer look at decisions regarding high school attendance in the school year 2016/2017 via data from a regional project in Upper Franconia. She measures relevant individual information and uses physical distances to potential schools instead of aggregate measures regarding school offering available in administrative data. Furthermore, she controls for the region the families live in. These analyses reveal two central findings. First, in case distances are used in the multivariate logistic regression models, there is no difference between children liv-

ing in rural and in urban areas in terms of the probability to attend a high school. Secondly, the further away the high school is from the place of residence, the less likely children attend high schools – independently of social origin. Unfortunately, these findings can be generalized for rather rural areas like Upper Franconia only, and it remains an open question if there is an effect of regional origin on educational opportunities at the national level.

In addition, the question remains whether there are any other regional structures that influence (individual) decisions at the transition to secondary education. Former studies have considered labor market indicators for the transition to vocational training and found some evidence for positive correlations (e.g. Bolder 1975, 1984; for a more current study, see Hillmert, Hartung, & Weßling, 2017). Sixt (2010, 2013) transfers this question to the transition from elementary to secondary school and argues in terms of a more or less long-term estimation of the benefits parents expect of a certain school track respectively school certificate. Staying in school for a longer time can be a strategy to avoid youth unemployment in a rather weak labor market. Reaching a high school certificate could also enhance the chances of getting an apprenticeship. In a strong labor market with a demand for differentiated qualified employees, the chances to get a job with another educational certificate than a high school degree are even higher, so that the benefits of alternative school tracks should increase. But, with regard to the results, the unemployment rate and the number of highly qualified persons at the level of “Raumordnungsregionen” show no effects. In her discussion, Sixt (2010, 2013) refers to the uncertainty whether the available operationalization of regional labor markets in terms of “Raumordnungsregionen” is the relevant one, as parents may have in mind an individually defined labor market context without correspondence to administrative borders.

Summarizing the literature discussed here, the effect of regional infrastructure on school choice seems to have decreased over time. However, the importance of the regional school offering for educational decisions has not disappeared, and the distance to school is one important factor determining parents' choice of school for their children. The studies mentioned so far typically work with indicators (size of residence, classification rural/urban area) for the reachability of schools and show that there is a gap between rural and urban areas regarding educational opportunities at the transition to secondary school, at least until the year 1990. Recent studies working with more direct measures for the offer or the reachability of schools are rare, while those of Sixt (2010, 2013) have other restrictions. This article is therefore concerned about the relevance of residence and region for educational success in Germany nowadays. We try to overcome the restrictions discussed above and ask again if the regional school offer and labor market conditions influence the educational decision of parents at the transition from elementary to secondary school.

### 3. Theoretical considerations

The theoretical considerations on the effects of regional infrastructure on educational decisions in the papers Sixt (2010, 2013) and in the paper at hand are based on a cost-benefit approach. Such approaches, especially those taking a sociological perspective on Rational Choice Theory (RCT), are currently – despite all criticism (Simon, 1993) – predominantly used as an explanatory approach, especially when it comes to educational inequality (Stocké, 2012). At the core of the socio-logical RCT, as found in Boudon (1974), Erikson and Jonsson (1996), Breen and Goldthorpe (1997) or Esser (1999), is the assumption that actors in decision-making situations weigh up the (subjectively expected) utilities of the possible alternatives against each other and then decide on those with the highest utility.

The utility is on the one hand based on the (subjectively expected) costs of an alternative. This includes direct costs like financial costs, as well as opportunity costs that arise from lost earnings due to longer schooling instead of working on the labor market, but also expenditures in the form of time. Regional infrastructure plays an important role in determining costs. A worse or less dense offer of schools in a region leads to longer journeys to schools for individuals. This implies higher direct costs in the form of time and money spent on traveling to school. Also, higher opportunity costs arise if time must be spent on the journey to school instead for homework, hobbies, or friends. Furthermore, a longer journey to school can go along with physical or psychological stress, which needs to be taken into account as a cost factor as well.

To calculate the utility of a specific school choice, the costs of this alternative are balanced with its (subjectively expected) benefits. These can be decomposed into benefits conditional on successfully achieving the corresponding educational certificate, and the corresponding success probabilities. With a high school degree (“Abitur”) it is argued that chances on the labor market, such as the chance of obtaining a job with a higher salary or with higher job security, are higher than with another degree. Thus regional labor market structures are possible determinants of the benefits of educational decisions, as these shape parental expectations regarding perceived benefits paying off the costs of education. Taking investment in education, e.g. the choice of a high school, as an investment in a child’s competitiveness regarding better jobs or job security, we would argue that the regional rate of unemployment has an impact on the kind of school track that is chosen at the end of primary education. We assume that the default choice favors a high school compared to other school tracks if there is a high rate of unemployment. Parents’ logic would be to bestow upon a child an advantage against the competition when it comes to apprenticeships, and in the long run, good jobs, by ensuring them a high school degree. An alternative interpretation could be that longer periods spent at school push back the transition to the labor market and may protect a child against youth unemployment. We further assume that individual expectations are formed by the regional labor market structure, as the investment in higher ed-

ucation may be easier to amortize in labor markets that require highly qualified people. Accordingly, a relationship between the chosen educational alternative and the qualification profile that is typically demanded on the labor market can be expected. If there is a high proportion of jobs where higher education is requested, the expected benefits of attending high school increase. If there is a low proportion of highly-skilled jobs, the benefits and the utility of the high school track are reduced.

Last but not least, *subjective expectations* are depending on the individual context, where this context relates to individual factors like age, the assessment of abilities, and skills (see also primary effects of social origin, Boudon, 1974, pp. 29ff.). To access these factors, school grades and school recommendations (which can also be seen as formal restrictions in some federal states) are taken into account. Furthermore, family factors like income, occupation or level of parent education play an important role. For example, educational inequalities according to social origin are in part explained by different subjective expectations with regard to success probabilities associated with alternative educational tracks given the same level of performance. This leads to social selective educational decisions, as parents with higher education levels expect a higher probability of their children successfully finishing the chosen school track and have higher educational aspirations (secondary effects of social origin, Boudon, 1974, pp. 29ff.; see e.g. Relikowski, 2012; Gresch 2012). In addition, gender (see e.g. Schuchart 2010) as well as migration background (see e.g. Kristen & Dollmann, 2009) are seen as important individual factors that lead to different educational aspirations and therefore different educational decisions.

Summarizing the argumentation with regard to regional infrastructure, we expect that (a) the better the regional offer of a certain educational option is, the more likely this option will be chosen due to lower costs. As the benefits of choosing this school type should increase, we expect that (b) the worse the regional labor market is regarding job security, and (c) the more highly qualified the occupations offered in a regional labor market are, the more likely a high school will be chosen.

#### **4. Description of data and sample characteristics**

To test the formulated theses, individual and aggregated regional data providing information on individual educational decisions, as well as regional structures need to be considered. Data from the National Educational Panel Study: Starting Cohort Grade 5, doi: 10.5157/NEPS:SC3:6.0.1. (NEPS SC3) is used to assess individual and family characteristics, as well as educational decisions. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research, funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg, in coop-

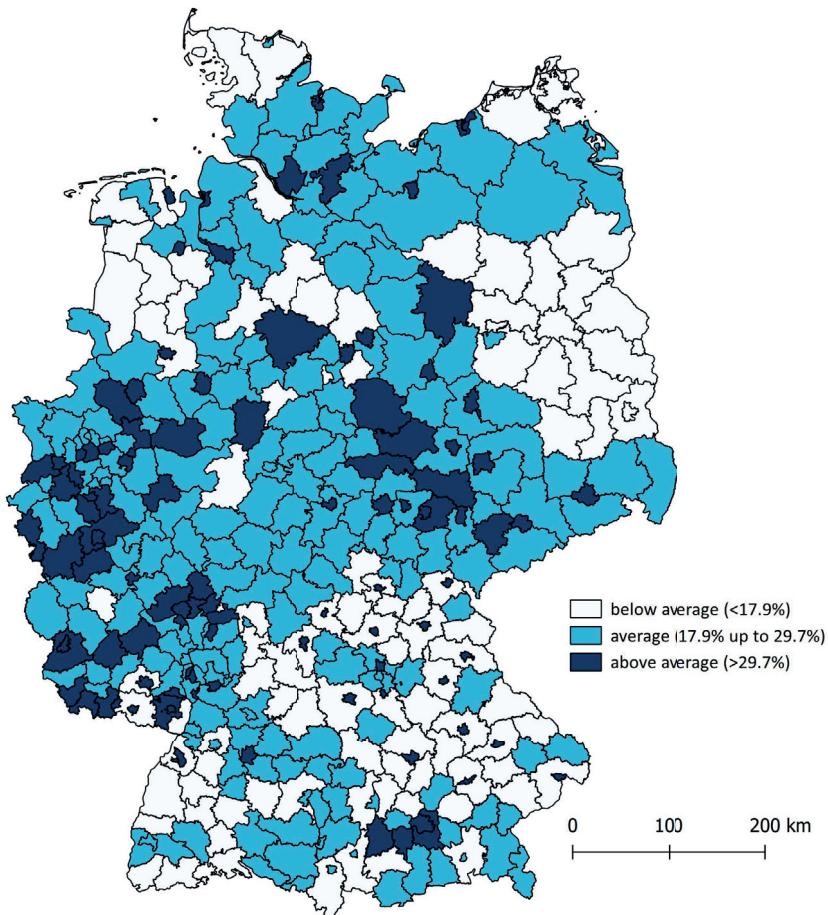
eration with a nationwide network (for a short overview see Blossfeld, von Maurice, & Schneider, 2011). It contains data for a sample of fifth-graders who opted at the end of the school year 2009/2010 for secondary school tracks (see below for information on the sampling approach). Hence, data on regional structures match to the time point of the decision whenever possible. The matching of regional data to NEPS SC3 is based on the municipality coding as of 31.12.2013, where there are 402 counties in Germany in 2013. Because of regional reforms, the number of counties varies from 2009 to 2013, especially due to reforms of the cutting of county regions in the eastern parts of Germany. In most cases, counties were simply pooled by the reforms, so that the corresponding structural information could be easily calculated for the new counties. In two cases the new borders do not fit with the old ones, so the structural information was adjusted proportional to population.

We use (analogous to Sixt 2010, 2013) the 4-stage *classification of counties* of the Federal Institute for Research on Building, Urban Affairs and Spatial Development (Bundesinstitut für Bau-, Stadt- und Raumforschung, BBSR). From the total of 402 German counties in 2009, 68 (16.9 percent) belong to metropolises (1, “Kreisfreie Großstädte”), 137 (34.1 percent) to urban counties (2, “Städtische Kreise”), 100 (24.9 percent) to rural counties with some agglomerations (3, “Ländliche Kreise mit Verdichtungsansätzen”), and 97 (24.1 percent) to rural counties with a sparse population (4, “Dünn besiedelte ländliche Kreise”). We use this indicator to establish a first impression of whether there is still a gap between rural and urban areas regarding high school attendance. To obtain a much closer measurement, we use information about the regional *school infrastructure* available from “Regionaldatenbank Deutschland”. It is a database which contains detailed official statistical data, e.g. information about the number of schools by type on the county level for several years. We generated the *proportion of high schools* (only “Gymnasien”, analogous to Sixt 2010, 2013) of a county in the school year 2009/2010 in relation to all other regular schools in the secondary school system (“Orientierungsstufe” + “Hauptschule” + “Schule mit mehreren Bildungsgängen” + “Realschule” + “Integrierte Gesamtschule” + “Waldorfschule”) to consider the relative offer of schools leading directly to university, and argue that the county level is the appropriate administrative area parents take into account when choosing a secondary school. The most appropriate measurement would be information about the physical distances from schools to parents’ houses. Such information would allow us to ignore the problem of the accurate definition of the relevant context, also known as the Modifiable-Area-Unit-Problem (MAUP), as discussed e.g. by Wu (2007). However, until such information is available, we can only use the data from the statistical offices, which are oriented towards administrative boundaries. With the exception of Lower Saxony, there is no regulation regarding school catchment areas, as it is the case with the “Schulsprengelprinzip” for elementary schools. Parents have the right to enroll their child at any secondary school in the respective federal state if the school has places available and the child meets the requirements for the respective type of school. However, using the federal state lev-

el as the context for measuring the considered school infrastructure maybe inadequate, as it can be assumed that the parents will always have a school in mind that can be reached by the child from their place of residence. As the level of communities is too small, the county level will be taken into account. Additionally, it is also worth considering that school planning is mostly based on this administrative area.

Figure 1 shows clear regional disparities on the county level for the proportions of high schools to all secondary schools in Germany 2009. The proportions range from 0.0 to 47.5 percent, with a median and mean (standard deviation) of 24.2 (8.5) percent. In order to describe the data, we classify the counties according to quartiles; counties with a proportion between 17.9 and 29.7 percent are described as counties with an average quantity of high schools, counties below 17.9 percent (lower quartile) are considered as under-average, whereas those with more than 29.7 percent (upper quartile) are considered as above average (see Table 1).

Figure 1: Proportion of high schools to all secondary schools per county in Germany 2009



*Note.* Source: Statistical offices of the Federation and the Länder, own illustration.

As shown by Table 1, the proportion of high schools to all secondary schools differs according to the classification of the county. The more urbanized a county is, the higher tends the proportion of high schools to be. The pairwise correlation amounts to -.409 with a corresponding  $p$  value below .001 when testing for zero correlation. Information about the regional labor market is taken from INKAR “Indikatoren und Karten zur Raum- und Stadtentwicklung”, an online tool edited by the BBSR. To measure later competitiveness concerning jobs or job security, we take the unemployment rate at the county level in 2009, and to measure the offerings of highly skilled jobs in a county, we take the proportion of highly skilled employees in 2012. Unfortunately, there is no timely closer measurement available. On average (not depicted), the proportion of high schools to all secondary schools increases from 9.7 percent in 2012 to 10.2 percent in 2013; 10.6 percent in 2014 and 11.1 percent in 2015 respectively. This could lead to an overestimation of the effects that we will have to keep in mind when interpreting the results.

To overcome the problem that we do not know the individually relevant regional context for the labor market (see Sixt, 2010, 2013 and discussion above), we consider weighted versions of both indicators, where weighting is based on commuter flows labor at the county level as recorded in the statistics of the Federal Employment Agency and released by the Institute for Employment Research in 2018. The weighted labor market indicator for a county is thus given as  $\tilde{x}_i = \sum_{j=1}^N P_{ji}/(\sum_{j=1}^N P_{ij}) x_i$ , with  $P_{ij}$  denoting the commuter flows from county  $i$  to  $j$ , and  $x_i$  the corresponding labor market indicator of county  $i$ . The weighted 2009 unemployment rates per county range from 3.2 to 17.3 percent, with a median of 6.9 and a mean (standard deviation) of 7.7 (3.4) percent, and the spatially weighted proportion of highly qualified employees 2012 ranges from 4.2 to 20.7 percent with a median of 8.8 and a mean (standard deviation) of 9.5 (2.9) percent. Note that the reported weighted indicators do not differ in a significant manner from the originals: the spatially unweighted unemployment rate of 2009 (not depicted) ranges from 2.2 to 17.8 percent, with a median of 6.7, and a mean (standard deviation) of 7.7 (3.4) percent, and the spatially unweighted proportion of highly qualified employees 2012 ranges from 4.0 to 29.7 percent, with a median of 8.5 and a mean (standard deviation) of 9.7 (4.4) percent.

As further shown by Table 1, the regional labor market indicators vary with regard to the classification of the county. Again, in order to describe the data, we classified the counties according to quartiles: counties with an unemployment rate between 4.9 and 9.9 percent are described as average, below 4.9 percent (lower quartile) are considered below average, and those counties with more than 9.9 percent (upper quartile) are above average. For the proportion of highly qualified employees, the lower quartile starts below 7.3 and the upper at 11.3 percent (see Table 1). While the unemployment rate reveals no clear pattern, the proportion of highly skilled employees is again positively correlated with urbanity. The pairwise correlation amounts to -.561 with a corresponding  $p$  value below .001 when testing for zero correlation.

Table 1: Regional structures by type of county

	type of county				
	metropolises (n = 68)	urban counties (n = 137)	rural counties with some agglomeration (n = 100)	rural counties with a sparse population (n = 97)	total (n = 402)
proportion of high schools to all secondary schools (pwcorr: -.409; p < .001)					
minimum	15.3 %	5.9 %	0.0 %	5.9 %	0.0 %
25 % percentile	26.4 %	19.8 %	16.7 %	14.3 %	17.9 %
median	31.1 %	24.6 %	22.2 %	20.0 %	24.2 %
M (SD)	31.1 % (7.1)	25.0 % (7.5)	22.1 % (8.6)	20.2 % (7.4)	24.2 % (8.5)
75 % percentile	35.1 %	29.8 %	28.0 %	25.0 %	29.7 %
maximum	47.5 %	45.5 %	47.1 %	38.5 %	47.5 %
unemployment rate (pwcorr: -.015, not significant)					
minimum	3.8 %	4.0 %	3.2 %	3.2 %	3.2 %
25 % percentile	8.0 %	4.6 %	4.4 %	5.1 %	4.9 %
median	9.5 %	5.8 %	6.3 %	7.9 %	6.9 %
M (SD)	9.9 % (2.8)	6.4 % (2.5)	7.3 % (3.5)	8.6 % (4.0)	7.7 % (3.4)
75 % percentile	12.2 %	7.6 %	10.2 %	11.6 %	9.9 %
maximum	14.2 %	14.1 %	15.2 %	17.3 %	17.3 %
proportion of highly qualified employee (pwcorr: -.5611; p < .001)					
minimum	7.5 %	5.8 %	4.5 %	4.2 %	4.2 %
25 % percentile	11.1 %	7.2 %	6.3 %	5.6 %	7.3 %
median	15.7 %	8.8 %	7.2 %	7.1 %	8.8 %
M (SD)	15.5 % (5.3)	9.7 % (3.6)	7.9 % (2.4)	7.4 % (2.2)	9.5 % (2.9)
75 % percentile	18.0 %	11.4 %	9.4 %	9.1 %	11.3 %
maximum	20.7 %	18.9 %	13.7 %	12.9 %	20.7 %

Notes. Source: BBSR, Regionaldatenbank Deutschland, INKAR, n = 402, own calculations. pwcorr denotes the pairwise correlation coefficient; SD denotes the standard deviation; p denotes the estimated p value; pwcorr for proportion high schools and unemployment rate is .2658 with a corresponding p value < .0001; pwcorr for proportion high schools/proportion of highly qualified employee is .4965 with corresponding p value < .0001; pwcorr for proportion of highly qualified employee and unemployment rate is .2080 with corresponding p value < .0001.

Unfortunately, we cannot measure some of the factors relevant for the educational decision before it has actually taken place, since the NEPS data collection starts after the transition to grade five. A potential way out could have been provided by NEPS Starting Cohort 2 data. However, after the transition to the secondary school track, the NEPS Starting Cohort 2 does not show sufficient heterogeneity to allow for disentangling the effects of regional infrastructure on the decision. But as the measurement in the NEPS SC3 is very early in grade five (November/December, with the school year starting in August/September), we assume that the measurements, especially with regard to competencies, are still comparable to the

situation before the transition (grades refer to grade four and are surveyed retrospective in the parent and the student questionnaires). The NEPS data are the only data representative for Germany at the moment, offering diverse control characteristics like grades, competencies and information about social and migration background – and offer the possibility to add regional data.

We use all cases of NEPS SC3 for our analyses where students attend a regular secondary school in the school year 2010/2011 and have participated in the first NEPS wave. As in Berlin and Brandenburg the transition from elementary to secondary school takes place two years later than in the other federal states, students from these locations are excluded. In the end, we use a subsample of 4,880 from the total of 6,112 NEPS SC3 cases. As described in the introduction, our dependent variable is dichotomous: attending high school in grade five vs. not attending high school. As shown in Table 2, 46.9 percent of the children in our sample enter a high school. The sample of students has been established via a two stage sampling procedure based on an explicit stratification across school types and an implicit stratification across federal states and regional classification relating to the forerunner of the DegUrba “Abgrenzung ländlicher von städtischen Gebieten” classification (for details see Aßmann et al., 2011; Steinhauer et al., 2015; and Schnorr-Bäcker, 2014). Schools were sampled proportional to the number of classes per school at the first stage. As only two classes should be selected in case of three or more classes per school two classes were sampled at the second stage in order to aim at equal inclusion probabilities across schools and school types. All students in the sampled classes were asked to participate in the NEPS survey. Further, as the sampling is based on an implicit stratification across federal states and no effect of the regional classification on school participation prevails (see Steinhauer et al. 2015, Table 5), one school per county is included in the sample on average. Only for the independent cities and city states more than one school is included. Given this, no bias with regard to the distribution of types of regions is imposed on the sample and hence no effect on the regression analysis assessing the impact of regional factors on the individual educational decisions is expected. Due to the individual participation decision varying across schools, the observed sample quota of 46.9 percent for entering a high school is slightly higher than the quota of 44.0 percent reported for school year 2016/2017 (see Statistisches Bundesamt, 2018). To check for robustness, we have also run weighted regressions, see the section on statistical methods for details.

Table 2: Sample description – categorical variables

	n	proportion	observed sample	imputed samples	
				M of imputed samples	SD between imputations
<b>high school attendance grade 5</b>					
no	2,592	0.531	-	-	-
yes	2,288	0.469	-	-	-
missing	0	-	-	-	-
<b>gender</b>					
female	2,360	0.484	-	-	-
male	2,520	0.516	-	-	-
missing	0	-	-	-	-
<b>migration background</b>					
no	4,121	0.844	0.846	< 0.001	
yes	749	0.153	0.154	< 0.001	
missing	10	0.002	-	-	-
<b>recommendation for high school</b>					
no	2,046	0.419	0.614	0.003	
yes	1,493	0.306	0.386	0.003	
missing	1,341	0.275	-	-	-
<b>idealistic educational aspiration of children for "Abitur"</b>					
no	1,385	0.284	0.300	0.001	
yes	3,313	0.679	0.700	0.001	
missing	182	0.037	-	-	-
<b>university degree (at least one parent ISCED 5B or higher)</b>					
no	1,470	0.301	0.476	0.005	
yes	1,813	0.372	0.524	0.005	
missing	1,597	0.327	-	-	-
<b>occupational status according to EGP classification (3 categories)</b>					
service class: EGP I + II	906	0.186	0.281	0.003	
working class: EGP V–VII	978	0.200	0.416	0.005	
others: EGP III–IV	895	0.183	0.303	0.004	
missing	2,101	0.431	-	-	-
<b>type of county</b>					
metropolises	653	0.134	0.200	0.001	
urban counties	1,587	0.325	0.450	0.001	
rural counties with some agglomeration	733	0.150	0.221	0.001	
rural counties with a sparse population	450	0.092	0.128	0.001	
missing	1,457	0.299	-	-	-

*Notes.* Source: NEPS SC3 (doi:10.5157/NEPS:SC3:6.0.1), own calculations. Due to rounding differences, deviations in the sum of the individual values are possible. *M* denotes the arithmetic sample mean and *SD* the sample standard deviation.

Table 3: Sample description – metric variables

		observed sample	imputed samples	
			<i>M</i> of imputed samples	<i>SD</i> between imputations
age in years and months	<i>n</i>	4,878		
	median	10.667	10.667	< 0.001
	<i>M</i>	10.722	10.722	< 0.001
	<i>SD</i>	0.523	0.523	0.001
reading comprehension (WLE)	missing	2	-	-
	<i>n</i>	4,867		
	median	-0.022	-0.022	< 0.001
	<i>M</i>	-0.008	-0.010	0.001
mathematical competence (WLE)	<i>SD</i>	1.255	1.255	0.001
	missing	13	-	-
	<i>n</i>	4,867		
	median	-0.010	-0.010	0.002
grade in mathematics	<i>M</i>	0.001	0.001	0.001
	<i>SD</i>	1.167	1.167	0.001
	missing	13	-	-
	<i>n</i>	4,753		
grade in German	median	2.000	2.000	< 0.001
	<i>M</i>	2.302	2.310	0.002
	<i>SD</i>	0.941	0.946	0.003
	missing	127	-	-
net equivalized income per month [€]	<i>n</i>	4,740		
	median	2.000	2.000	< 0.001
	<i>M</i>	2.351	2.360	0.002
	<i>SD</i>	0.877	0.881	0.002
proportion of high schools to all secondary schools	missing	140	-	-
	<i>n</i>	3,317		
	median	1428.571	1401.439	21.001
	<i>M</i>	1612.550	1574.155	28.054
proportion of highly qual- fied employees	<i>SD</i>	1447.164	1378.932	97.396
	missing	1,563	-	-
	<i>n</i>	3,423		
	median	0.242	0.245	0.022
regional unemployment rate	<i>M</i>	0.241	0.243	0.021
	<i>SD</i>	7.310	7.448	0.012
	missing	1,457	-	-
	<i>n</i>	3,423		
	median	0.067	0.068	0.004
	<i>M</i>	0.072	0.072	0.006
	<i>SD</i>	2.569	2.604	0.006
	missing	1,457	-	-
	<i>n</i>	3,423		
	median	0.092	0.093	0.009
	<i>M</i>	0.100	0.100	0.008
	<i>SD</i>	3.064	3.064	0.005
	missing	1,457	-	-

Notes. Source: NEPS SC3 (doi:10.5157/NEPS:SC3:6.0.1), own calculations. *M* denotes the arithmetic sample mean and *SD* the sample standard deviation.

Table 2 provides a further description of the categorical sample variables, whereas Table 3 refers to the metric sample variables. It is shown that most of our students are on average (standard deviation) 10.7 (0.5) years old, and nearly half of the sample is female (48.4 percent). Most of the students have no migration background (84.4 percent), meaning that they or their parents are not migrants. The average WLE-estimator (standard deviation) for reading comprehension amounts to -0.008 (1.255); for mathematical competence to 0.001 (1.167). The median grade in German, as well as Mathematics, is 2.0. Regarding the school recommendation, we find 30.6 percent with a positive recommendation for high school, while 67.9 percent of the children express a desire to obtain a high school degree (idealistic educational aspiration). Unfortunately, there is information missing from 29.9 percent of the parents, as parent participation was not coupled to student participation, see also below. Consequently, and together with item-non-response, 32.7 percent of information is missing with regard to the educational background of the family, as measured by the highest educational degree of the parents. 37.2 percent of the children in the sample have at least one parent with an ISCED-level of 5B or even higher. 18.6 percent of the families belong to EGP-Class I or II; 20.0 percent to EGP-Class V-VII, and 18.3 percent to EGP-Class III and IV, while information for 43.1 percent is missing. The monthly average net equivalised income, based on information for 68.0 percent of the families, is 1,612.55€, with the standard deviation of 1,447.16€ and the median at 1,428.57€. Regarding the regional origin, 13.4 percent of families live in a metropolis, 32.5 percent in urban counties, 15.0 percent in rural counties with some agglomerations, and 9.2 percent of the counties are classified with a sparsely population. The average (standard deviation) proportion of high schools to all secondary schools in the counties the students live in is 24.1 (7.3) percent (median 24.2), the unemployment rate 7.2 (2.6) percent (median 6.7), and the proportion of highly skilled employees 10.0 (3.1) percent (median 9.2). Again, due to the missing interview with the parents, we are missing information on residence addresses, and therefore the regional setting of 1,457 families, which makes up 29.9 percent of the sample. To handle missing values, we adapt a multiple imputation strategy, as described in the next section. As shown in the last two columns of Tables 2 and 3, the variation between the different imputed data sets is, however, only modest.

## 5. Statistical methods

To assess the effect of regional school infrastructure on individual educational decisions, we use binary logistic regressions accounting for the hierarchical regional structures with students nested within living regions and schools via random intercepts. Let  $y_{ij}$  denote the decision of an individual  $j$  in region  $i$  to join a high school, i.e.  $y_{ij} \in \{0,1\}$  for all regions  $i = 1, \dots, N$  and  $j = 1, \dots, J_i$ , where  $J_i$  denotes the number of individuals observed in living region  $i$ . Then the regression is implied by

$y_{ij} = I[X_{ij}\beta + W_i\gamma + \alpha_i + \epsilon_{ij} > 0]$  with  $I[\cdot]$  denoting the indicator function,  $\alpha_i$  denotes an unconditional independently and identically normally distributed regional effect with mean zero and variance  $\sigma_\alpha^2$ , whereas  $\epsilon_{ij}$  is identically and identically standard normally distributed (see Greene, 2004). We denote the set of regressors with variation at the individual and regional level including a constant as  $X_{ij}$ , and the regressor varying only at the regional level as  $W_i$ . Parameter estimation is directly accessible towards Maximum Likelihood using quadrature based integration (see Butler & Moffitt, 1982).

To handle the uncertainty within parameter estimation stemming from missing information, (see Table 2 and Table 3), we make use of multiple imputation via chained equations (MICE, see Rubin, 1976; van Buuren & Groothuis-Oudshoorn, 2011). Note that missing information in the regressor variables related to  $X_{ij}$  can be handled, as typically in MICE, via specification of a set of full conditional distribution models adapted to the scale of each variable and considering all other (completed) variables as conditioning factors. The information about in which region a child lives may be missing, as this information is surveyed within the parental interview and not all parents participated in the survey. As the region of schooling and the region of living relevant for the individual educational decision do not coincide per se, regional information cannot be matched to the individual level based on the known region of school. However, as each child is clustered within a school, we make use of the observed within school distribution of living regions to fill in the missing regional information on children's places of residence. Missing values are thereby imputed by drawing from the observed within school distribution of children's living regions via Bayesian Bootstrapping. Overall, we perform a total of  $M = 50$  imputations, where for each imputation, we first impute the missing regional link information and then perform the imputation for missing individual specific values. Given  $M = 50$  imputed data sets, we use combining rules for asymptotic normally distributed estimators in order to arrive at a correct assessment of estimation uncertainty, resulting in valid parameter inference. In order to assess relative fit of different (nested) model specifications, we refer to likelihood ratio tests and use the median pooling rule as suggested by Eekhout, van de Wiel, and Heymans (2017). We do so, as alternative combining rules for likelihood ratio test statistics currently discussed in the literature do not apply in the considered hierarchical context (see Chan & Meng, 2019). All computations have been conducted using R and Stata, where the R package MICE has been adapted to provide the multiple imputations, whereas the hierarchical binary logistic regressions have been conducted in Stata using the xtlogit and gllamm commands respectively. Combining rules to provide the final estimates are based on own R code. All codes for data handling and estimation are available from the authors upon request.

Note that we apply weighting within the considered hierarchical binary regression analysis only to check for the larger proportion of high school attendants (46.9) compared to the quota of 44.0 percent reported for school year 2016/2017 (see Statistisches Bundesamt, 2018). In contrast to a linear regression, where

weights provide one way to account for heteroscedasticity, variance and mean parameter estimation are interrelated in binary regression analysis. Thus, no theoretical argument ensures that use of weights would help to discover the true regression relationship. The robustness is checked via assigning an individual weight factor of 0.820 to high school attendants and an individual weight of 1 to all others. These weight factors imply a weighted sample quota for high school attendance of 42.0 percent. Corresponding weighted regression estimates do not differ from unweighted regression analysis, see discussion of empirical results below.

## 6. Empirical results

The analysis of the individual decision to enter a high school is performed in two steps. The first step investigates the dependence between the degree of urbanization and high school attendance rates, as shown in Table 4. The results show higher attendance rates in urbanized counties (51.6 percent and 52.7 percent) compared to rural counties (38.9 percent and 33.0 percent). Looking at regional differences

**Table 4:** High school attendance and regional structures

	high school attendance grade 5 (school year 2009/2010)	
	yes	no
type of county		
metropolises	51.6 %	48.4 %
urban counties	52.7 %	47.3 %
rural counties with some agglomeration	38.9 %	61.1 %
rural counties with a sparse population	33.0 %	67.0 %
proportion of high schools to all secondary schools		
under average	45.6 %	54.4 %
average	46.1 %	53.9 %
above average	83.2 %	16.8 %
unemployment rate		
under average	48.2 %	51.9 %
average	51.0 %	49.0 %
above average	32.1 %	67.9 %
proportion of highly qualified employee		
under average	28.5 %	71.6 %
average	43.3 %	56.7 %
above average	51.2 %	48.8 %

*Notes.* Source: NEPS SC3 (doi:10.5157/NEPS:SC3:6.0.1), n = 4,880, imputed data, own calculations. Due to rounding differences, deviations in the sum of the individual values are possible.

in terms of fraction of high schools to all schools at the county level, we also find a strong relationship with attendance rates, as an above average supply is accompanied by higher attendance rates. The relationship between unemployment rates and school attendance rates is not in line with the expected *ceteris paribus* effect, thus pointing at general structural differences at the county level. The proportion of highly qualified employees is again in line with the expected effects with a higher fraction of highly qualified employees coinciding with a higher attendance rate for high school.

The second step of the analysis explicitly focusses on the individual level. Using hierarchical logistic regressions, we assess the effects of regional infrastructure on high school attendance. We consider five model specifications, denoted as I, IIa-c, III, IV and V. Model specification I considers no control variables at all, serving thus as the rudimentary benchmark specification allowing for nested model comparison, while specification II considers the regional infrastructure. Model specification IIa considers the type of county, whereas specification IIb considers the regional school offer, and specification IIc involves the regional labor market conditions. A joint consideration of the regional infrastructure variables is provided in model specification III. Model specification IV relates the individual background characteristics to the observed decision to enter a high school. V is the full model specification, involving all considered variables jointly. The corresponding estimates are provided in Table 5 with corresponding confidence intervals given in Table 6.

The results from model specification I highlight the regional clustering structure of the sample arising from the underlying sampling of schools. This implies a correlation of individuals at the county level of 0.72. Consequently, we would like to point out that our estimation results are expected to be more conservative compared to results that would arise from an individual sample not clustered within schools. Further, model specification I serves as a reference point to assess the joint significance of the considered explaining factors. We provide pairwise comparison of model specification I vs. IIa, I vs. IIb, I vs. IIc, I vs. III, I vs. IV, I vs. V, and IV vs. V. The results on the corresponding log likelihood values and likelihood ratio tests are given in Table 5. Overall, the comparison of model specification I with model specifications IIa, IIb, and IIc reveals that each set of regional factors is jointly significant in predicting the binary dependent variable. The same holds when considering all regional indicators jointly in model specification III. Comparing models I and IV indicates a strong increase in model fit arising from the set of individual control variables. However, the regional indicators add to this as indicated by the corresponding likelihood ratio test comparing model specifications IV and V with a median *p* value of .001. This confirms that the regional indicators are jointly significant predictors of the individual educational decision to enter a high school.

The effects of the regional infrastructure shown in model specifications IIa, IIb, IIc, and III are in line with the theoretical expectations. Regressing the individual decision on the type of county classification reveals that the probability of at-

Table 5: Multivariate models for the probability of high school attendance – parameter estimates, log-likelihood values, and  $p$  values

	I est. coef.	IIa est. coef.	IIb est. coef.	IIIc est. coef.	III est. coef.	IV est. coef.	V est. coef.
type of county (ref. metropolises)							
urban counties	-1.850	0.168			1.979	0.132	0.074
rural counties with some agglomeration	-5.066	0.015			-0.040	0.494	-1.493
rural counties with a sparse population	-6.713	0.002			-0.414	0.442	-1.711
proportion of high schools to all secondary schools		0.276	0.000		0.287	0.002	0.177
unemployment rate			0.326	0.091	-0.154	0.306	-0.237
proportion of highly qualified employee gender (ref: female)			0.460	0.005	0.067	0.385	-0.051
male							
age in years and months						-0.113	0.270
migration background (ref: no)						-0.524	0.007
yes						-0.518	0.008
reading comprehension (WLE)							
mathematical competence (WLE)							
grade in mathematics							
grade in German							
recommendation for high school (ref: no)							
yes							
idealistic educational aspiration of children for "Abitur" (ref: no)							
yes							
university degree (at least on parent ISCED 5B or higher; ref: no)							
yes							
net equivalized income per month [€]							
occupational status according to EGP classification (3 categories; ref: service class: EGP I + II)							
working class: EGP V–VII							
others: EGP III–IV							
constant	-1.648	0.000	1.989	0.090	-7.630	0.000	0.454
variance parameter of random effect for living region		4.298	0.000	4.446	0.000	4.545	0.000
median log likelihood value (number of parameters)							
	-1.471.891 (2)	-1.460.373 (5)	-1.445.840 (3)	-1.457.197 (4)	-1.443.917 (8)	-751.123 (15)	-736.956 (21)

Notes. Source: NEPS SC3 (doi:10.5157/NEPS:SC3:6.0.1),  $n = 4,880$ , imputed data, own calculations. Est. coef. denotes the parameter estimate based on multiple imputation;  $p$  denotes the estimated  $p$  value; df denotes the degree of freedom; SD denotes the standard deviation.

Table 6: Multivariate models for the probability of high school attendance – estimated confidence intervals

	I 95 % CI	IIa 95 % CI	IIb 95 % CI	IIc 95 % CI	III 95 % CI	IV 95 % CI	V 95 % CI
type of county (ref. metropolises)							
urban counties	[-5.621;1.922]				[-1.799;5.756]		[-3.394;3.541]
rural counties with some agglomeration	[-9.655;-0.477]				[-5.035;4.955]		[-5.824;2.838]
rural counties with a sparse population	[-11.233;-2.193]				[-5.937;5.108]		[-6.709;3.287]
proportion of high schools to all secondary schools	[0.135;0.417]				[0.092;0.483]		[0.026;0.327]
unemployment rate							
proportion of highly qualified employee gender (ref: female)							
male							
age in years and months							
migration background (ref: no)							
yes							
reading comprehension (WLE)							
mathematical competence (WLE)							
grade in mathematics							
grade in German							
recommendation for high school (ref: no)							
yes							
idealistic educational aspiration of children for "Abitur" (ref: no)							
yes							
university degree (at least on parent ISCED 5B or higher; ref: no)							
yes							
net equivalized income per month [€]							
occupational status according to EGP classification (3 categories; ref: service class: EGP I + II)							
working class: EGP V – VII							
others: EGP III – IV							
constant	[-2.215;-1.082]	[-0.915;4.892]	[-11.378;-3.881]	[-12.942;-2.381]	[-15.957;0.307]		
variance parameter of random effect for living region	[3.846;4.751]	[3.947;4.944]	[4.083;5.007]	[3.933;4.944]	[4.070;5.035]	[3.255;4.124]	[3.341;4.266]

Notes. Source: NEPS SC3 (doi:10.5157/NEPS:SC3:6.0.1), n = 4,880, imputed data, own calculations. 95 % CI denotes the estimated 95 % confidence interval in terms of the 2.5 % and 97.5 % quantile.

tending a high school is significantly lower for rural counties. Also, as indicated by specification IIb, we find a positive influence of the proportion of high schools to all secondary schools in a county on the probability of attending a high school. With regard to regional labor market conditions, we find positive effects for both the regional unemployment rate and the share of highly qualified employees in the regional labor force. Bringing all regional infrastructure variables together in model specification III reveals that, given the substantial level of multicollinearity among the regional variables, only the regional school offer remains a significant determinant for the individual school decision.

When considering individual characteristics, the estimates for model specification IV are also in line with the theoretical expectations. Higher values in mathematical and reading competencies are associated with a higher probability of attending high school. Similar effects are documented for grades in mathematics and German, with worse grades reducing the probability of ending up in high school. We also find positive significant effects for educational aspiration and the school recommendation at the end of elementary schooling. Moreover, the effect of a migration background is in line with theoretical expectations and empirical results of other studies. As we control for the individual competence level, a migration background increases educational aspirations *ceteris paribus*, and thus the probability of attending high school. Positive significant effects are also found for the educational background of parents, where an university degree, as indicated by ISCED classification code 5b or higher, increases the probability of attending a high school. Furthermore, a higher household income is strongly associated with a higher probability of high school attendance. The documented effect for age reveals that fifth-graders at high schools are on average younger than fifth-graders not attending high schools. We interpret this age effect as being related to developmental tempo. Children who develop more slowly may either enter school later, stay longer at primary school, or may even change from higher school tracks to lower schools and repeat grades. Note that no significant difference between boys and girls is documented, and this also holds for occupational status as provided in terms of the EGP-class index.

The results from the full specification V reveal the robustness of the estimated effects obtained in specification III and IV respectively. This indicates that the regional school offer is still an important determinant of the individual educational decision regarding secondary schooling, even nowadays. Furthermore, it points at necessary extensions allowing to consider further operationalizations of the regional school offer more closely related to the costs implied by specific school choice, e.g. distance to schools. Note that we have checked the robustness of the documented regression relations via using individual weights in model specifications IV and V, see Tables 7 and 8 respectively. The corresponding results reveal no substantial difference compared to the unweighted regression analyses and thus confirm the still present effect of school infrastructure on the individual educational decision to enter a high school.

The influence of regional school infrastructure and labor market conditions |

Table 7: Multivariate models for the probability of high school attendance – parameter estimates, log-likelihood values, and p-values for weighted regressions

	IV weighted		V weighted	
	est. coef.	p	est. coef.	p
type of county (ref. metropolises)				
urban counties	0.911	0.307		
rural counties with some agglomeration	-0.196	0.465		
rural counties with a sparse population	-0.202	0.465		
proportion of high schools to all secondary schools	0.162	0.015		
unemployment rate	-0.166	0.283		
proportion of highly qualified employee	0.040	0.420		
gender (ref: female)				
male	-0.189	0.169	-0.163	0.202
age in years and months	-0.496	0.011	-0.541	0.007
migration background (ref: no)				
yes	Nested model comparison via likelihood ratio tests (median p / df / SD)			
reading comprehension (WLE)	median p	df	SD	
mathematical competence (WLE)	0.001	6	0.251	
grade in mathematics	0.600	0.014	0.836	0.002
grade in German	0.205	0.016	0.216	0.014
recommendation for high school (ref: no)				
yes	0.1045	0.000	0.967	0.000
idealistic educational aspiration of children for "Abitur" (ref: no)	-0.388	0.002	-0.437	0.001
university degree (at least on parent ISCED 5B or higher; ref: no)	-0.657	0.000	-0.658	0.000
net equivalized income per month [€]	2.056	0.000	2.128	0.000
occupational status according to EGP classification (3 categories; ref: service class: EGP I + II)				
working class: EGP V–VII	3.551	0.000	3.675	0.000
others: EGP III–IV	0.518	0.014	0.507	0.019
constant	0.117	0.097	0.114	0.120
variance parameter of random effect for living region	-0.153	0.283	-0.079	0.388
median log likelihood value (number of parameters)	-0.187	0.236	-0.258	0.164
	3.395	0.068	-0.041	0.496
	4.722	0.000	5.097	0.000
	-706.365 (15)		-693.973 (21)	

Notes. Source: NEPS SC3 (doi:10.5157/NEPS;SC3:6.0.1), n = 4,880, imputed data, own calculations, effective sample size for weighted data is 4,469. Est. coef. denotes the parameter estimate based on multiple imputation; p denotes the estimated p value; df denotes the degree of freedom; SD denotes the standard deviation.

Table 8: Multivariate models for the probability of high school attendance – estimated confidence intervals for weighted regressions

	IV weighted 95 % CI	V weighted 95 % CI
type of county (ref. metropolises)		
urban counties	[-2.629;4.452] [-4.591;4.199] [-4.717;4.312]	
rural counties with some agglomeration		
rural counties with a sparse population		
proportion of high schools to all secondary schools	[0.015;0.309] [-0.733;0.401] [-0.345;0.424]	
unemployment rate		
proportion of highly qualified employee		
gender (ref: female)		
male	[-0.575;0.197] [-0.920;-0.071] [-0.973;-0.109]	
age in years and months		
migration background (ref: no)		
yes	[0.062;1.138] [0.018;0.392] [0.024;0.407] [0.788;1.302] [0.709;1.125]	
reading comprehension (WLE)		
mathematical competence (WLE)		
grade in mathematics	[-0.648;-0.128] [-0.942;-0.372]	
grade in German		
recommendation for high school (ref: no)		
yes	[1.632;2.480] [1.652;2.604]	
idealistic educational aspiration of children for “Abitur” (ref: no)		
yes	[2.828;4.275] [2.945;4.406]	
university degree (at least on parent ISCED 5B or higher; ref: no)		
yes	[0.053;0.983] [-0.060;0.293] [0.029;0.985] [-0.076;0.304]	
net equivalized income per month [€]		
occupational status according to EGP classification (3 categories; ref: service class: EGP I + II)		
working class: EGP V – VII	[-0.676;0.370] [-0.619;0.461]	
others: EGP III – IV	[-0.698;0.324] [-0.776;0.259]	
constant	[-1.130;8.320] [4.080;5.364]	
variance parameter of random effect for living region		
		[-8.850;8.769] [3.973;6.221]

Notes. Source: NEPS SC3 (doi:10.5457/NEPS;SC3;6.0;1),  $n = 44880$ , imputed data, own calculations, effective sample size for weighted data is 4469, 95 % CI denotes the estimated 95 % confidence interval in terms of the 2.5 % and 97.5 % quantile.

## 7. Conclusion

In this paper we analyze the role of regional disparities on educational opportunities. In empirical research this topic has been discussed irregularly in Germany since the 1960ies. The obtained results have shown that children from rural areas are systematically worse off compared to children in urban areas. An important factor explaining this disadvantage has been the relative lack of educational infrastructure in rural areas as especially Sixt (2010, 2013) has shown with analysis of the GSOEP regarding educational decisions in the years from 1990 to 2006. No effects are found for labor market structures. Unfortunately, there was no possibility to control some of the most relevant variables such as competencies or grades at the individual level in these former studies and especially for labor market structures the questions about the appropriate context measuring the effect is open. Hence, the paper at hand tries to replicate the results with more recent and comprehensive data provided by NEPS SC3, where the focus is set on the transition to secondary education in the school year 2010/2011.

First, the findings of this paper confirm that children from rural areas in Germany are disadvantaged in terms of educational opportunities. Operationalizing regional school infrastructure analogous to Sixt (2010, 2013) via the proportion of high schools to all secondary schools in a county, it is revealed that the rural effect can be explained by the school infrastructure and that the probability of a transition to high school is lower the lower the proportion of high schools in a county is. Although there is a conceptual difference in the sample definition (only first borns and birth cohorts from the western part of Germany are considered by Sixt, 2010, 2013) we can obtain the same results in this first step. Second, no significant effects can be found for labor market indicators given the operationalization of regional labor market conditions at hand. This operationalization differs from Sixt (2010, 2013) and should help to overcome the criticism that in administrative regions, i.e. "Raumordnungsregionen" are possibly not the relevant regional context when parents think about labor market benefits of different school tracks. But also the consideration of commuter flows does not lead to substantial effects. It remains still an open question whether this parental decision is not influenced by labor market structures at all or whether the assessment of the regional context of the regional labor market is still not adequate. Future research could re-examine these effects and consider alternative operationalizations, e.g. subjectively defined relevant labor markets without reference to administrative boarders. However, currently no appropriate database is available to address these issues. Third, the most interesting part of the results is whether the regional effect of the school infrastructure still persists when important control variables such as competencies, grades, school recommendations, and educational aspirations at the individual level are considered. Note that these were missing in the former studies using the GSOEP data. The analyses at hand with the NEPS SC3 reveal a persistent effect indicating

that the regional school offer is a further important determinant of the individual educational decision regarding secondary schooling still today.

Even though the NEPS SC3 data are the most adequate data at the moment, some caveats and restrictions apply. First, the sample size and the two stage sampling procedure of NEPS SC3 causes the individual sample to be nested within schools, with typically no more than two schools in a county. Therefore, the informational content on variation at the regional level is systematically lower than it would be in an individual sample not nested at the school level. This may cause an underestimation of the effects of regional infrastructure indicators relative to an investigation carried out with an individual sample. A second limitation arises from surveying individual characteristics at the beginning of grade five, i.e. retrospective to the transition of interest. This may be crucial at least for measured educational aspirations, which need to be better measured in the years before the transition to allow for disentangling the infrastructure effect and the effect of high aspirations. Unfortunately, no data with prospectively measured controls are available yet which would allow for such an investigation of the transition processes to high schools. As addressed above, due to missing heterogeneity, NEPS SC2 does not provide a better data basis either. Furthermore, the database at hand is plagued by high fraction of missing information with regard to the county of residence of the students (see section 4) surveyed within the parental interviews.

A further issue, already discussed by Sixt (2010, 2013) relates to the fact that residential segregation cannot be controlled for. This means that we have no information on the spatial distribution of families with, and without, access to supportive educational resources in the counties. Possibly, families with access to supportive resources may live nearer to high schools than those with less. In this case, the presented result would not be an effect of infrastructure, but of residential segregation. To overcome this problem, it would be necessary to directly measure and operationalize distances to schools and control for the place of residence. Sixt (2018) provides a first study, albeit regionally limited to rather rural regions, and shows that after controlling for residential segregation, an effect of school infrastructure on educational decisions is still present. Furthermore, a closer look on smaller spatial contexts would be necessary as the school infrastructure as well as other infrastructure like public transport differs systematically between rural and urban areas and even worse between city districts (e.g. Hauf, 2007).

Despite these restrictions, evidence with regard to still existing regional inequalities in educational opportunities connected to regional disparities in school provision is strong. As the National Educational Report shows, the number of existing schools decreased, especially in rural areas, due to socio-demographic change (Autorengruppe Bildungsberichterstattung, 2016, p. 32), thus further contributing to regional educational inequality. Even if the trend towards declining birth rates, declining student numbers (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2018), and thus school closures (Weishaupt, 2006) seems to have come to an halt, the observation at hand is alarming as the structural inequalities observed in the 1960ies still pre-

vail. Although education was, and is, still in the competence of the federal states, with the municipalities being responsible for school planning, regional disparities in educational offerings can be managed to a certain extent. It seems to be unrealistic to offer a two or three tracked school system so widespread that at an individual level the physical journeys to schools do not lead to systematically varying distances.

The equivalence of living conditions plays a prominent role in the current political debate, and is even enshrined in Article 72 of the Basic Law. This raises the question of other options for decoupling spatial access to educational institutions from access to educational opportunities. Therefore, regional opportunity structures like educational offerings should be more present in current political as well as scientific debate.

## Acknowledgments

The authors thank Stef van Buuren for discussing the applied imputation strategy at the Aschauer Soiree, held at Glücksburg castle, Germany, 15th to 17th November 2013, Tobias Koberg for matching the regional information with the NEPS data, and Thorsten Schnapp for realizing smooth interactions of Stata and R code. Furthermore, the authors acknowledge support from the Oberfrankenstiftung under grant P.Nr. 04307, and the Deutsche Forschungsgesellschaft (DFG) within priority program SPP 1646 under grants AS 368/3-1 and AS 368/3-2. Finally, we thank two anonymous reviewers and the editors for their helpful comments and suggestions that helped to improve the manuscript considerably.

## References

- Aßmann, C., Steinhauer, H. W., Kiesl, H., Koch, S., Schönberger, B., Müller-Kuller, A., Rohwer, G., Rässler, S., & Blossfeld, H.-P. (2011). Sampling design of the National Educational Panel Study: Challenges and solutions. In H.-P. Blossfeld, H.-G. Roßbach & J. von Maurice (Eds.), *Education as a lifelong process. The German National Educational Panel Study (NEPS)* (pp. 87–113). Wiesbaden: VS.
- Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: wbv.
- Bartels, J. (1975). *Interdependenzen zwischen dem Angebot und der Nachfrage nach Plätzen der Sekundarschulstufe bei Angehörigen unterschiedlicher Schichtzugehörigkeit in einer Großstadt* (unpublished). Freie Universität Berlin, Fachbereichs Gesellschafts- und Planungswissenschaften.
- Bartl, W., & Sackmann, R. (2014). Path dependency, demographic change, and the (de-)differentiation of the German secondary school system. *Zeitschrift für Soziologie*, 43(1), 50–69.
- Baumert, J., Trautwein, U., & Artelt, C. (2003). Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens. In J. Baumert et al. (Eds.), *PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (pp. 261–331). Wiesbaden: VS.

- Baur, R. (1972). *Elternhaus und Bildungschancen. Eine Untersuchung über die Bedeutung des Elternhauses für die Schulwahl nach der 4. Klasse Grundschule.* Weinheim: Beltz.
- Becker, R., & Lauterbach, W. (2016). *Bildung als Privileg: Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit.* Wiesbaden: VS.
- Blossfeld, H.-P. (1988). Bildungsverläufe im historischen Wandel: Eine Längsschnitt-analyse über die Veränderungen der Bildungsbeteiligung im Lebenslauf dreier Geburtskohorten. In H.-J. Bodenhöfer (Ed.), *Bildung, Beruf, Arbeitsmarkt* (pp. 259–302). Berlin: Duncker & Humblot.
- Blossfeld, H. P., Blossfeld, G. J., & Blossfeld, P. N. (2019). Soziale Ungleichheiten und Bildungsentscheidungen im Lebensverlauf. Die Perspektive der Bildungssoziologie. *Journal for Educational Research Online*, 11(1), 16–30.
- Blossfeld, H.-P., Maurice, J. von, & Schneider, T. (2011). The National Educational Panel Study: Need, main features, and research potential. *Zeitschrift für Erziehungswissenschaft*, 14(2), 5–17.
- Bolder, A. (1975). *Strukturelle und sozioökonomische Determinanten des Ausbildungswegs. Zu den Auswirkungen horizontaler und vertikaler Disparitäten im Bildungsbereich.* (Dissertation). Köln: Universität zu Köln.
- Bolder, A. (1984). Regionale Brechungen des Zusammenhangs zwischen regionalen Umwelten und individuellen Ausbildungschancen. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 17(3), 407–418.
- Boudon, R. (1974). *Education, opportunity, and social inequality. Changing prospects in western society.* New York, NY: Wiley.
- Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials. Towards a formal rational action theory. *Rationality and Society*, 9(3), 275–305.
- Buchholz, S., & Pratter, M. (2017). Wer profitiert von alternativen Bildungswegen? Alles eine Frage des Blickwinkels! Eine systematische Rekonstruktion des Effektes sozialer Herkunft für alternative Wege zur Hochschulreife. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69(3), 409–443.
- Butler, J., & Moffitt, R. (1982). A computationally efficient quadrature procedure for the one-factor multinomial probit model. *Econometrica*, 50(3), 761–764.
- Chan, K. W., & Meng, X.-L. (2019). Multiple improvements of multiple imputation likelihood ratio tests. arXiv:1711.08822v2.
- Clausen, M. (2006). Warum wählen Sie genau diese Schule? Eine inhaltsanalytische Untersuchung elterlicher Begründungen der Wahl der Einzelschule innerhalb eines Bildungsgangs. *Zeitschrift für Pädagogik*, 52(1), 69–90.
- Clausen, M. (2008). *Einzelschulwahl. Zur Wahl der konkreten weiterführenden Einzelschule aus Sicht von Bildungsnachfragenden und Bildungsanbietern.* (Habilitationsschrift). Mannheim: Universität Mannheim, Fakultät für Sozialwissenschaften.
- Ditton, H., Elsäßer, S., Götz, N., Stahn, V., & Wohlkinger, F. (2017). Soziale Disparitäten im Bildungsverlauf. Schulische Laufbahnen von der 2. bis zur 7. Jahrgangsstufe. In H. G. Holtappels (Ed.), *Entwicklung und Qualität des deutschen Schulsystems: Neue empirische Befunde und Entwicklungstendenzen* (pp. 143–171). Münster: Waxmann.
- Eekhout, I., van de Wiel, M. A. & Heymans, M. W. (2017). Methods for significance testing of categorical covariates in logistic regression models after multiple imputation: power and applicability analysis. *BMC Medical Research Methodology*, 17(1), 129, <https://doi.org/10.1186/s12874-017-0404-7>.
- Erikson, R., & Jonsson, J. O. (1996). *Can education be equalized? The Swedish case in comparative perspective.* Boulder, San Francisco: Westview.
- Esser, H. (1999). *Soziologie. Spezielle Grundlagen. Band 1: Situationslogik und Handeln.* Frankfurt: Campus.

- Fickermann, D. (1997). Soziale Aspekte der Bildungsbeteiligung. In P. Zedler & H. Weishaupt (Eds.), *Kontinuität und Wandel. Thüringer Schulen im Urteil von Schülern, Lehrern und Eltern* (pp. 147–168). Weinheim: Deutscher Studienverlag.
- Gehrman, S. (2019). *Aspirationen, kulturelles Kapital und soziale Herkunft*. Wiesbaden: VS.
- Georg, W. (2006) *Soziale Ungleichheit im Bildungssystem. Eine empirisch-theoretische Bestandsaufnahme*. Konstanz: UVK.
- Greene, W. (2004). The behaviour of the maximum likelihood estimator for limited dependent variable models in the presence of fixed effects. *Econometrics Journal*, 7, 98–119.
- Gresch, C. (2012). *Der Übergang in die Sekundarstufe I. Leistungsbeurteilung, Bildungsaspiration und rechtlicher Kontext bei Kindern mit Migrationshintergrund*. Wiesbaden: VS.
- Hansen, R. (1993). *Quantitative Entwicklungen und strukturelle Veränderungen der Schule in der BRD*. Dortmund: Projekt Verlag.
- Hauf, T. (2007). Innerstädtische Bildungsdisparitäten an der Übergangsschwelle von den Grundschulen zum Sekundarschulsystem. *Zeitschrift für Pädagogik*, 53, 299–313.
- Helbig, M., & Schneider, T. (2014). *Auf der Suche nach dem katholischen Arbeitermädchen vom Lande*. Wiesbaden: VS.
- Henz, U., & Maas, I. (1995). Chancengleichheit durch die Bildungsexpansion? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 47, 605–633.
- Hillmert, S., & Jacob, M. (2005). Die zweite Chance im Schulsystem? Zur sozialen Selektivität bei “späteren” Bildungsscheidungen. In A. P. Berger & H. Kahlert (Eds.), *Institutionalisierte Ungleichheiten. Wie das Bildungswesen Chancen blockiert* (pp. 155–179). Weinheim: Juventa.
- Hillmert, S., Hartung, A., & Weßling, K. (2017). A decomposition of local labour-market conditions and their relevance for inequalities transitions to vocational training. *European Sociological Review*, 33(4), 534–550.
- Kristen, C., & Dollmann, J. (2009). Sekundäre Effekte der ethnischen Herkunft: Kinder aus türkischen Familien am ersten Bildungsübergang. *Zeitschrift für Erziehungswissenschaft*, 12, 205–229.
- Nowey, W. (1983). *Das Bildungsverhalten in den urbanen, urbanisierten und ländlichen Kleinräumen Bayerns. Schulwege, Übertritte an weiterführende Schulen, Schülerneigungen und -leistungen, Lehrereinstellungen*. Mering: W. Nowey.
- Peisert, H. (1967). *Soziale Lage und Bildungschancen in Deutschland*. München: Piper.
- Relikowski I. (2012). *Primäre und sekundäre Effekte am Übertritt in die Sekundarstufe I*. Wiesbaden: VS.
- Riedel, A., Schneider, K., Schuchart, C., & Weishaupt, H. (2010). School choice in German primary schools: How binding are school districts? *Journal for Educational Research Online*, 2(1), 94–120.
- Rösner, E. (2005). Von erfolgreichen Verlierern und verlustreichen Absteigern. In H. G. Holtappels (Ed.), *Schulentwicklung und Schulwirksamkeit. Systemsteuerung, Bildungschancen und Entwicklung der Schule: 30 Jahre Institut für Schulentwicklungsforschung* (pp. 131–139). Weinheim: Juventa.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581–592.
- Schindler, S. (2015). Soziale Ungleichheit im Bildungsverlauf – alte Befunde und neue Schlüsse? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 67(3), 509–537.
- Schindler, S. (2017). School tracking, educational mobility and inequality in German secondary education: Developments across cohorts. *European Societies*, 19(1), 28–48.
- Schnorr-Bäcker, S. (2014). Regionalstatistische Daten für Deutschland, Europa und weltweit. *Zeitschrift für amtliche Statistik*, 3, 12–27.

- Schuchart C. (2010) Warum interessieren sich Hauptschülerinnen und Hauptschüler für einen Realschulabschluss? Eine Analyse individueller Überzeugungen unter besonderer Beachtung geschlechts- spezifischer Differenzen. In J. Baumert, K. Maaz, U. Trautwein (Eds.), *Bildungsentscheidungen* (pp. 373–397). Wiesbaden: VS.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2018). Vorausberechnung der Schüler- und Absolventenzahlen 2016 bis 2030. *Statistische Veröffentlichungen der Kultusministerkonferenz Nr. 213*. Berlin.
- Shavit, Y., & Blossfeld, H.-P. (1993). *Persistent inequality. Chancing educational attainment in thirteen European countries*. Boulder: Westview.
- Simon, H. A. (1993). *Homo rationalis. Die Vernunft im menschlichen Leben*. Frankfurt am Main: Campus.
- Sixt, M. (2010). Regionale Strukturen als herkunftsspezifische Determinanten von Bildungsentscheidungen. (Dissertation). Kassel: Universität Kassel. Retrieved from <https://kobra.bibliothek.uni-kassel.de/handle/urn:nbn:de:hebis:34-2010110934909>
- Sixt, M. (2013). Wohnort, Region und Bildungserfolg. Die strukturelle Dimension bei der Erklärung von regionaler Bildungsungleichheit. In R. Becker & A. Schulze (Eds.), *Bildungskontexte. Strukturelle Voraussetzungen und Ursachen ungleicher Bildungschancen* (pp. 457–481). Wiesbaden: VS.
- Sixt, M. (2018). Wahl der Schulart und Schulinfrastruktur. Die Bedeutung der räumlichen Verteilung von weiterführenden Schulen bei der Wahl der Schulart am Übergang von der Primar- in die Sekundarstufe. In M. Sixt, M. Bayer, & D. Müller (Eds.), *Bildungsentscheidungen und lokales Angebot. Die Bedeutung der Infrastruktur für Bildungsentscheidungen im Lebensverlauf* (pp. 87–113). Münster: Waxmann.
- Stanat P., Bergann S., & Taraszow T. (2016). Geschlechtsbezogene Disparitäten im deutschen Bildungswesen. In R. Tippelt & B. Schmidt-Hertha (Eds.), *Handbuch Bildungsforschung. Springer Reference Sozialwissenschaften* (pp. 1321–1338). Wiesbaden: VS.
- Statistisches Bundesamt. (2018). *Schulen auf einen Blick*. Retrieved from [https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft\\_derivate\\_00035140/Schulen\\_auf\\_einen\\_Blick\\_2018\\_Web\\_bf.pdf;jsessionid=5BBFAA19E06C8B05F31D4EF0E0326230](https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft_derivate_00035140/Schulen_auf_einen_Blick_2018_Web_bf.pdf;jsessionid=5BBFAA19E06C8B05F31D4EF0E0326230)
- Steinhauer, H.W., Aßmann, C., Zinn, S., Goßmann, S., & Rässler, S. (2015). Sampling and weighting cohort samples in institutional contexts. *AStA Wirtschafts- und Sozialstatistisches Archiv*, 9(2), 135–157.
- Stocké, V. (2012). Das Rational-Choice Paradigma in der Bildungssoziologie. In U. Bauer, U. H. Bittlingmayer, & A. Scherr (Eds.), *Handbuch Bildungs- und Erziehungssoziologie* (pp. 423–436). Wiesbaden: VS.
- Trommer-Krug, L. (1980). Soziale Herkunft und Schulbesuch. In Projektgruppe Bildungsbericht. (Ed.), *Bildung in der Bundesrepublik Deutschland: Daten und Analysen Band 1* (pp. 217–281). Stuttgart: Klett-Cotta.
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Multivariate imputation by chained equations. *Journal of Statistical Software*, 45, 1–67.
- Weishaupt, H. (2006). Veränderungen im elementaren und sekundären Bildungsbereich durch demographischen Wandel. In DESTATIS. (Ed.), *Statistik und Wissenschaft. Demographischer Wandel – Auswirkungen auf das Bildungssystem* (pp. 1–17). Wiesbaden: Statistisches Bundesamt.
- Wu, J. (2007). *Scale and scaling: a cross-disciplinary perspective. Key topics of landscape ecology*. Cambridge: University Press.
- Zielonka, M. (2017). *Revision und Stabilisierung von Schulwahlentscheidungen in der Sekundarstufe 1*. Berlin: Logos.

Florian Wohlkinger & Michael Bayer

## Typologie des Schülerhabitus in der Sekundarstufe I: Ein quantitativer Operationalisierungsansatz mit Daten des Nationalen Bildungspanels

### Zusammenfassung

Im Rahmen einer qualitativen Studie entwickelten Kramer et al. (2009) eine Typologie des Schülerhabitus, die an die Arbeiten von Bourdieu anknüpft und auf die Analyse der Reproduktion sozialer Ungleichheit im Bildungsbereich abzielt. In quantitativen Untersuchungen blieb dieser Ansatz bislang jedoch weitgehend unberücksichtigt. Mit dem vorliegenden Beitrag versuchen wir eine Verbindung herzustellen und formulieren mithilfe von Daten des Nationalen Bildungspanels einen Vorschlag für eine quantitative Abbildung des Schülerhabituskonzepts. In einer Reihe von Latent-Class-Analysen mit ausgewählten Variablen zeigte sich, dass es sowohl formal geeignete als auch inhaltlich sinnvoll interpretierbare Lösungen gibt, die verschiedene Schülerhabitustypen beschreiben und zudem Zusammenhänge mit weiteren ungleichheitsrelevanten Merkmalen aufweisen. Daneben werden aber auch die Grenzen einer solchen Analytik deutlich, die sich aus der anspruchsvollen Hintergrundtheorie und der Breite der Indikatoren ergeben.

### Schlagworte

Schülerhabitus; Bourdieu; Latent-Class-Analyse; Soziale Ungleichheit

---

Dr. Florian Wohlkinger (corresponding author), Lehrstuhl für allgemeine Pädagogik, Erziehungs- und Sozialisationsforschung, Ludwig-Maximilians-Universität München, Martiusstr. 4, 80802 München, Deutschland  
E-Mail: [florian.wohlkinger@edu.lmu.de](mailto:florian.wohlkinger@edu.lmu.de)

Prof. Dr. Michael Bayer, Evangelische Hochschule Nürnberg, Bärenschanzstraße 4, 90429 Nürnberg, Deutschland & Leibniz-Institut für Bildungsverläufe e.V. (LIfBi), Bamberg, Deutschland  
E-Mail: [michael.bayer@lifbi.de](mailto:michael.bayer@lifbi.de)

# Typology of students' academic habitus in lower secondary education: A quantitative operationalization approach with data from the National Educational Panel Study

## Abstract

*With a qualitative approach based on Bourdieu's ideas, Kramer et al. (2009) developed a typology of students' academic habitus in order to analyze the reproduction of social inequality in education. So far, however, their approach has been widely disregarded in quantitative studies. With this paper, we aim at connecting both perspectives and propose an approach to operationalize the concept of students' academic habitus with quantitative data from the German National Educational Panel Study. In a series of latent-class-analyses with selected variables, we found formally adequate and meaningfully interpretable models, which describe distinct types of students' academic habitus and show connections to further variables with relevance for educational inequality. However, we also found limits in our analytical approach. These limits are a product of a sophisticated background theory and a broad range of indicators.*

## Keywords

*Students' academic habitus; Bourdieu; Latent-Class-Analysis; Social inequality*

## 1. Einleitung

Die Bildungsforschung befasst sich mit Fragen der Ungleichheitsreproduktion oftmals im Rahmen eines *Statuserwerbsmodells*, wie es etwa von Boudon (1974) mit der Unterscheidung von primären und sekundären Herkunftseffekten programmatisch ausgearbeitet wurde. Als eine zweite Traditionslinie finden sich daneben Modelle, welche die Statusfrage als eine der *Zuweisung* untersuchen (vgl. hierzu auch Hofstetter, 2017, S. 272). In dieser Traditionslinie finden sich insbesondere Arbeiten, die auf Pierre Bourdieu rekurrieren. Bourdieu beschäftigte sich seit den 1960er Jahren insbesondere mit dem französischen Schulsystem und brachte seine Forschungen Anfang der 1970er Jahre zusammen mit Jean-Claude Passeron auf den diagnostischen Titel „die Illusion der Chancengleichheit“ (Bourdieu & Passeron, 1971). Während sich das Gros der Bildungsforschung vor allem auf der Ebene der Individuen und unter Verwendung von Ressourcenansätzen der Frage widmet, ob und auf welchen Wegen sich individuelle Ressourcen in Bildungserfolge übersetzen (vgl. Becker & Lauterbach, 2016), waren die Bildungsinstitutionen in ihrer Strukturierung und in den dort verwendeten Curricula für Bourdieu ein Produkt der herrschenden Klassen, durch welches diese ihre eigene Reproduktion absicherten (Bourdieu, 2018). Diese stärker machtsoziologisch inspirierte Perspektive auf Bildung, Bildungschancen und Bildungserfolge führt in erster Linie

zu anderen Fragen, die im Sinne eines Statuszuweisungsmodells das innerschulische Geschehen als eine Konsequenz der sozialen Strukturiertheit von Gesellschaft verstehen. Aus einer solchen Perspektive steht dann die Frage im Zentrum, wie Bildungsinstitutionen diese soziale Privilegierung sicherstellen. Anknüpfend an Bourdieus Aussage, dass die soziale Realität zweimal existiert, „in den Sachen und in den Köpfen, in den Feldern und in den Habitus, innerhalb und außerhalb der Akteure“ (Bourdieu & Wacquant, 1996, S. 161), fokussiert man mit dem Habitus kein individuelles Einstellungsprofil, sondern ein in und durch die sich kategorial unterscheidenden Herkunftsmilieus und Institutionen produziertes, sozial strukturiertes latentes Muster.

Es sind vor allem qualitativ operierende Forschungszugänge, die sich dieser Frage zuwenden und die insbesondere das Habituskonzept von Bourdieu als Schlüsselmechanismus der Ungleichheitsreproduktion stark machen (van Essen, 2013). Hierdurch konnten durchaus wichtige empirische Umsetzungen des Bourdieu'schen Ansatzes geleistet werden, etwa die vertiefende Einsicht in die Muster symbolischer Gewaltausübung, die in Bildungsinstitutionen strukturell verankert sind. Gleichzeitig konnte im Kontext dieser Forschungen gezeigt werden, dass und wie Schulen, aber auch Lehrkräfte selbst ihren Schülerinnen und Schülern mit Habituserwartungen gegenüberstehen, die beispielsweise in Schulkulturen strukturell verankert sind (vgl. Helsper, Böhme, Kramer, & Lingkost, 2001).

In den mit quantitativen Daten operierenden und auf statistisch abgesicherte Verallgemeinerbarkeit setzenden Zugängen hat insbesondere das Konzept des kulturellen Kapitals eine gewisse Prominenz erlangt (DiMaggio, 1982). Weniger genutzt wird hingegen die Habituskonzeption, was neben grundsätzlichen theoretischen Vorbehalten (Goldthorpe, 2007), die jedoch teilweise auf Missverständnissen basieren (vgl. hierzu Lizardo, 2008), auch mit Herausforderungen bzgl. der Operationalisierung zu tun hat. Ungeachtet der Popularität des Kulturkapitalkonzepts lässt sich der „Habitus als das ‚Kern‘- oder ‚Herzstück‘ der theoretischen Überlegungen von Pierre Bourdieu“ verstehen (Kramer, 2011, S. 46).

Es gibt jedoch mittlerweile einige Versuche, das Konzept des Habitus auch mit Mitteln der quantitativen Forschung fruchtbar zu machen, was es ermöglichen würde, Statuszuweisungsprozesse in und durch das Bildungssystem auch in Hinblick auf die Stabilitätsbedingungen der Vermittlungsmechanismen zu untersuchen. Unser Beitrag knüpft an die bereits vorhandenen Versuche an, um einen Vorschlag zu formulieren, der die in den qualitativen Zugängen genutzten typologischen Verfahren aufgreift und sie clusteranalytisch reformuliert. Damit nutzen wir die in den quantitativen Zugängen vorhandenen Verfahren, die bereits seit Langem zur Identifizierung von latenten, überindividuellen Strukturen genutzt werden (Hagenaars & Halman, 1989).

In einem ersten Schritt diskutieren wir die theoretischen Hintergründe der Bourdieu'schen Habituskonzeption sowie die darauf aufbauenden Konzepte zu einem auf das Feld der Bildung bezogenen Bildungs- bzw. Schülerhabitus, bevor wir im nächsten Schritt die vorhandenen Vorschläge zur Operationalisierung vor-

stellen und kritisch diskutieren. Im daran anschließenden empirischen Teil formulieren wir einen eigenen Operationalisierungsvorschlag, um diesen mit Daten des National Bildungspanels in einer ersten explorativen Studie hinsichtlich der Umsetzbarkeit zu prüfen. Dies wird abschließend hinsichtlich notwendiger nächster Schritte diskutiert.

## 2. Theoretische Grundlagen

### 2.1 Das Bourdieu'sche Habituskonzept

Mit dem Konzept des Habitus, der – wie Eva Barlösius anmerkt – eine Art „Hilfskonstruktion“ (Barlösius, 2011, S. 46) darstellt, unternimmt Bourdieu den Versuch eine theoretische Figur zu etablieren, mit der sich die verwunderliche Abgestimmtheit von Verhaltensweisen erklären lässt, die im Kontext ihrer jeweiligen Milieu- oder Klassenspezifität eben keine explizite Abstimmung benötigt. Der den Personen zugehörige Habitus als Erzeugungsprinzip und Klassifikationssystem auf der einen und als durch gesellschaftliche Notwendigkeiten und Zwänge strukturierte Struktur auf der anderen Seite (Bourdieu, 1982, S. 277ff.) stellt mithin eine Art Vermittlungsinstanz zwischen gesellschaftlicher Struktur und Individuum dar. Hierbei fungiert der Habitus als ein in sich und in der ihn prägenden sozialen Umwelt kohärentes System von Einstellungen, Wahrnehmungen und Orientierungen. Mit dem Habituskonzept sucht Bourdieu zu zeigen, dass und wie sich soziale Ungleichheiten in doppelter Form gesellschaftlich verankern und in der Folge auch reproduzieren.

Damit entwickelt er eine Perspektive, mit der sich zeigen lässt, dass Individuen auf zweifache Weise strukturell begrenzt sind, nämlich einerseits hinsichtlich der mit sozialen Positionen verknüpften Ressourcenunterschiede und andererseits durch die mit sozialen Positionen einhergehenden Anforderungen hinsichtlich mentaler, aber auch körperbezogener Erwartungen.

Der *primäre Habitus* entwickelt sich im sozialen Herkunftskontext eines Individuums, was seine klassen- bzw. milieubezogene Prägung und damit die Abgestimmtheit auf diesen Kontext verursacht. So ermöglicht er vor allem ein unkompliziertes Sich-Bewegen, ohne dass der Einzelne vor ständige Entscheidungsnotwendigkeiten gestellt wird.

Wechselt man die Perspektive auf die Seite der Institutionen bzw. der *Felder*, wie Bourdieu diese unter Betonung der sie prägenden relationalen Struktur nennt, dann zeigen sich die jeweiligen Habitusanforderungen, welche durch die Institutionen ausgeprägt werden. In Bezug auf die Schule zeigt Bourdieu, dass es vor allem die Kultur der Elite ist, die der Kultur der Schule nahesteht, und zwar „so nah, dass die Kinder aus einem kleinbürgerlichen [...] Milieu das nur mühsam erwerben können, was den Kindern der gebildeten Klasse gegeben ist: den Stil, den Geschmack, die Gesinnung ...“ (Bourdieu, 2018, S. 25). In den Augen Bourdieus

konzentriert sich die schulbezogene Forschung jedoch meist nur auf die augenfälligsten Aspekte der Ausübung von Privilegien bildungssystemnaher Klassen wie die elterliche Unterstützung bei Hausaufgaben, die Hinzuziehung von Nachhilfe, die Informiertheit über das Bildungs- und Ausbildungssystem oder die Beziehungen zu Lehrkräften. „In Wirklichkeit jedoch vermittelt jede Familie ihren Kindern auf eher indirekten als direkten Wegen ein bestimmtes kulturelles Kapital und ein bestimmtes Ethos, ein System impliziter und tief verinnerlichter Werte, das unter anderem auch die Einstellungen zum kulturellen Kapital und zur schulischen Institution entscheidend beeinflusst“ (Bourdieu, 2018, S. 8).

## **2.2 Bildungshabitus (als Erklärungsansatz für Bildungsungleichheiten)**

Ausgehend von dieser feldbezogenen Konzeption eines *sekundären Habitus* finden sich in der Bildungsforschung Begriffe wie *Bildungshabitus* (Grgic & Bayer, 2015; Kramer & Helsper, 2011; Niestradt & Ricken, 2014) oder auch *schulbezogener Habitus* von Kindern (de Moll, 2018). Das Bildungssystem ist, wie Kramer (2011) es im Anschluss an Bourdieu ausdrückt, ein durch eine Eigenlogik spezifiziertes Feld, welches jedoch nicht nur ein eigenständiges (und durch die Institution gewünschtes) schulisches Kapital zu definieren imstande ist, sondern welches darüber hinaus institutionelle Erwartungen hinsichtlich eines passenden Habitus in Richtung der Schülerinnen und Schüler und damit auch des familiären Lernumfeldes beinhaltet. „So entscheiden letztendlich die vorhandenen Passungsverhältnisse dieser beiden Lernorte darüber, ob die kulturellen Transferbeziehungen in der Familie dem Statuserhalt oder der Statusverbesserung förderlich sind oder nicht“ (Ecarius & Wahl, 2009, S. 17).

In einer breit angelegten qualitativen Längsschnittstudie rekonstruieren Helsper, Kramer und Kollegen (Helsper, Kramer, Thiersch, & Ziems, 2009; Kramer & Helsper, 2011) den durch das schulische Feld definierten idealtypischen sekundären Habitus und untersuchen die Passungsfähigkeiten des familiär (und letztlich klassenspezifisch) erzeugten primären Habitus von Schülerinnen und Schülern zu den institutionell erwarteten Habitus. „Dadurch stehen spezifische Schulkulturen zu sozialen Milieus in einem korrespondierenden Verhältnis der Homologie, der Nähe oder Distanz bis hin zur Abstoßung“ (Kramer & Helsper, 2011, S. 110). Unter dem Begriff des Bildungshabitus identifizieren sie Typen unterschiedlicher Habitus, welche sich entlang einer Achse von Bildungs- bzw. Schulaffinität voneinander unterscheiden lassen. Unter Bildungshabitus verstehen sie grundlegende Haltungen „mit deutlich unterschiedlichen Bezügen auf schulische Bildung, auf optionale Schulkarrieren und auf verschiedene Schulabschlüsse“ (ebd., S. 116).

Dieser mit großem interpretatorischem Aufwand verknüpften qualitativen Rekonstruktion von bildungsbezogenen Habitus stehen nur vereinzelte Versuche einer quantitativ operierenden Operationalisierung und Typologisierung gegenüber. Bourdieu selbst sieht Kognitions- und Motivationsstrukturen als konstituti-

ve Merkmale des Habitus an (Bourdieu, 1993, S. 104), da sich die soziale Herkunft eben auch in Unterschieden gegenstandsbezogener Motivation ausdrückt. Nicht weil es quasi schichtspezifische Unterschiede in der grundsätzlichen Motiviertheit gibt, sondern weil Motivationen immer auch Ergebnis sozial spezifischer Erfahrungen darstellen. Grgic und Bayer (2015) können auf dieser Grundlage zeigen, dass und wie sich familiäre Ressourcen (Kapitalien) bei Schülerinnen und Schülern in einem Bildungshabitus ausdrücken, mit dem sich zumindest Grade der Orientierung an herrschenden Bildungsnormen (hier: das Abitur) unterscheiden lassen und diese in ihren Auswirkungen auf schulische Erfolge untersuchbar werden.

## 2.3 Schülerhabitus

Bereits Anfang der 1970er Jahren befassten sich Bourdieu und Passeron (1971) mit der Entstehung sozialer Ungleichheiten im Bildungssystem. Ihrem Ansatz zufolge hängt es maßgeblich vom Herkunfts米尔ieu und den dort vorherrschenden „kul-turellen Gewohnheiten“ (Bourdieu & Passeron, 1971, S. 32) ab, inwieweit Schule den einzigen Zugang zu Bildung oder lediglich eine Fortführung bzw. Ergänzung zu familialen Bildungspraktiken darstellt. Dahinter steht der Grundgedanke, dass Familie und Schule nicht isoliert voneinander zu betrachten sind, sondern dass es die *Passung* der beiden Sphären zueinander ist, die mit systematischen Vor- bzw. Nachteilen einhergeht und damit die Ursache von Bildungsungleichheiten darstellt. Demnach stehen auf der einen Seite die Kinder, die je nach sozialer Herkunft unterschiedliche Eigenschaften (Habitus) mitbringen, während auf der anderen Seite die Institutionen des Bildungssystems stehen, welche bestimmte Erwartungen an die Kinder richten und ihre Entwicklung beurteilen und bewerten. Eltern aus privilegierten Milieus vermitteln ihren Kindern von Anfang an Wissen und sprachliche Ausdrucksfähigkeiten mit Nähe zum Schul- und Bildungssystem (Bildungssprache). Diese Kinder haben damit bereits beim Eintritt in die Schule einen Vorteil gegenüber Kindern aus unterprivilegierten Milieus, deren Vorwissen und Ausdrucksweisen eine größere Distanz zum Bildungssystem aufweisen. Hinzu kommt, dass sich die in der Schule geltenden Verhaltensnormen eher an den Gewohnheiten der privilegierten Schichten orientieren, wovon diese Kinder insbesondere bei Einschätzungen und Bewertungen durch die Lehrkräfte zusätzlich profitieren. Infolge der unterschiedlichen Anerkennung ihres Verhaltens in der Schule erscheinen den Kindern schulische Erfolge (bzw. Misserfolge) je nach sozialer Herkunft entweder als Bestätigung ihrer (fehlenden) Begabung, als Ergebnis ihrer (unzureichenden) Bemühungen oder als willkürliche Schicksalsfügung. Die soziale Herkunft wirkt sich also neben der Beeinflussung der Leistungsentwicklung der Kinder auch auf ihre Wahrnehmung von sich selbst sowie auf ihre Einstellungen und Haltungen gegenüber Bildung aus.

Die Annahmen von Bourdieu und Passeron (1971) aufgreifend werten Kramer und Kollegen (2009) Interviews mit Schülerinnen und Schülern

mithilfe der dokumentarischen Methode aus und arbeiteten eine Typologie von schul- und bildungsbezogenen Habitusformationen heraus, die sich bereits bei Grundschulkindern der vierten Jahrgangsstufe abzeichnet. „Diese frühen habituellen Haltungen zu Schule und Bildung führen zu unterschiedlichen Übergangserfahrungen, Transformationsprozessen des Orientierungsrahmens und Passungskonstellationen zur weiterführenden Schule“ (Kramer, Helsper, Thiersch, & Ziems, 2009, S. 131) und sind damit einerseits von hoher Relevanz für die Entwicklung von Bildungsverläufen sowie andererseits für dabei entstehende soziale Ungleichheiten. Es werden die folgenden vier Typen von Schülerhabitus identifiziert: der *Habitus der Bildungsexzellenz bzw. -distinktion*, der *Habitus der Strebenden*, der *Habitus der Bildungskonformität bzw. -notwendigkeit* sowie der *Habitus der Bildungsfremdheit* (für eine ausführlichere Darstellung vgl. Kramer et al., 2009, S. 131ff.). Diese Schülerhabitustypen lassen sich entlang einer sehr basalen Dimension der Bildungsaffinität vs. -ferne anordnen, welche anzeigt, in welchem Passungsverhältnis der jeweilige Habitus zu Schule und Bildung steht. Entsprechend der rekonstruktiven Erzeugungslogik der Typen lässt sich dies jedoch im Rahmen eines quantitativen Zugangs nicht im Sinne eines linearen Modells interpretieren, zumal sich in den Konturierungen der einzelnen Habitus verschiedene Dimensionen wiederfinden, die man für die Identifikation von Indikatoren im Rahmen des hier gewählten Zugangs nutzen kann.

Zusammenfassend unterscheiden sich die von Kramer et al. (2009) rekonstruierten Typen von Schülerhabitus systematisch in Hinblick auf drei wesentliche Kernbereiche. Ein besonders relevantes Merkmal ist *erstens* die zugrundeliegende *Bildungsorientierung*, was in Teilen auch der zentralen Dimension (Bildungsaffinität vs. -ferne) des Gesamtmodells entspricht. Alle vier Habitus-typen weisen in Hinblick auf Aspekte wie Bildungsaspirationen, Leistungs- bzw. Anstrengungsbereitschaft, die Wahrnehmung der Schullandschaft oder Hochkulturorientierung ausgeprägte Unterschiede auf. *Zweitens* lassen sich hinsichtlich des *Selbstbildes* der Schülerinnen und Schüler (Selbstkonzept, Hilflosigkeit) Unterschiede in den verschiedenen Typen identifizieren. *Drittens* spielt der hier-von noch zu unterscheidende Aspekt der *Sicherheit bzw. Unsicherheit*, die die Kinder innerhalb der schulischen Umgebung empfinden (Wohlbefinden) eine typ-konstituierende Rolle. Die Identifikation dieser drei Kernbereiche stellt bereits einen ersten Schritt in Richtung Operationalisierung dar. Neben diesen zentralen Merkmalen sprechen Kramer et al. (2009) gelegentlich auch Aspekte wie Fleiß und Anstrengungsbereitschaft, Selbst- vs. Fremdbestimmung hinsichtlich schulischer Leistungserwartungen oder den Stellenwert von Mitschülerinnen und Mitschülern an, allerdings werden nicht immer alle vier Habitusformen auf diesen Dimensionen verortet, so dass sich von einer geringeren Relevanz für die Typologie ausgehen lässt. Die auf den drei Kernbereichen basierenden Schülerhabitustypen lassen sich bereits bei Kindern im Grundschulalter beobachten, wodurch auch die mit dem Übergang in die weiterführenden Schulen der Sekundarstufe I verbundenen Erfahrungen unterschiedlich ausfallen: inwieweit sich im Anschluss an die Wahl einer Schule ein „Gefühl der rechtmäßigen Platzierung“ einstellt oder stattdessen

der Eindruck entsteht, „nicht am rechten Platz, sondern ‚in der Fremde‘ zu sein“ (Kramer et al., 2009, S. 161), wird dieser kulturtheoretischen Perspektive zufolge maßgeblich vom Habitusotyp und der daraus resultierenden Passungskonstellation zur jeweiligen Schulkultur mitbestimmt. Zugleich lässt sich auch die Entwicklung der nachfolgenden Schullaufbahn als Produkt der Passung zwischen Schülerhabitus und Schulkultur begreifen. Damit stellt die von Kramer und Kollegen entworfene Konzeption des Schülerhabitus gerade für die Untersuchung der Entstehung und Reproduktion von Bildungsungleichheiten einen vielversprechenden Ansatz dar. Für quantitative Analysen ergeben sich daraus jedoch insbesondere in Hinblick auf die Operationalisierung einige Fragen, die es zunächst zu klären gilt.

### 3. Forschungsstand und Fragestellung

Einige wenige Arbeiten unternehmen den Versuch, das Bourdieu'sche Habituskonzept im Rahmen quantitativer Analysen umzusetzen. Bereits sehr früh wies McClelland darauf hin, dass sich Aspirationen als *internalisierte objektive Erfolgswahrscheinlichkeiten* (McClelland, 1990, S. 103) als Bestandteile des Habitus interpretieren lassen. Entsprechend nutzten alle darauf aufbauenden Beiträge, die den Versuch der Operationalisierung eines bildungsbezogenen Habitus unternahmen, zumindest auch bildungsbezogene Aspirationen als Kernbestandteil von Habitus und als einen zentralen Indikator für den Aspekt der *Bildungsorientierung* (vgl. etwa Gaddis, 2013; Grgic & Bayer, 2015). De Moll (2018) weist hingegen die Nutzung von Aspirationen als Element einer latenten Habituskonstruktion zurück und nutzt für seine Operationalisierung im Rahmen eines faktorenanalytischen Modells die Aspekte Besorgtheit, Selbstvertrauen, Selbstbeschränkung und Ohnmachtserleben (vgl. de Moll, 2018, 371ff.). Bodovski (2015) argumentiert demgegenüber durchaus nachvollziehbar, dass die Bildungserwartungen (die bildungsbezogenen Aspirationen) einen zentralen Aspekt des sich herausbildenden Habitus von Heranwachsenden darstellen. Sie nutzt darüber hinaus das Fähigkeitselfstkonzept als einen weiteren Indikator für die Bestimmung des Habitus, was sich bei Grgic und Bayer (2015) ebenfalls findet und was bei de Moll (2018) über den Faktor Selbstvertrauen zumindest vergleichbar angelegt ist. Neben der Bildungsorientierung stellen die hier verwendeten Indikatoren eine zweite Habituskomponente dar, die insgesamt vor allem Aspekte des *Selbstbildes* der Personen abdeckt.

Neben Aspirationen und dem Selbstkonzept nutzt Bodovski den Locus of Control als weiteren Indikator zur Habitusbestimmung, während Grgic und Bayer (2015) stattdessen das Selbstwertgefühl in ihre latente Modellierung einfügen. Edgerton, Roberts und Peter (2013) konzentrieren sich in ihrer Arbeit sogar auf nur einen additiven Index aus Bildungserwartungen, Einstellungen zu Lehrkräften und Einstellungen zu post-sekundärer Bildung als Indikator für Habitus. Insofern sind es vor allem die beiden Komponenten *Bildungsorientierung* sowie das

*Selbstbild* der Person, welches in je spezifischen Operationalisierungen Eingang in die Habituskonzeptionen findet. Bei den von de Moll verwendeten Indikatoren (vor allem Ohnmachtserleben bzw. Besorgtheit) wird eine weitere Habituskomponente sichtbar, die man als *Unsicherheitskomponente* bezeichnen könnte und die im Hinblick auf die hier anstehende Konzeptionalisierung von Schülerhabitus das Element der Passungsfähigkeit des Bildungshabitus im Feld der Schule ausdrückt.

Insgesamt zeigt sich in allen bisherigen Arbeiten, dass die Operationalisierung von Bildungshabitus mit relativ wenigen Indikatoren durchgeführt wird, was hinsichtlich der Reduktion auf eine einzelne Dimension zwar zweckdienlich erscheinen mag, dem Habituskonzept in seiner inhaltlichen Breite jedoch kaum gerecht wird. Festhalten lässt sich allerdings, dass Bildungserwartungen bzw. Aspirationen in fast allen Modellen eine zentrale Rolle spielen, wobei sich jedoch deutliche Unterschiede bei der methodischen Umsetzung der Operationalisierung zeigen.

Während einige Autoren Habitus entweder als Beschreibungszusammenhang spezifischer Variablen nutzen (so etwa Bodovski, 2015; Gaddis, 2013) oder als Begriff für einen Index (Edgerton et al., 2013), modellieren Grgic und Bayer (2015) und de Moll (2018) den Habitus als latentes Konstrukt, was auch stärker der hier gewählten Vorgehensweise entspricht. Fast alle Autoren operationalisieren Habitus jedoch als eine Art eindimensionale Skala, auf der Personen dann eine stärkere oder weniger starke Ausprägung aufweisen können, was zwar letzten Endes auch im Rahmen der Typologie von Kramer et al. (2009) entlang der Dimension Bildungsaffinität bzw. -ferne angedeutet wird, in Hinblick auf Bourdieus ursprüngliche Konzeption jedoch durchaus kritisch gesehen werden kann (vgl. hierzu Bremer & Lange-Vester, 2014). Hierin sehen wir die deutlichste Herausforderung im Hinblick auf eine mit quantitativen Daten arbeitende Umsetzung des Habituskonzepts, gleichzeitig aber auch den zentralen Mehrwert. Sowohl die Arbeiten von Bourdieu selbst wie auch die darauf aufsetzende qualitative Bildungsforschung betonen die kategoriale Struktur von Habitus. Entsprechend sollte auch im Bereich quantitativer Analysen verstärkter auf typologisierende Methoden wie etwa Clusteranalysen oder Latent-Class-Analysen zurückgegriffen werden. Die folgenden Analysen stellen hierbei einen ersten Versuch und Vorschlag für eine entsprechende Vorgehensweise dar. Das übergeordnete Ziel dieser Untersuchung ist es dabei, die von Kramer und Kollegen (2009) mit qualitativen Längsschnittdaten entwickelte Schülerhabitustypologie anhand eines quantitativen Zugangs nachzuvollziehen und so eine Verbindung zwischen den typischerweise getrennten Forschungsbereichen herzustellen.

## 4. Daten und Methode

### 4.1 Datenbasis

Unsere Untersuchung basiert auf Daten der Startkohorte 3 (SC3) des Nationalen Bildungspanels (NEPS; vgl. Blossfeld, Roßbach, & Maurice, 2011).<sup>1</sup> Die Grundgesamtheit besteht aus Schülerinnen und Schülern an deutschen Regelschulen, welche im Schuljahr 2010/2011 die fünfte Klassenstufe besuchten. Die Stichprobenziehung erfolgte anhand eines geschichteten Zufallsverfahrens. Die Datenerhebung erfolgte bei den ausgewählten Kindern einmal jährlich im Klassenkontext mithilfe von paper-and-pencil-interviews (PAPI) sowie mit Leistungstests in jährlich wechselnden Fachbereichen. Ergänzend wurden auch bei Schulleitungen und Lehrkräften der Zielpersonen Informationen per PAPI sowie bei den Eltern per Telefonbefragung erhoben.

Unsere Untersuchung stützt sich ausschließlich auf Angaben aus den Schülerfragebögen und den Elterninterviews. Weil sie sich in einer grundlegend andersartigen Situation innerhalb des Bildungssystems befinden, wurden Fälle, die sich zum ersten Messzeitpunkt nicht in einer Regelschule im Sekundarschulbereich befanden (z.B., weil sie die Klassenstufe 5 an einer Grundschule in Berlin oder an einer Förderschule besuchten), in unseren Analysen nicht berücksichtigt. Ferner haben wir uns *a priori* auf die Verwendung von Instrumenten der ersten beiden Messzeitpunkte beschränkt, um dabei so weit wie möglich auszuschließen, dass Prozesse der „Habituustransformation“ – d.h. Entwicklungs- und Veränderungsprozesse im Verlauf der Sekundarstufe (etwa der sich verändernde Stellenwert von Peers; vgl. hierzu Krüger & Deppe, 2014) – zu Unschärfen hinsichtlich der beobachteten Dispositionen des Schülerhabitus führen. Da die 5. und 6. Klasse einen Zeitraum darstellt, in der Eltern und Lehrkräfte die „richtige“ Platzierung der Kinder sicherstellen, richten wir mit der Beschränkung auf dieses Zeitfenster den Blick gezielt auf eine für die Schülerinnen und Schüler wichtige Phase, in der sie sich an der neuen Schule bewähren und ihre Passung zur Lernumgebung unter Beweis stellen müssen. Insgesamt gingen Angaben von  $n = 4990$  Schülerinnen und Schülern in unsere Berechnungen ein. Mit einem Anteil von 48.1 % Mädchen sowie 51.9 % Jungen ist das Geschlechterverhältnis nahezu ausgeglichen.

---

<sup>1</sup> Diese Arbeit nutzt Daten des Nationalen Bildungspanels (NEPS): Startkohorte Klasse 5, doi:10.5157/NEPS:SC3:8.0.1. Die Daten des NEPS wurden von 2008 bis 2013 als Teil des Rahmenprogramms zur Förderung der empirischen Bildungsforschung erhoben, welches vom Bundesministerium für Bildung und Forschung (BMBF) finanziert wurde. Seit 2014 wird NEPS vom Leibniz-Institut für Bildungsverläufe e.V. (LIfBi) an der Otto-Friedrich-Universität Bamberg in Kooperation mit einem deutschlandweiten Netzwerk weitergeführt.

## 4.2 Analytischer Ansatz

Zur Beantwortung unserer Fragestellung stützen wir uns auf Latent-Class-Analysen (LCA) um herauszufinden, inwieweit sich die Schülerinnen und Schüler aus unserer Stichprobe anhand ausgewählter Merkmale in Gruppen („Klassen“) einteilen lassen, die den oben skizzierten Idealtypen nahekommen. Unsere Vorgehensweise bestand dabei aus mehreren Schritten, wobei bei jedem Schritt ggf. unpassende Modelle identifiziert und von der weiteren Betrachtung ausgeschlossen wurden. Ausgangspunkt war die Schätzung einer Reihe von Modellen mit bis zu acht latenten Klassen in *Mplus* (Version 6.12), die wir zunächst formal mithilfe verschiedener Güteindizes vergleichend untersuchten. Anschließend erfolgte die inhaltliche Prüfung, bei der wir die latenten Klassen auf Basis der zur Typisierung herangezogenen Merkmale hinsichtlich inhaltlicher Interpretierbarkeit und Trennschärfe inspizierten. Den Abschluss unserer Analysen bildete die Begutachtung der Klassen anhand einiger ausgewählter externer Validierungskriterien. Die beobachteten Unterschiede wurden bei kategorialen Merkmalen mit  $\chi^2$ -Tests auf Signifikanz untersucht. Bei metrischen Merkmalen wurden Mittelwertunterschiede in Post-Hoc-Analysen anhand von Scheffé-Tests geprüft.

Die Berechnung der LCA in *Mplus* erfolgte als Maximum-Likelihood-Parameterschätzung mit robusten Standardfehlern. Fehlende Werte wurden mithilfe des Full-Information-Maximum-Likelihood-Verfahrens geschätzt.

## 4.3 Instrumente

Mit Rückbezug auf die im theoretischen Teil in Auseinandersetzung mit dem Modell von Kramer und Kollegen (2009) als relevant identifizierten Komponenten von Schülerhabitus (Bildungsorientierung, Selbstbild, Sicherheit/Unsicherheit) sowie unter Berücksichtigung der im Forschungsstand diskutierten Operationalisierungen verwenden wir in unseren Analysen die folgende Zusammenstellung von Indikatoren.

Die *Bildungsorientierung* der Schülerinnen und Schüler wird anhand von drei Teilverfacetten abgebildet: Als Indikator für die Bildungsambitionen berücksichtigen wir die *idealistische Bildungsaspiration* („*Egal, welche Schule du gerade besuchst und wie gut deine Noten sind: Welchen Schulabschluss wünschst du dir?*“; Antwortkategorien: *Schule ohne Abschluss verlassen / Hauptschulabschluss / Realschulabschluss / Abitur*; Quelle: Stocké, 2014; Angabe aus Welle 1). Die Wahrnehmung der Schullandschaft operationalisieren wir anhand einer Skala zur *subjektiven Informiertheit* (Skala aus zwei Items; Beispiel-Item: „*Wie gut weißt du darüber Bescheid, welche Schulabschlüsse man in Deutschland machen kann?*“; fünfstufige Antwortskala von „*sehr schlecht*“ bis „*sehr gut*“; Quelle: NEPS-Eigenentwicklung; Angaben aus Welle 2; Cronbachs Alpha  $\alpha = .75$ ). Der über den schulischen Rahmen hinausgehende Kulturbezug im Habitus wird über eine Skala zur *Leselust* erfasst (Skala aus drei Items; Beispiel-Item: „*Es macht mir Spaß,*

*Bücher zu lesen*“; vierstufige Antwortskala von „stimme gar nicht zu“ bis „stimme völlig zu“; Quelle: Möller & Bonerad, 2007; Angaben aus Welle 1;  $\alpha = .90$ ).

Das *Selbstbild* der Schülerinnen und Schüler erfassen wir anhand einer Skala zum allgemeinen *schulischen Selbstkonzept* (Skala aus drei Items; Beispiel-Item: „Ich bin in den meisten Schulfächern gut“; vierstufige Antwortskala von „trifft gar nicht zu“ bis „trifft völlig zu“; Quelle: Kunter et al., 2002; Angaben aus Welle 1;  $\alpha = .82$ ).

Die Komponente der empfundenen *Sicherheit* bzw. *Unsicherheit* berücksichtigen wir anhand eines Items zum *Wohlfühlen in der Klasse* („In der neuen Klasse fühle ich mich im Vergleich zur Grundschulklasse wohler“; fünfstufige Antwortskala von „trifft gar nicht zu“ bis „trifft völlig zu“; Quelle: NEPS-Eigenentwicklung, angelehnt an Wild, Rammert, & Siegmund, 2006; Angabe aus Welle 1) sowie über ein Item zur *Zufriedenheit mit Schule* („Wie zufrieden bist du mit deiner schulischen Situation?“; elfstufige Antwortskala von „ganz und gar unzufrieden“ bis „ganz und gar zufrieden“; Quelle: NEPS-Eigenentwicklung; Angabe aus Welle 1). Aufgrund der extrem schiefen Verteilung der Antworten wurde die Angabe zur Zufriedenheit mit Schule über den Mittelwert dichotomisiert und ging ebenso wie die idealistische Bildungsaspiration als kategoriales Merkmal in die Berechnungen ein. Alle anderen Variablen wurden in z-standardisierter Form berücksichtigt.

Die von uns im Anschluss vorgenommene Validierung zielt nicht darauf ab, die Erklärungskraft der gefundenen Typen für bildungs- bzw. leistungsbezogene Aspekte im Rahmen eines linearen Modellansatzes zu demonstrieren; vielmehr wollen wir – ganz im Sinne des Bourdieu'schen Vorgehens – zeigen, ob sich hier spezifische *Korrespondenzen* zwischen den Habitustypen und einerseits herkunftsbezogenen Merkmalen sowie andererseits dem Geschlecht nachweisen lassen, die von Bourdieu als die beiden zentralen habituskonstituierenden Merkmale bezeichnet wurden (Bourdieu, Dölling, & Steinrücke, 1997). Als herkunftsbezogenes Merkmal stützen wir uns auf das *Bildungsniveau der Eltern*, welches einen Indikator für das kulturelle Kapital des Kindes darstellt. Dieses basiert auf der höchsten verfügbaren Angabe (Wellen 1 bis 8) des befragten Elternteils bzw. des Partners zum ISCED (International Standard Classification of Education) und wurde zu einer dreistufigen Information zusammengefasst („niedrig“: max. Realschulabschluss; „mittel“: Hochschulzugangsberechtigung; „hoch“: Hochschulabschluss). Das *Geschlecht* des Kindes markiert ein Merkmal, von dem wir annehmen, dass sich hier bereits vorhandene Bildungserfahrungen (in der Grundschule) in einer spezifischen Weise wiederfinden. So unterscheiden sich die beiden Geschlechter bereits in der Grundschule hinsichtlich der Umsetzung kognitiver Kompetenzen in Schulnoten (Helbig, 2010). Darüber hinaus untersuchen wir die Korrespondenz zwischen den Habitustypen und den *Schulformen*, da dies im Hinblick auf die weitere Schul- bzw. Bildungskarriere und die damit verknüpften Chancen im stratifizierten Bildungssystem von elementarer Bedeutung ist.

## 5. Ergebnisse

### 5.1 Latent-Class-Analysen

Insgesamt wurden Modelle mit bis zu acht Klassen berechnet, die wir eingangs formal auf ihre Eignung hin untersuchten. Tabelle 1 weist für alle Modelle eine Reihe statistischer Gütemaße aus, die Hinweise auf die Modellpassung geben. Es zeigt sich, dass der Wert der Log-Likelihood-Funktion (LL) im ersten Modell am niedrigsten ausfällt und sich in Modellen mit steigenden Klassenzahlen zunehmend verbessert. Dies ist als Hinweis darauf zu verstehen, dass die sechs Modellvariablen tatsächlich systematisch miteinander in Beziehung stehen und sich sinnvoll zu latenten Klassen zusammenfassen lassen. Folglich wurde das 1-Klassen-Modell verworfen und blieb von der weiteren Untersuchung ausgeschlossen. Auch die anderen LL-basierten Maße – d.h. das Akaike Information Criterion (AIC), das Bayesian Information Criterion (BIC) sowie das Sample-size Adjusted BIC (SABIC) – zeigen übereinstimmend an, dass Modelle mit höherer Klassenzahl jeweils besser auf die Daten passen als sparsamere Modelle. Aus Sicht dieser Indizes wäre demnach die 8-Klassen-Lösung zu bevorzugen. Allerdings lässt sich mit Blick auf die prozentuale Verbesserung gegenüber dem Nullmodell (PVo) bzw. gegenüber dem jeweils vorausgehenden Modell (PV1) beobachten, dass der Zugewinn mit steigender Klassenzahl tendenziell abnimmt und zum Schluss vergleichsweise gering ausfällt. Zieht man zusätzlich den adjustierten Vuong-Lo-Mendell-Rubin-Tests (p LMR; vgl. Lo, Mendell, & Rubin, 2001) zu Rate, so ist die von den LL-basierten Kriterien präferierte 8-Klassen-Lösung zurückzuweisen, da diese zu keiner signifikanten Verbesserung gegenüber dem Modell mit einer Klasse weniger führt. Aus diesem Grund wurde das 8-Klassen-Modell ausgeschlossen und bei den anschließenden Analyseschritten nicht weiter untersucht. Die Entropie schließlich stellt ein Globalmaß für die Zuverlässigkeit der Klassifikation dar, wobei Werte nahe 1 eine hohe Sicherheit der Klassifikation anzeigen (vgl. Geiser, 2011, S. 270). Diesem Maß zufolge sind die Modelle mit 4 und 5 Klassen den anderen Lösungen vorzuziehen.

Tabelle 1: Güteindizes der Latent-Class-Analysen zu Schülerhabitus bei unterschiedlicher Klassenanzahl

k	npar	LL	PVo	PV1	AIC	BIC	SABIC	p LMR	Entropie
1	12	-31987.6	-	-	63999.2	64077.4	64039.3	-	-
2	21	-31314.8	2.1 %	2.1 %	62671.5	62808.4	62741.6	0.000	0.721
3	30	-30847.8	3.6 %	1.5 %	61755.5	61951.0	61855.6	0.000	0.786
4	39	-30323.8	5.2 %	1.7 %	60725.5	60979.6	60855.7	0.000	0.846
5	48	-30017.7	6.2 %	1.0 %	60131.4	60444.1	60291.6	0.000	0.870
6	57	-29905.6	6.5 %	0.4 %	59925.2	60296.5	60115.4	0.000	0.794
7	66	-29838.2	6.7 %	0.2 %	59808.4	60238.4	60028.7	0.002	0.801
8	75	-29784.3	6.9 %	0.2 %	59718.5	60207.1	59968.8	0.289	0.762

Anmerkungen. Quelle: NEPS SC3;  $n = 4990$ . k = Anzahl latenter Klassen im Modell; npar = Zahl der Parameter; LL = Wert der Log-Likelihood-Funktion; PVo = prozentuale Verbesserung ggü. dem Nullmodell ( $k = 1$ ); PV1 = prozentuale Verbesserung ggü. dem vorherigen Modell; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; SABIC = Sample-size Adjusted BIC; p LMR = adjustierter Vuong-Lo-Mendell-Rubin-Test.

Ein weiteres Gütekriterium besteht in den mittleren Klassenzuordnungswahrscheinlichkeiten (vgl. Tabellen 2 und 3), die aufzeigen, wie sicher sich Personen anhand ihres Antwortmusters einer bestimmten Klasse zuordnen lassen. Werte auf der Hauptdiagonale (fett hervorgehoben) sollten für jede Klasse möglichst über .80 liegen (vgl. Geiser, 2011, S. 250). Die Modelle mit bis zu 5 Klassen erzielen hier gute bis sehr gute Resultate, während die Modelle mit 6 und 7 Klassen jeweils zwei Klassen beinhalten, die sich offenbar nicht präzise voneinander abgrenzen lassen. Aus diesem Grund werden auch diese beiden Modelle bei den nachfolgenden Untersuchungen ausgeblendet.

Tabelle 2: Mittlere Klassenzuordnungswahrscheinlichkeiten der Modelle mit 2 bis 5 Klassen

k	2-Klassen-Modell		3-Klassen-Modell			4-Klassen-Modell				5-Klassen-Modell				
	1	2	1	2	3	1	2	3	4	1	2	3	4	5
1	<b>.88</b>	.12	<b>.91</b>	.00	.08	<b>.94</b>	.02	.00	.04	<b>.95</b>	.01	.01	.01	.02
2	.07	<b>.93</b>	.02	<b>.89</b>	.09	.02	<b>.97</b>	.01	.01	.00	<b>.99</b>	.00	.00	.00
3			.05	.06	<b>.89</b>	.02	.01	<b>.90</b>	.06	.05	.02	<b>.88</b>	.01	.04
4						.03	.01	.06	<b>.90</b>	.02	.02	.01	<b>.94</b>	.02
5										.02	.02	.00	.00	<b>.96</b>

Anmerkungen. **Fett** hervorgehoben: Werte auf der Hauptdiagonale.

Tabelle 3: Mittlere Klassenzuordnungswahrscheinlichkeiten der Modelle mit 6 und 7 Klassen

k	6-Klassen-Modell						7-Klassen-Modell						
	1	2	3	4	5	6	1	2	3	4	5	6	7
1	<b>.94</b>	.01	.01	.03	.01	.01	<b>.92</b>	.01	.01	.00	.02	.01	.03
2	.01	<b>.95</b>	.01	.01	.01	.00	.00	<b>.94</b>	.00	.04	.00	.01	.00
3	.00	.00	<b>.99</b>	.00	.00	.00	.05	.02	<b>.74</b>	.00	.13	.02	.04
4	.02	.00	.01	<b>.96</b>	.00	.00	.01	.13	.00	<b>.81</b>	.02	.01	.01
5	.03	.02	.02	.02	<b>.74</b>	.18	.03	.01	.19	.01	<b>.73</b>	.02	.02
6	.06	.02	.02	.04	.12	<b>.75</b>	.00	.00	.00	.00	.00	<b>.99</b>	.00
7							.02	.00	.00	.00	.00	.02	<b>.96</b>

Anmerkungen. **Fett** hervorgehoben: Werte auf der Hauptdiagonale.

Zusammenfassend lässt sich an dieser Stelle festhalten, dass sich unsere Modellvariablen durchaus sinnvoll zur Bildung latenter Klassen eignen. Anhand der herangezogenen formalen Gütekriterien konnten einige Modelle identifiziert werden, die eine geringe Passung zu den Daten aufweisen und daher aussortiert wurden. Unter den verbliebenen Modellen scheinen die 4- bzw. 5-Klassen-Lösung tendenziell etwas besser auf die Daten zu passen als die sparsameren Modelle, weshalb bei der inhaltlichen Prüfung der Fokus auf diese beiden Modelle gelegt wurde.

## 5.2 Inhaltliche Prüfung

Bei der inhaltlichen Prüfung geht es in erster Linie darum herauszufinden, inwieviel die latenten Klassen im Inneren homogene und nach außen hin heterogene Merkmalsausprägungen aufweisen. Tabelle 4 beinhaltet für die beiden verbliebenen Modelle die Häufigkeitsverteilungen bzw. Mittelwerte der berücksichtigten Variablen.

Das *4-Klassen-Modell* besteht aus einer großen Klasse mit 46 % der Fälle (Klasse 3), einer mittleren Klasse mit 33 % (Klasse 4) sowie zwei kleineren Klassen mit 13 % (Klasse 1) bzw. 8 % (Klasse 2) der Fälle. Hinsichtlich der Bildungsaspiration sticht Klasse 2 durch einen hohen Anteil an Schülerinnen bzw. Schülern heraus, die sich einen Hauptschulabschluss wünschen (14 %), während der Wunsch nach dem Abitur in dieser Klasse besonders schwach ausgeprägt ist (41 %). Den Gegenhorizont hierzu bildet Klasse 3, in welcher der Hauptschulabschluss mit lediglich 3 % kaum Erwähnung findet, wohingegen sich über 80 % das Abitur wünschen. Bei der Informiertheit zeigen sich nur geringe Unterschiede, die lediglich im Vergleich der Klassen 2 und 3 sowie 3 und 4 signifikant ausfallen. Die Leselust weist hochsignifikante Differenzen zwischen allen Klassen auf, und auch das Selbstkonzept unterscheidet sich mit Ausnahme der

Klassen 1 und 2 signifikant voneinander. Beim Wohlfühlen sind dagegen lediglich kleine Unterschiede zwischen den Klassen 1 und 3 sowie 3 und 4 zu beobachten. Die Zufriedenheit schließlich fällt für die Klassen 1 und 4 ähnlich hoch aus, unterscheidet sich aber ansonsten signifikant voneinander.

**Tabelle 4:** Häufigkeitsverteilungen, Mittelwerte und Klassengrößen der Modelle mit 4 und 5 Klassen

Merkmal	4-Klassen-Modell					5-Klassen-Modell				
	1	2	3	4		1	2	3	4	5
n	645	405	2309	1631		1047	410	2349	419	765
Relative Klassengröße	12.9 %	8.1 %	46.3 %	32.7 %		21.0 %	8.2 %	47.1 %	8.4 %	15.3 %
<b>Idealistische Aspiration</b>										
kein Abschluss	0.5 %	0.5 %	0.3 %	0.3 %		0.2 %	0.7 %	0.3 %	0.5 %	0.4 %
Hauptschulabschluss	7.0 %	13.6 %	3.2 %	3.9 %		3.5 %	6.5 %	3.2 %	13.8 %	5.6 %
Realschulabschluss	33.1 %	44.5 %	16.1 %	24.5 %		24.7 %	33.8 %	16.1 %	44.9 %	25.7 %
Abitur	59.4 %	41.3 %	80.5 %	71.2 %		71.6 %	58.9 %	80.5 %	40.7 %	68.3 %
Subjektive Informiertheit	-0.04	-0.14	0.08	-0.06		-0.03	-0.03	0.08	-0.14	-0.06
Leselust	-1.08	-2.14	0.91	-0.14		0.05	-1.26	0.93	-2.16	-0.54
Schulisches Selbstkonzept	-0.32	-0.42	0.27	-0.11		-0.06	-0.33	0.27	-0.42	-0.19
Wohlfühlen in der Klasse	-0.08	-0.06	0.08	-0.05		-0.05	-0.11	0.08	-0.06	-0.03
<b>Zufriedenheit mit Schule</b>										
gering	52.6 %	57.9 %	36.2 %	48.4 %		48.3 %	53.3 %	36.4 %	57.9 %	47.6 %
hoch	47.4 %	42.1 %	63.8 %	51.6 %		51.7 %	46.7 %	63.6 %	42.1 %	52.4 %

Im *5-Klassen-Modell* findet sich mit 47 % der Fälle (Klasse 3) erneut eine große Klasse, gefolgt von zwei mittelgroßen Klassen mit 21 % (Klasse 1) bzw. 15 % (Klasse 5) der Fälle. Die beiden kleinen Klassen 2 und 4 umfassen jeweils rund 8 % der Fälle. Die Bildungsaspiration ist bei Klasse 3 am stärksten in Richtung Abitur ausgeprägt, während Klasse 4 den Gegenpol dazu darstellt. Die subjektive Informiertheit zeigt auch bei dieser Lösung sehr geringe Unterschiede und ist lediglich zwischen Klasse 3 und 4 signifikant. Die Leselust ist dagegen in allen fünf Klassen unterschiedlich ausgeprägt. Das Selbstkonzept fällt bei Klasse 3 am stärksten aus, während die Klassen 2 und 4 hier unterdurchschnittliche Werte aufweisen. Beim Wohlfühlen in der Klasse finden sich kaum Unterschiede. Die Zufriedenheit unterscheidet sich am stärksten zwischen Klasse 3 und Klasse 4.

Im direkten Vergleich der beiden Modelle zeigen sich deutliche Parallelen, wobei bei jeweils mehrere Klassen nahezu identisch verlaufen. Aufgrund der großen inhaltlichen Nähe und der sehr ähnlich ausfallenden formalen Kriterien ist keines der beiden Modelle eindeutig dem anderen überlegen. Daher haben wir uns beim letzten Analyseschritt der Validierung anhand zusätzlicher Kovariaten für das sparsamere 4-Klassen-Modell entschieden. Verglichen mit der Typologie von Kramer et

al. (2009) ergaben sich Parallelen, aber auch Abweichungen: Klasse 3 unterscheidet sich in jeglicher Hinsicht deutlich von den anderen Klassen, wobei die *ausgeprägte Bildungsorientierung* sowie das *positive Selbstbild* und die *hohe Sicherheit* die Vermutung nahelegen, dass es sich dabei um den *Habitus der Strebenden* handeln könnte. Den Gegenhorizont hierfür findet man in Klasse 2, die sich durch eine *sehr geringe ausgeprägte Bildungsorientierung*, ein *negatives Selbstbild* und eine *geringe Sicherheit* auszeichnet und damit am ehesten dem Typen des *Habitus der Bildungsfremdheit* entsprechen. Klasse 1 befindet sich bei praktisch allen Merkmalen zwischen den beiden Extremgruppen (mit leichter Tendenz zu unterdurchschnittlichen Ausprägungen) und könnte demnach als *Habitus der Bildungskonformität und -notwendigkeit* bezeichnet werden. Einzig Klasse 4 lässt sich trotz der vielen zugeordneten Fälle aufgrund der unauffälligen Antwortmuster keinem Habitustyp eindeutig zuordnen. Da viele Variablen einen nahe beim Nullpunkt liegenden Mittelwert aufweisen, lässt sich vermuten, dass diese Klasse möglicherweise eine Art Residualgruppe darstellt, die sich noch aus mehreren Untertypen zusammensetzt. Bezogen auf den *Habitus der Bildungsexzellenz bzw. -distinktion* fanden sich in unseren Ergebnissen keine Hinweise – allerdings ist letztlich fraglich, inwieweit die untersuchte Stichprobe überhaupt Fälle mit einer Habitusform beinhaltet, die nur für Schülerinnen und Schülern rekonstruiert werden konnte, welche anschließend „auf ein exklusives Gymnasium mit zusätzlicher Eingangsprüfung wechselten“ (Kramer et al., 2009, S. 132).

### 5.3 Externe Validierung

Um zu überprüfen, inwieweit die vorgenommene Klassifikation sich zur Erklärung von Bildungsungleichheiten eignet, wurden die Zusammenhänge der latenten Klassen mit drei bildungs- bzw. ungleichheitsrelevanten Merkmalen untersucht. Tabelle 5 weist die Häufigkeitsverteilungen und Mittelwerte der vier Klassen hinsichtlich der Merkmale *Schulform*, *Geschlecht* sowie *Bildungsniveau der Eltern* aus. Bei der in der fünften Jahrgangsstufe besuchten Schulform zeigen sich deutliche Unterschiede zwischen den vier latenten Klassen. Auffällig ist hierbei vor allem der hohe Anteil an Hauptschülerinnen bzw. -schülern in Klasse 2 sowie der hohe Anteil an Gymnasiastinnen bzw. Gymnasiasten in Klasse 3. Auch beim Geschlecht sind Differenzen zwischen den Klassen erkennbar, wobei erneut Klasse 2 einen auffallend hohen Anteil an männlichen Schülern umfasst. Und schließlich zeigt sich auch beim elterlichen Bildungsniveau ein auffälliger Unterschied: Kinder in Klasse 3 haben häufiger hoch gebildete Eltern als Kinder in den anderen Klassen.

Tabelle 5: Häufigkeitsverteilungen der Validierungsmerkmale im finalen 4-Klassen-Modell

Merkmal	4-Klassen-Modell			
	1	2	3	4
<b>Schulform</b>				
Hauptschule	15.8 %	31.6 %	10.0 %	13.3 %
Schule mit mehreren Bildungsgängen	13.3 %	16.8 %	8.8 %	9.8 %
Realschule	27.4 %	26.9 %	19.4 %	25.0 %
Integrierte Gesamtschule	5.4 %	5.2 %	6.0 %	5.5 %
Gymnasium	38.0 %	19.5 %	55.8 %	46.5 %
<b>Geschlecht</b>				
männlich	64.3 %	70.5 %	43.9 %	53.7 %
weiblich	35.7 %	29.5 %	56.1 %	46.4 %
<b>Bildungsniveau der Eltern</b>				
niedrig	5.3 %	7.2 %	3.4 %	4.8 %
mittel	51.4 %	49.4 %	33.7 %	41.4 %
hoch	43.3 %	43.4 %	63.0 %	53.8 %

Die beobachteten Zusammenhänge zu den zusätzlichen Validierungsmerkmalen legen nahe, dass sich die mithilfe unserer Variablenauswahl generierten latenten Klassen durchaus für Anschlussuntersuchungen zu verschiedenen ungleichheitsbezogenen Fragestellungen eignen.

## 6. Diskussion

Die vorliegende Untersuchung stellt einen Versuch dar, zwischen zwei typischerweise getrennten Bereichen eine Brücke zu errichten: Ziel war die quantitative Rekonstruktion des mit qualitativen Daten entwickelten Schülerhabituskonzepts. Auf Basis der Typologie von Kramer et al. (2009) wurden drei grundsätzliche Merkmalsbereiche (Bildungsorientierung, Selbstbild, (Un-)Sicherheit) identifiziert, zu deren Abbildung wir insgesamt sechs in den NEPS-SC3-Daten verfügbare Merkmale (idealistische Bildungsaspiration, subjektive Informiertheit, Leselust, schulisches Selbstkonzept, Wohlfühlen in der Klasse, Zufriedenheit mit Schule) in einer Reihe von Latent-Class-Analysen untersuchten. Unseren Befunden zufolge lassen sich anhand der herangezogenen Merkmale vier distinkte latente Klassen bilden, von denen drei als klar voneinander unterscheidbare Ausprägungen von Schülerhabitus aufgefasst werden können. Besonders deutliche Unterschiede zwischen den Klassen zeigten sich bei der idealistischen Bildungsaspiration sowie bei der Leselust, daneben aber auch beim schulischen Selbstkonzept und bei der generellen Zufriedenheit mit Schule. Über die zur

Typologisierung herangezogenen Merkmale hinaus weisen die identifizierten latenten Klassen Zusammenhänge mit weiteren ungleichheitsrelevanten Merkmalen (Schulform, Geschlecht, Bildungsniveau der Eltern) auf. Inwieweit die gebildeten Typen sich für die Erklärung bildungsbezogener Ungleichheiten eignen, konnte im Rahmen unserer Analysen nur exemplarisch aufgezeigt werden. Die beobachteten Korrespondenzen zwischen den Habitustypen und Merkmalen wie dem *Geschlecht*, dem *Bildungsniveau der Eltern* oder der zu Beginn der Sekundarstufe besuchten *Schulform* legen jedoch nahe, dass dies für weitergehende Untersuchungen ein vielversprechender Weg sein könnte.

Die Befunde deuten für einen am Ansatz Bourdieus orientierten bildungssoziologischen Zugangs jedoch auch auf einige Herausforderungen hin, die sowohl mit dem methodologischen als auch analytischen Vorgehen in Zusammenhang stehen. Während die bei Kramer et al. (2009) beschriebenen Typen im Sinne von Idealtypen auf Basis eines maximal kontrastierenden Designs ermittelt wurden, führen quantitative Vorgehensweisen wie die hier vorgestellte dazu, dass die an Mittelwerten orientierten Typen in sich durchaus heterogen sind. Diese Heterogenität ließe sich durch eine größere Anzahl an Klassen zumindest in Teilen reduzieren und würde es dann auch ermöglichen, die Befunde stärker auch mit den (hier ausgeblendet gebliebenen) Untervarianten der Habitustypen von Kramer et al. (2009) zu vergleichen. Hierfür wäre jedoch zunächst eine Weiterentwicklung der Modelle insbesondere hinsichtlich der hierfür heranzuziehenden Indikatoren notwendig. Die von uns genutzten Indikatoren decken unserer Einschätzung nach die zentralen Komponenten von Schülerhabitus ab, vernachlässigen zugleich aber einige theoretisch möglicherweise ebenfalls relevante Aspekte wie Fleiß und Anstrengungsbereitschaft, Selbst- vs. Fremdbestimmung hinsichtlich schulischer Leistungserwartungen oder der Stellenwert von Mitschülerinnen und Mitschülern. Eine Erweiterung der Typologie um derartige Facetten könnte dabei behilflich sein, die bislang wenig scharf konturierte Klasse 4 präziser in homogene Untergruppen zu zerlegen (wenngleich der Berücksichtigung zusätzlicher Variablen modellanalytisch natürlich auch Grenzen gesetzt sind). Da die on uns genutzten Daten keine systematische Operationalisierung von Habitustypen beinhalten, ist der hier gewählte Zugang zusätzlich mit Herausforderungen wie unterschiedlichen Antwortskalen etc. konfrontiert.

Die hier vorgestellte Analytik sollte zukünftig noch stärker die bereits theoretisch angelegte Mehrebenenstruktur in die Modellierung aufnehmen, womit an die in der Validierung sichtbar gewordenen entsprechenden Varianzen angeknüpft werden kann. So lassen sich über eine systematische Berücksichtigung der Klassenebene und unter Einbezug von Lehrkraftinformationen (wobei die Daten des Nationalen Bildungspanels hier nur eingeschränkt verwendbar sind) Habituserwartungen und damit auch Nähe und Distanz zwischen dem Habitus von Lehrkräften und den Habitus der Schülerinnen und Schüler abbilden. Ähnliches gilt für die Ebene der Schule und ihrer habitusrelevanten schulkulturellen Verortung, worauf die Befunde von Helsper et al. (2001) hindeuten. Dadurch las-

sen sich dann die Habitustypen noch besser im relationalen Feld der Schule(n) verorten.

Da es sich beim Bourdieu'schen Habituskonzept um ein strukturelles Konzept handelt, ist es für zukünftige Forschungen zudem angezeigt, die langfristige Stabilität bzw. den Wandel der Typen über die Zeit hinweg zu untersuchen. Gleichzeitig lassen sich dann auch auf individueller Ebene stattfindende Veränderungsprozesse, etwa in Richtung empirisch sichtbarer Habitustransformationen, innerhalb der Struktur von Habitus und hinsichtlich der damit verbundenen Auswirkungen auf Bildungswege untersuchen. Das konnte mit den Daten des Nationalen Bildungspanels nicht geleistet werden, da es aufgrund fehlender Messwiederholungen nur eingeschränkte Möglichkeiten gibt, die Stabilität der Habitustypen über die Zeit empirisch zu untersuchen.

Insgesamt zeigt die Analyse jedoch, dass die hier vorgestellte kategoriale Umsetzung zu inhaltlich interpretierbaren Befunden führt und gleichzeitig eine Brücke zu den qualitativen Ansätzen bilden kann, deren sehr detaillierte Typbeschreibungen wiederum von einem stärker in die Breite gehenden und die Verteilung der Typen ermittelnden quantitativen Zugang profitieren können.

## Literatur

- Barlösius, E. (2011). *Pierre Bourdieu* (2. Aufl.). Frankfurt a. M.: Campus.
- Becker, R. & Lauterbach, W. (2016). Bildung als Privileg – Ursachen, Mechanismen, Prozesse und Wirkungen. In R. Becker & W. Lauterbach (Hrsg.), *Bildung als Privileg: Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit* (5. Aufl., S. 3–53). Wiesbaden: VS.
- Blossfeld, H.-P., Roßbach, H.-G. & Maurice, J. von (Hrsg.). (2011). *Education as a life-long process. The German National Educational Panel Study (NEPS)* (vol. 14). Wiesbaden: VS.
- Bodovski, K. (2015). From parental to adolescents' habitus: Challenges and insights when quantifying Bourdieu. In M. Murphy & C. Costa (Hrsg.), *Bourdieu, habitus and social research: The art of application* (S. 39–54). New York, NY: Palgrave Macmillan.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in western society*. New York, NY: Wiley-Interscience.
- Bourdieu, P. (1982). *Die feinen Unterschiede: Kritik der gesellschaftlichen Urteilskraft*. Frankfurt a. M.: Suhrkamp.
- Bourdieu, P. (1993). *Sozialer Sinn: Kritik der theoretischen Vernunft*. Frankfurt a. M.: Suhrkamp.
- Bourdieu, P. (2018). *Bildung. Schriften zur Kultursoziologie: Vol. 2*. Berlin: Suhrkamp.
- Bourdieu, P., Dölling, I. & Steinrücke, M. (1997). Eine sanfte Gewalt. In I. Dölling & B. Krais (Hrsg.), *Ein alltägliches Spiel. Geschlechterkonstruktion in der sozialen Praxis* (S. 218–230). Frankfurt a. M.: Suhrkamp.
- Bourdieu, P. & Passeron, J.-C. (1971). *Die Illusion der Chancengleichheit: Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs*. Stuttgart: Klett.
- Bourdieu, P. & Wacquant, L. (1996). *Reflexive Anthropologie*. Frankfurt a. M.: Suhrkamp.

- Bremer, H., & Lange-Vester, A. (2014). Die Pluralität der Habitus- und Milieuformen bei Lernenden und Lehrenden. Theoretische und methodologische Überlegungen zum Verhältnis von Habitus und sozialem Raum. In W. Helsper, R.-T. Kramer & S. Thiersch (Hrsg.), *Schülerhabitus. Theoretische und empirische Analysen zum Bourdieuschen Theorem der kulturellen Passung* (S. 56–81). Wiesbaden: VS.
- de Moll, F. (2018). *Familiale Bildungspraxis und Schülerhabitus: Außerschulische Reproduktionsmechanismen von Bildungsungleichheit in der Grundschulzeit*. Weinheim: Beltz Juventa.
- DiMaggio, P. (1982). Cultural capital and school success: The impact of status culture participation on the grades of U.S. high school students. *American Sociological Review*, 47(2), 189–201.
- Ecarius, J. & Wahl, K. (2009). Bildungsbedeutsamkeit von Familie und Schule. Familienhabitus, Bildungsstandards und soziale Reproduktion – Überlegungen im Anschluss an Pierre Bourdieu. In J. Ecarius, C. Groppe & H. Malmede (Hrsg.), *Familie und öffentliche Erziehung* (S. 13–33). Wiesbaden: VS.
- Edgerton, J. D., Roberts, L. W. & Peter, T. (2013). Disparities in academic achievement: Assessing the role of habitus and practice. *Social Indicators Research*, 114(2), 303–322.
- Gaddis, S. M. (2013). The influence of habitus in the relationship between cultural capital and academic achievement. *Social Science Research*, 42(1), 1–13.
- Geiser, C. (2011). *Datenanalyse mit Mplus: Eine anwendungsorientierte Einführung* (2. Aufl.). Wiesbaden: VS.
- Goldthorpe, J. H. (2007). „Cultural Capital“: Some Critical Observations. *Sociologica*, 2, 1–23.
- Grgic, M. & Bayer, M. (2015). Eltern und Geschwister als Bildungsressourcen? Der Beitrag von familialem Kapital für Bildungsaspirationen, Selbstkonzept und Schulerfolg von Kindern. *Zeitschrift für Familienforschung*, 27(2), 173–192.
- Hagenaars, J. A. & Halman, L. C. (1989). Searching for ideal types: the potentialities of latent class analysis. *European Sociological Review*, 5(1), 8196.
- Helbig, M. (2010). Sind Lehrerinnen für den geringeren Schulerfolg von Jungen verantwortlich? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 93–111.
- Helsper, W., Böhme, J., Kramer, R.-T., & Lingkost, A. (Hrsg.). (2001). *Schulkultur und Schulmythos: Rekonstruktionen zur Schulkultur I*. Opladen: Leske + Budrich.
- Helsper, W., Kramer, R.-T., Thiersch, S. & Ziems, C. (2009). Bildungshabitus und Übergangserfahrungen bei Kindern. In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen. Zeitschrift für Erziehungswissenschaft, Sonderheft 12/2009* (S. 126–152). Wiesbaden: VS.
- Hofstetter, D. (2017). *Die schulische Selektion als soziale Praxis: Aushandlungen von Bildungsentscheidungen beim Übergang von der Primarschule in die Sekundarstufe I*. Weinheim: Beltz Juventa.
- Kramer, R.-T. (2011). *Abschied von Bourdieu? Perspektiven ungleichheitsbezogener Bildungsforschung*. Wiesbaden: VS.
- Kramer, R.-T., & Helsper, W. (2011). Kulturelle Passung und Bildungsungleichheit – Potenziale einer an Bourdieu orientierten Analyse der Bildungsungleichheit. In H.-H. Krüger, U. Rabe-Kleberg, R.-T. Kramer, & J. Budde (Hrsg.), *Bildungsungleichheit revisited: Bildung und soziale Ungleichheit vom Kindergarten bis zur Hochschule* (2. Aufl., S. 103–125). Wiesbaden: VS.
- Kramer, R.-T., Helsper, W., Thiersch, S. & Ziems, C. (2009). *Selektion und Schulkarriere: Kindliche Orientierungsrahmen beim Übergang in die Sekundarstufe I*. Wiesbaden: VS.
- Krüger, H.-H., & Deppe, U. (2014). Habitutransformationen von Schülerinnen im Verlauf der Sekundarstufe I und die Bedeutung der Peers. In W. Helsper, R.-T. Kramer & S. Thiersch (Hrsg.), *Schülerhabitus. Theoretische und empirische*

- Analysen zum Bourdieuschen Theorem der kulturellen Passung* (S. 250–273). Wiesbaden: VS.
- Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., ... Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Lizardo, O. (2008). Comment on John Goldthorpe/5. Three Cheers for Unoriginality. *Sociologica*, 1, 116.
- Lo, Y., Mendell, N. R. & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88(3), 767–778.
- McClelland, K. (1990). Cumulative disadvantage among the highly ambitious. *Sociology of Education*, 64(2), 102–121.
- Möller, J. & Bonerad, E.-M. (2007). Fragebogen zur habituellen Lesemotivation. *Psychologie in Erziehung und Unterricht*, 54, 259–267.
- Niestradt, F. & Ricken, N. (2014). Bildung als Habitus – Überlegungen zum Konzept eines Bildungshabitus. In W. Helsper, R.-T. Kramer & S. Thiersch (Hrsg.), *Schülerhabitus. Theoretische und empirische Analysen zum Bourdieuschen Theorem der kulturellen Passung* (S. 99–124). Wiesbaden: VS.
- Stocké, V. (2014). Idealistische Bildungsaspirationen. *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*.
- van Essen, F. (2013). *Soziale Ungleichheit, Bildung und Habitus: Möglichkeitsräume ehemaliger Förderschüler*. Wiesbaden: VS.
- Wild, E., Rammert, M. & Siegmund, A. (2006). Die Förderung selbstbestimmter Formen der Lernmotivation in Elternhaus und Schule. In M. Prenzel & L. Allolio-Näcke, (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule* (S. 370–397). Münster: Waxmann.

Jennifer Lorenz & Tobias C. Stubbe

## Private tutoring as a means for maintaining social status

### Abstract

*In this study, we analyze to what extent students' social background and, particularly, the motive of status maintenance, as proposed by the rational choice theory, are related to parents' decision to provide private tutoring in secondary school. Given the large share of students attending the highest type of secondary school in Germany (Gymnasium) in recent years, several researchers suggest that this type of school might no longer serve its long-lived purpose of ensuring that families maintain their social status. We propose that parents might employ private tutoring as a new means for status maintenance. Analyzing data from the National Educational Panel Study (NEPS), we find that the decision to take private tutoring in grades 5–8 depends on one aspect of students' social capital, that is, the school-related support that parents offer at home. Additional effects can be identified for the type of school that students attend, a migration history in the family, and students' performance in mathematics and German. Regarding the motive of status maintenance, we find no statistical support for the assumption that parents employ private tutoring to maintain their families' social status. Our analyses indicate, however, that at Gymnasium, parents might use private tutoring to increase their children's social status beyond that of their own.*

### Keywords

*Private tutoring; Social status; Status maintenance; Rational choice theory*

---

Dr. Jennifer Lorenz (corresponding author), Zentrum für Statistik, Georg-August-Universität Göttingen, Humboldtallee 3, 37073 Göttingen, Germany  
e-mail: jennifer.lorenz@uni-goettingen.de

Prof. Dr. Tobias C. Stubbe, Institut für Erziehungswissenschaft, Georg-August-Universität Göttingen, Waldweg 26, 37073 Göttingen, Germany  
e-mail: tstubbe@uni-goettingen.de

## Nachhilfe als Mittel zum sozialen Statuserhalt

### Zusammenfassung

Diese Studie geht der Frage nach, inwiefern die soziale Herkunft von Schülerinnen und Schülern und insbesondere das Motiv des Statuserhalts, wie es von der Rational Choice Theorie vorgeschlagen wird, mit der Entscheidung für Nachhilfeunterricht in der Sekundarstufe I zusammenhängen. Angesichts des Anteils von Schülerinnen und Schülern, die in den vergangenen Jahren das Gymnasium besuchen, gehen verschiedene Forscherinnen und Forscher davon aus, dass diese Schulform ihrem langjährigen Anspruch, Familien den Erhalt ihres sozialen Status zu ermöglichen, nicht mehr gerecht wird. Wir nehmen an, dass Eltern Nachhilfe als neues Mittel einsetzen könnten, um den sozialen Status ihrer Familie zu erhalten. In Analysen mit Daten des Nationalen Bildungspanels (NEPS) zeigt sich, dass die Entscheidung für Nachhilfeunterricht in den Klassen 5–8 mit einem Aspekt des sozialen Kapitals der Schülerinnen und Schüler zusammenhängt, nämlich mit der Unterstützung der Eltern für schulbezogene Aufgaben. Darauf hinaus zeigen sich Effekte der besuchten Schulform, eines Migrationshintergrundes und der Schulleistungen in Mathematik und Deutsch. Im Hinblick auf das Motiv des Statuserhalts finden wir keinen statistischen Beleg für die Annahme, dass Eltern Nachhilfe einsetzen, um ihren Status zu erhalten. Die Analysen liefern jedoch Hinweise darauf, dass Eltern von Gymnasiastinnen und Gymnasiasten Nachhilfe nutzen, um ihren Kindern einen höheren sozialen Status zu verschaffen als ihren eigenen.

### Schlagworte

*Nachhilfeunterricht; Soziale Herkunft; Statuserhalt; Rational Choice Theorie*

## 1. Introduction

Nowadays, half of the German students attend a school at the secondary level that leads to the highest school graduation certificate (*Abitur*) (Autorengruppe Bildungsberichterstattung, 2018). This can be seen as a consequence of the educational expansion, which started in the 1960s and aimed to increase the overall education level to overcome a shortage in the qualified labor force and diminish the relationship between social status and educational attainment (Geißler, 2014). As today's larger share of higher educated students suggests, the goal has been somewhat accomplished, although social disparities in education are still substantially higher in Germany than in many other countries (Weis et al., 2018). Next to the more obvious changes the educational expansion provoked in the schooling system and the labor market, it also had an impact on the social structure. According to the theory of social distinction (Bourdieu, 1986), families of higher social status seek to pass their social status on to their children and thus main-

tain their families' distinctive position in the social structure. Before the educational expansion, this could be accomplished by having children attend the highest secondary school (*Gymnasium*). In the 1970s, for example, this right was reserved to roughly ten percent of all students from a birth cohort (Autorengruppe Bildungsberichterstattung, 2018). Given today's attendance rate, merely sending children to this type of school will probably not suffice to maintain a family's higher social status. Several researchers suggest that families of higher social status might seek out other options to distinguish themselves from families of lower social status (Reimer & Pollack, 2010). As Lucas (2001) puts it for the US American context, "Once that level of schooling [that is the high school level, the authors] becomes nearly universal, the socioeconomically advantaged seek out whatever qualitative differences there are *at that level* and use their advantages to secure quantitatively similar but qualitatively better education" (p. 1652). Several such qualitative measures have been proposed in recent studies, for example, stays abroad (Gerhards & Hans, 2012), class profiles with a special focus on subjects such as Musical Education or Science (Nonte, Haas, Krieg, & Stubbe, 2017), and private tutoring (Schneider, 2005).

In this paper, we build on the idea that parents of higher social status might use private tutoring as a mechanism to ensure their social status is passed to their children. This idea was also brought forth by Schneider (2005), who analyzed data from the German *Socio-economic Panel* (SOEP) and found that families from Western Germany might employ private tutoring strategically to further their children's educational advancement. Owing to the data source, Schneider (2005) did not have a specific measure for status maintenance at hand but rather analyzed it implicitly by looking at students' social backgrounds. Furthermore, his analyses were restricted to a cross-sectional perspective and did not include students' academic performance, which is the driving factor in the decision for private tutoring (Guill & Bonsen, 2011). In another study, Luplow and Schneider (2014) overcame many of these limitations, for example, by including students' academic performance and utilizing the measure of the importance that parents attach to status maintenance. However, their study focused on students in primary schools where private tutoring is less prevalent than at the secondary level (Autorengruppe Bildungsberichterstattung, 2018). Furthermore, they did not model time-varying aspects such as the development of students' grades. With data from the *National Educational Panel Study* (NEPS), it is now possible to examine the motive of status maintenance with a measure explicitly developed within the framework of rational choice theory (Steinberg & Hoenig, 2018). We can also analyze its relationship with private tutoring in secondary school from a panel perspective, including students' performance and changes in their performance.

## 2. Family's social status and private tutoring

### 2.1 Theoretical model of social status in educational research

Students' social background and its impact in the context of education are usually analyzed by referring to Bourdieu's theory of capital (1986). He differentiates three forms of capital that make up a person's social status: economic, cultural, and social capital. A person's professional status and financial means are their economic capital. Cultural capital refers to a person's level of education and possession of culturally valuable objects. It can be divided further into three subcategories: incorporated, objectified, and institutionalized cultural capital. The first is education or knowledge accumulated by a person. The second is capital in its material form, for example, musical instruments or books. The last is defined as education that has been recognized institutionally, for example, certificates. Compared to economic and cultural capital, social capital is rather abstract as it is defined as the resources that others can provide for a person. The amount of social capital a person has, therefore, depends on the size and variety of their social network. In the context of education, Bourdieu's (1986) ideas on social capital are often supplemented by Coleman's (1988) description of the matter, as he refers explicitly to education. According to Coleman (1988), social capital lies, on the one hand, in the social ties of the community, and on the other hand, inside the family. In the context of education, the latter can be interpreted as parents' support for school related-tasks, or more generally, as their involvement in their children's lives, which is an important factor in children's educational success (Coleman, 1988).

### 2.2 Private tutoring in the framework of rational choice theory

Different variations of the rational choice theory (e.g., Breen & Goldthorpe, 1997; Erikson & Jonsson, 1996; Esser, 1999) have proven useful in research on educational decisions, including the decision for private tutoring (e.g., Schneider, 2005). The most commonly used adaptation of the rational choice theory for the German education context was proposed by Esser (1999). Like other rational choice theorists, Esser (1999) assumes that individuals act rationally, that is, they weigh the costs and benefits of alternatives they face. Unlike others, Esser (1999) does not assume that individuals consider the objective conditions of the alternatives. He suggests that costs and benefits are assessed individually, that is, according to a person's perception of the circumstances of their decision. This implies that subjective self-reports, for example, surveys, can be used to assess these circumstances.

According to Esser (1999), the following elements are central to an educational decision, which in this paper is the decision to utilize private tutoring:

- benefit ( $U$ ): enhancing students' school performance to achieve good school graduation certificates;

- cost ( $C$ ): expenditures for private tutoring; and
- the possibility of not maintaining the family's social status ( $-SV$ ): not being able to reach a similar professional status as the parents.

Esser (1999) claims that the direct and indirect costs (i.e., monetary or opportunity costs) for education are the same for everybody in Germany; however, the benefit ( $U$ ) and status maintenance ( $SV$ ) may or may not be realized and thus vary:

- with the probability of success ( $p$ ) of realizing the benefit ( $U$ ) and
- with the probability ( $c$ ) of not maintaining the family's social status ( $-SV$ ).

The following formulas can be employed to calculate the value-expectation weights ( $EU$ ) for a decision with two alternatives. A higher  $EU$  weight will determine the alternative an individual will decide on (Esser, 1999). In this paper, we examine the decision for private tutoring ( $A$ ), which may be realized ( $A_1$ ) or not ( $A_2$ ):

$$EU(A_1) = pU + (1 - p)c(-SV) - C$$

where the benefit ( $U$ ) is realized with a certain probability ( $p$ ) at a certain cost ( $C$ ). Status maintenance might not be attained ( $-SV$ ) by a certain probability ( $c$ ) if the benefit ( $U$ ) is not realized, which might happen with the counter-probability ( $1 - p$ ).

$$EU(A_2) = c(-SV)$$

where, due to not arranging private tutoring, the benefit ( $U$ ) will not be realized by any probability ( $p$ ), and there will be no expenditures ( $C$ ). What remains is the possibility of not maintaining the families' social status ( $-SV$ ), which might occur with a certain probability ( $c$ ).

In the context of the decision for private tutoring, parents will decide to arrange it for their children, when  $EU(A_1) > EU(A_2)$ . By integrating the above formulas accordingly, the following inequality is obtained:

$$U + c(-SV) > \frac{C}{p}$$

In line with Esser's (1999) considerations for educational decisions, we can assume that the costs ( $C$ ) for private tutoring will not differ by parents' social status, as tutoring institutions in Germany have fixed prices. Similarly, the benefit ( $U$ ) of private tutoring, that is, obtaining good grades and good certificates, should not vary by social status. However, the probability of success ( $p$ ) in this endeavor might be positively linked with parents' social status. Parents who have higher education might be able to offer additional help with school-related tasks and thus increase the probability ( $p$ ) of realizing the benefit ( $U$ ). Furthermore, the higher a family's social status already is, the risk of not maintaining a family's social status ( $c(-SV)$ ) is more pronounced. Thus, for families of higher social status, the motivation for arranging private tutoring should be higher.

## 2.3 Factors driving the decision for private tutoring

In the course of their secondary education, one out of five German students receives private tutoring (Autorengruppe Bildungsberichterstattung, 2018). Nevertheless, for decades, this topic has attracted little attention in educational research (Hass & Jäger, 2011). Only in recent years has the number of studies increased (Luplow & Schneider, 2016). These show that the main factor driving the demand for private tutoring is students' performance (Guill & Bonsen, 2011; Ireson & Rushfort, 2014), although the effectiveness of private tutoring to attain this goal has not yet been proven (Guill, Köller, & Lüdtke, 2019). However, other factors also have an impact on parents' decision to arrange private tutoring for their children. One of these is parents' social status, as measured by their income, professional status, and educational attainment.

### 2.3.1 Parents' income

Using data from the German SOEP, Schneider (2005, 2006) found strong positive effects of parents' income on the chances of 17-year olds having received private tutoring at least once during their school career. Supplementing the SOEP data with data from a study on German families (*Familien in Deutschland* (FiD)), Hille, Spieß, and Staneva (2016) also reported that the arrangement of private tutoring is dependent on parents' income. Parents with higher and, in recent years, also middle incomes, are more likely to arrange private tutoring for their children than parents with lower incomes. Another study with data from students in the seventh and eighth grades from schools in Hamburg reported mixed evidence: a higher income is only a significant factor in the decision for private tutoring when students' performance is controlled for. The authors interpret that parents with higher incomes would already react to small decreases in their children's performance by arranging private tutoring (Guill & Bonsen, 2011). For younger students in primary school, Luplow and Schneider (2014) are unable to identify any significant relationship between parents' income and private tutoring.

### 2.3.2 Parents' professional status

When examining the relationship between parents' professional status and the chance that their children receive private tutoring, the evidence is again mixed. While Schneider (2005) cannot find any effects, Guill (2012) shows that in Hamburg, chances for private tutoring are slightly higher for students with parents who have a higher professional status.

### 2.3.3 Parents' educational attainment

Ireson and Rushforth (2014), who analyzed data from students in grades 6, 11, and 13 from South England, found that parents who have attained higher levels of education are more likely to arrange private tutoring for their children. For Germany, Schneider (2006) showed that the highest rates for private tutoring are found in families where at least one of the parents has obtained the highest German school certificate (*Abitur*). Similarly, Hille et al. (2016) observed that students whose mothers had received vocational training compared to those who had not were more likely to receive private tutoring. However, this effect is not present for higher educational certificates such as university degrees. For primary school students, Luplow and Schneider (2014) found evidence for the reversed effect: students whose parents have a higher educational status are less likely to arrange private tutoring for their children. Finally, Guill (2012) found no statistical evidence for the effect of parents' educational level on the decision to arrange private tutoring for children in Hamburg's secondary schools.

### 2.3.4 Parents' support at home

Luplow and Schneider (2014) suggested that the effect they found, that is, that parents of higher social status are less likely to arrange private tutoring for their children in primary school might stem from parents' ability to provide school-related support at home. Indirectly, they found evidence for this hypothesis in the higher shares of private tutoring that they observed when parents do not feel capable of offering support for school-related tasks, which should be especially true for parents with lower levels of education. This hypothesis is also supported by the findings of another study: the more children advance in their school careers, the less their parents support them with school-related tasks (Ireson & Rushforth, 2014; Luplow & Schneider, 2016). One interpretation for this is that parents might not know enough about the school curriculum and thus do not feel capable of helping their children with their homework (Ireson & Rushfort, 2014). Irrespective of the grade or type of school attended, another study revealed that when students feel sufficiently supported by their parents, they are also more likely to receive private tutoring (Birkelbach, Dobischat, & Dobischat, 2017). This might seem contradictory to the above findings, but it could also be interpreted as a general interest of parents in their children's school careers: those that offer support at home are also willing to pay for additional external support.

### 2.3.5 Motive of status maintenance

In their study with primary school students, Luplow and Schneider (2014) found statistical evidence for the role of the motive of status maintenance in the decision

for private tutoring. In the study they analyzed, parents were asked how important the motive of status maintenance was for them. The analyses showed that the importance of status maintenance has a significant positive effect on the decision to take up private tutoring in German. In their final model, this is one of the few effects that remain significant although several indicators for social status (e.g., parents' educational and professional status) were also included.

### **2.3.6 School-related factors**

Guill (2012) found that not only parents' characteristics but also the specific schools students attend have an impact on private tutoring. In the multilevel model that she used for her analyses, the school attended has additional explanatory power in the decision for private tutoring. Adding information describing the schools, Guill (2012) revealed that lower rates of private tutoring are observed when the school attended provides after school hours where students are supported when doing their homework or when it is an all-day school. In a study with the same data preceding the before-mentioned, the effect for all-day schools is not found – this might, however, be due to the application of a different statistical model (Guill & Bonsen, 2011). The school-level effects reported by Guill (2012) might also be a phenomenon specific to the school system in Hamburg, as no similar effect can be found for students from across Germany (Schneider, 2006).

Regarding the difference between types of schools, the evidence is mixed. Guill (2012) reported higher rates of private tutoring at more challenging types of schools in Hamburg. The same holds for Birkelbach et al. (2017), who found that students at *Gymnasium* are more likely to receive private tutoring than those attending the lowest type of secondary school (*Hauptschule*). Hille et al. (2016), conversely, did not find any significant differences between the types of schools in Germany.

### **2.3.7 Migration background and students' sex**

One characteristic that is regularly included in German education studies is a migration background measured by parents' country of birth (Hußmann et al., 2016). As this characteristic is usually strongly correlated with indicators of social background (Stanat, 2003), it is not surprising that mixed findings have been reported for its relationship with private tutoring. For Hamburg, Guill and Bonsen (2011) found statistical evidence for the higher use of private tutoring for families where at least one of the parents was born outside Germany. Contrarily, Luplow and Schneider (2014) could not find any significant relationship with private tutoring in primary school once students' performance is controlled for.

Regarding students' sex, most studies do not report any significant effects, with one exception: Luplow and Schneider (2014). They first found that girls have a sig-

nificantly higher chance of receiving private tutoring in mathematics in primary school. However, the effect becomes insignificant once students' performance is controlled for, indicating that it stems from differences in male and female students' performance.

### 3. Research questions

Current research on the relationship between students' social background and their demand for private tutoring has rendered mixed results. This might stem from the different scopes of the studies, which are specific for a certain region (Guill, 2012; Hamburg; Ireson & Rushfort, 2014; South England) or restricted to younger students (Luplow & Schneider, 2014: primary school). The ambiguous results might also be due to the data used for the analyses, which were not in all studies primarily designed for educational research, and therefore, did not contain central variables such as students' performance (Schneider, 2005, 2006: SOEP data). Furthermore, all the studies cited above adopt a cross-sectional perspective and do not model time-dependent factors in the decision for private tutoring explicitly. These include changes in students' performance, which might be a decisive factor for participation in private tutoring in a specific school year. Against the background of these mixed findings and open questions, in this study, we use NEPS data to determine the relationship between private tutoring and social background with a representative sample for secondary schools across Germany from a panel perspective. Accordingly, our first research question is as follows:

Research question 1: Does social background affect parents' decision to arrange private tutoring for their children in secondary school?

As Esser's (1999) adaptation of the rational choice theory shows, two elements are crucial for the decision to utilize private tutoring and its relationship with social background: the probability of not maintaining the family's social status and the probability of success for a given benefit, which in our case is enhancing students' school performance and attaining good school graduation certificates. From this, we derive the second research question:

Research question 2: Are the motive of status maintenance and the probability of success, as proposed by the rational choice theory, driving factors in parents' decision to arrange private tutoring for their children?

We assume that in the past, status maintenance in the German education system could be achieved by having children attend *Gymnasium*. Nowadays, as almost half of all German students attend this type of school, it may no longer serve that purpose. We propose that the effect of status maintenance is most pronounced at *Gymnasium*, as this is what private tutoring is used as a substitute for. Therefore, our third research question is as follows:

Research question 3: Is the motive of status maintenance more relevant for the decision for private tutoring at *Gymnasium* than at other types of schools?

## 4. Data and methods

For the following analyses, we use data from NEPS Starting Cohort 3 (Blossfeld, Roßbach, & von Maurice, 2011). The original panel started in November 2010 with students in fifth grade (see Abmann et al., 2011). Additional students were sampled in the third wave when students were in seventh grade. In this paper, we analyze students from the original panel cohort to ensure continuity of the data ( $n = 5525$ ). Data use is restricted to the first four waves (students in grades 5–8), as parents were not surveyed in the fifth wave, and thus, central information for our analyses, such as the extent of parental support at home or students' grades, is missing.

### 4.1 Measures

#### 4.1.1 Private tutoring

Students' participation in private tutoring is measured by parents' responses to the question of whether their child currently receives private tutoring, defined as external, regular learning support.<sup>1</sup>

#### 4.1.2 Parents' social status

The theoretical construct of social status is measured based on Bourdieu's theory of capital (see section 2.1). Economic capital is measured by the families' net household income. For reasons of interpretability, it is aggregated into four categories, which are defined by the quartiles of the income distribution in the sample. In addition to income, parents' occupation forms part of a family's economic capital. In the NEPS questionnaire, the interviewed parent was asked to state both parents' professions. We use the higher of the two stated professions, coded according to the *Socio-economic Index of Occupational Status 2008* (ISEI) (Ganzeboom, 2010). The number of books at home and parents' highest educational attainment are used to indicate parents' cultural capital. Similar to parents' occupational status, the higher of the parents' educational certificates is used. It is coded according to the *International Standard Classification of Education* (ISCED) (UNESCO, 1997). Parents' social capital is measured by their support for school-related tasks at home (four items) and the number of different occupations in their social network, as measured by the position generator (Lin & Dumin, 1986). With the first indicator, we refer to social capital within the family. The latter indicator is for social capital that is accessible through social networks outside the family. We use

---

<sup>1</sup> Question in the survey: "Now I would like to move on to the subject of private tuition. Does <target child's name> currently receive private tuition? «Private tuition includes all external educational, systematic, mainly regular support of pupils to overcome any learning issues or to improve their learning performance.»" (FDZ-LIfBi, 2019).

this to extend the view on social capital usually adopted in educational research, which is often limited to social capital within the family (see Lorenz & Stubbe, 2020).

#### **4.1.3 Status maintenance**

The probability of status maintenance is measured by parents' rating for two questions on the chances that their child would be able to attain a similar or better job than their mother's or father's when finishing school with a middle secondary school certificate, that is, the certificate of *Realschule*. Like the measurement of professional status and educational attainment, in the case of differing ratings, the higher of both ratings is used. Consequently, the resulting variable indicates whose job position can more easily be reached or exceeded with the school graduation certificate of *Realschule*; and thus, how far at least the lower of the mother's or father's status can be attained. The idea behind this measure is that the more likely it is for a family to attain status maintenance with the school graduation certificate of *Realschule*, the less afraid they would have to be of losing their social status. This is because, in general, this certificate can be attained more easily than higher ones, and, specifically, more easily when children attend *Gymnasium*. Thus, the higher the probability for status maintenance with the certificate of *Realschule*, the less interested parents should be to enhance their childrens' performance with private tutoring, specifically at *Gymnasium*, where they would also be less interested in their children attaining a better school certificate (*Abitur*).

#### **4.1.4 Probability of success**

The probability of success for a given benefit, which here is enhancing students' performance to attain good school certificates, specifically, a good *Abitur* at the *Gymnasium*, is measured by parents' answers to the question of how likely they think it is that their child will successfully attain the *Abitur*. The lower parents rate this probability, the more motivated they should be to arrange private tutoring.

#### **4.1.5 Type of school**

To determine whether the demand for private tutoring differs by type of school, we include the types of schools students attend. These are secondary school with the highest general education (*Gymnasium*), integrated secondary school (*Integrierte Gesamtschule*), secondary schools with several courses of education (*Schule mit Bildungsgängen*), more extensive general education (*Realschule*), and basic gen-

eral education (*Hauptschule*), and primary school (*Grundschule*).<sup>2</sup> Whether the effect of the motive of status maintenance on the decision for private tutoring is different at *Gymnasium* compared to other types of schools is analyzed using a binary variable: *Gymnasium* vs. other types of schools.

#### 4.1.6 Students' performance

Given that students' performance is a driving factor in the decision for private tutoring (Guill & Bonsen, 2011), students' grades in mathematics and German are included as control variables in the analyses. We use grades because we assume that for parents, they serve as an indicator of children's performance rather than their actual competencies, which would be measured by standardized tests. Additionally, for both subjects, we include the differences in grades compared to the previous school year. The idea behind this is to control for a possible delayed effect of grades from the previous school year, which might cause participation in private tutoring in the ongoing school year.

#### 4.1.7 Control variables

Further control variables for all analyses are as follows: migration background, which is given if at least one parent was born outside Germany; the type of school recommendation students received after primary school, as those that were not recommended for *Gymnasium* might have to utilize private tutoring to ensure they can remain in this type of school; and students' sex, as this might be related to differences in performance and thus differences in the necessity for private tutoring.

### 4.2 Analyses

To model the binary outcome of the decision for private tutoring, we employ logistic regression analysis. The following variables were measured in all four waves and vary over time: private tutoring, parents' support at home, students' performance, and the type of school attended. The elements of the rational choice theory were measured in two waves. Given this data structure, we include a random intercept for the student level to capture unexplained variation between students.<sup>3</sup>

To answer research question 1, we first analyze the impact of all indicators of social background on parents' decision to arrange private tutoring for their chil-

<sup>2</sup> For a comprehensive overview of the German education system, please refer to KMK (2017).

<sup>3</sup> Additionally, we tested whether there is unexplained variation at the school level by including a random intercept for the schools that students attend but found no substantial variation ( $\sigma_{\text{school}} \leq 0.01$ ).

dren (Model 1). In the second model, we include students' performance and changes in their performance alongside other control variables (recommendation for *Gymnasium*, type of school attended<sup>4</sup>, migration background, and sex). For research question 2, we look at a model with the elements of the rational choice theory (Model 3) to determine whether they are statistically relevant for the decision for private tutoring. In the fourth model, we include an interaction effect for attending the *Gymnasium* compared to other types of schools<sup>5</sup> and the motive of status maintenance to analyze whether the interaction proposed in research question 3 can be found. To answer research question 3, we add all social background, performance, and control variables (Model 5). All analyses were performed in R (R Development Core Team, 2019).<sup>6</sup>

### 4.3 Imputation of missing values

For those characteristics that are assumed to be time-invariant and were not available for all waves analyzed here (1–4), we imputed the values deterministically from subsequent and/or preceding waves (see Schulte Nordholt & Hooft Van Huijsdijnen, 1997). This applies to the following variables: parents' income, parents' professional and educational status, the number of professions in parents' social networks (position generator), the number of books at home, students' recommendation for *Gymnasium*, students' migration background, and students' sex. Subsequently, all missing values (including those of variables that were already missing in the wave used for deterministic imputation) were multiply imputed using the R-package *mice* (van Buuren & Groothuis-Oudshoorn, 2011). All analyses were conducted for five multiply imputed datasets, and the results were pooled according to Rubin's rule (Rubin, 1987), where possible.<sup>7</sup>

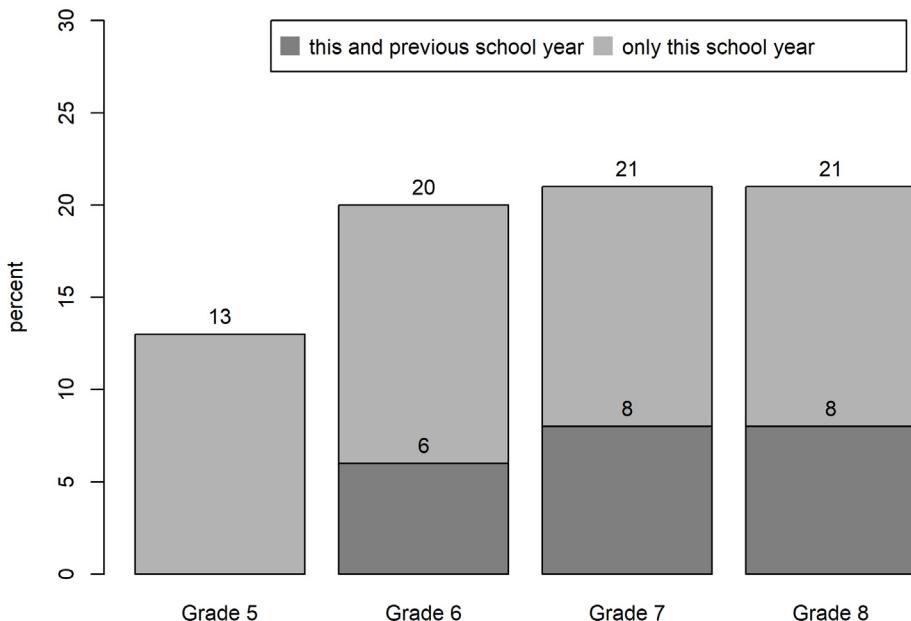
## 5. Results

### 5.1 Descriptive results

The share of students who receive private tutoring indicates that in fifth grade, 13 percent of students receive private tutoring. After the first year of secondary school,

- 
- 4 In Model 1 and 2 we include a polytomous variable for the type of school as we are interested in differences in the demand for private tutoring between all types of schools.
  - 5 In Model 4 and 5 we include a binary variable for the type of school as we are interested in the effect of the motive of status maintenance at *Gymnasium* compared to any other type of school.
  - 6 The following packages were used: *BIFESurvey* (Robitzsch & Oberwimmer, 2019), *dplyr* (Wickham, François, Henry, & Müller, 2018), *lme4* (Bates, Maechler, Bolker, & Walker, 2015), *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017), *merTools* (Knowles, Frederick, & Whitworth, 2019) and *miceadds* (Robitzsch & Grund, 2019).
  - 7 In lack of a pooling formula for pseudo- $R^2$  measures, the arithmetic mean of the five imputed data sets is reported.

Figure 1: Percentage of students receiving private tutoring by school year and previous school year



there is a rise in demand, which stabilizes at approximately 20 percent. Six to eight percent of students who received private tutoring in grades 6–8 had received private tutoring in the previous school year (see Figure 1).

Parents' characteristics that we consider for our analyses are distributed as follows in the NEPS sample. Parents have a mean net income of approximately 3,700 euros per month, which varies by almost the same amount around the mean. Mothers' or fathers' highest ISEI is at 56.7 on a scale from 12–80 points (see Ganzeboom, 2010). The most common educational degree is a vocational training certificate or the highest school graduation certificate at the secondary level, that is, the *Abitur*. Approximately one in four families has at least one parent with a university degree. Most families have less than 500 books at home, and only 16 percent stated that they had more. Almost one-third of students received a recommendation for *Gymnasium* after the fourth grade. Forty-eight percent of all students are female, and 22 percent have at least one parent born outside Germany (see Table 1).

Regarding the time-varying characteristics (see Table 2), overall, not much variation can be observed over the four school years analyzed here. For support at home, the average rating is always above 2.5 (theoretical average) on a scale of 1 “never” to 4 “often,” showing that throughout grades 5–8, parents support their children rather often with school-related tasks. In the first year of secondary school, the mean of students' grades is 2.4 in mathematics and 2.3 in German. Both increase slightly over the following years. This shows that student perfor-

mance decreases as the grading scale in Germany ranges from 1 “very good” to 6 “insufficient”. Parents ratings of the probability of success for the *Abitur* have a mean of 3.7 on a scale of 1 “not probable at all” to 5 “very probable” in the waves in which it was measured (grades 5 and 7). The chance for status maintenance with the school certificate of *Realschule* is on average, slightly lower, with a mean of 3.4 and 3.2 on a scale of 1 “very bad” to 5 “very good” in grades 5 and 7 when they were measured. In grades 5–8, 43.7 to 50 percent of students attend *Gymnasium*, which makes it the type of school with the highest attendance rates. Overall, the majority of students still attend the more “traditional” types of schools (*Gymnasium*, *Realschule*, *Hauptschule*), whereas between 15.3 and 16 percent attend secondary schools with several courses of education (*Schule mit mehreren Bildungsgängen, integrierte Gesamtschule*) in grades 5–8.

Table 1: Time-invariant characteristics of students and parents

	<i>M</i>	<i>SD</i>
Monthly income (in Euro)	3713.9	3533.0
Highest ISEI (12–80 points)	56.7	19.4
Position generator (1–13 professions in social network)	8.3	2.6
Percentage		
<i>Highest educational degree</i>		
None up to middle secondary school level <sup>a</sup>	4.2	
Vocational training or highest secondary school level ( <i>Abitur</i> ) <sup>b</sup>	42.8	
Tertiary education below university level <sup>c</sup>	21.0	
University level <sup>d</sup>	26.4	
PhD or higher <sup>e</sup>	5.5	
<i>Number of books at home</i>		
up to 100	37.4	
101– 500	47.1	
501 or more	15.5	
Recommendation for <i>Gymnasium</i>	29.0	
Migration background	21.6	
Female	48.3	

<sup>a</sup>ISCED level 0, 1, 2. <sup>b</sup>ISCED level 3, 4, 5, 6, 7. <sup>c</sup>ISCED level 5B. <sup>d</sup>ISCED level 5A. <sup>e</sup>ISCED level 6.

Table 2: Time-varying characteristics of students and parents

	Grade 5		Grade 6		Grade 7		Grade 8	
	M	SD	M	SD	M	SD	M	SD
<b>Support at home<sup>a</sup></b>								
Talk to child about literature that was read together	3.2	0.9	3.0	1.0	2.8	1.0	2.7	1.0
Buying additional learning material	2.8	0.9	2.6	0.9	2.7	0.9	2.5	0.9
Searching for information on the internet together with child	2.8	0.9	2.8	0.8	2.8	0.8	2.6	0.8
Support with talks and presentations	3.1	0.9	2.9	0.9	2.8	0.9	2.7	0.9
<b>Students' performance<sup>b</sup></b>								
German grade	2.4	0.8	2.6	0.8	2.6	0.8	2.7	0.8
Mathematics grade	2.3	0.9	2.6	0.9	2.7	0.9	2.7	0.9
Difference in German grade to previous school year			0.2	0.9	0.1	0.7	0.1	0.7
Difference in Mathematics grade to previous school year			0.2	0.8	0.1	0.6	0.0	0.6
<b>Rational Choice Elements</b>								
Probability of status maintenance with <i>Realschule</i> <sup>c</sup>	3.4	1.2	3.4	1.2	3.2	1.2	3.2	1.2
Probability of successful completion of <i>Abitur</i> <sup>d</sup>	3.7	1.2	3.7	1.2	3.7	1.3	3.7	1.3
	Perc.		Perc.		Perc.		Perc.	
<i>Type of school attended</i>								
Secondary school with highest general education ( <i>Gymnasium</i> )	43.7		45.9		50.6		49.8	
Integrated secondary school ( <i>Integrierte Gesamtschule</i> )	5.9		6.2		6.9		7.2	
Secondary school with several courses of education ( <i>Schule mit mehreren Bildungsgängen</i> )	9.5		9.8		8.4		8.8	
Secondary school with more extensive general education ( <i>Realschule</i> )	21.3		21.0		22.7		23.4	
Secondary school with basic general education ( <i>Hauptschule</i> )	13.5		11.5		11.2		10.7	
Primary school ( <i>Grundschule</i> ) <sup>e</sup>	6.1		5.6		0.0		0.0	

<sup>a</sup>Scale from 1 (never) to 4 (often). <sup>b</sup>Scale from 1 (very good) to 6 (insufficient). <sup>c</sup>Scale from 1 (very bad) to 5 (very good). <sup>d</sup>Scale from 1 (not at all probable) to 5 (very probable). <sup>e</sup>In some German *Länder* primary school ends after grade 6 instead of grade 4.

## 5.2 Results from multivariate analyses

In Model 1 (see Table 3), parents' decision to arrange private tutoring for their children is analyzed in relation to their social status. The model shows significant effects for the measures for social capital within the family: parents who buy additional learning materials for their children are more likely to arrange private tutor-

Table 3: Effects on the decision for private tutoring (Model 1–2) [odds ratios]

	Model 1	Model 2
Intercept ( <i>logit</i> )	-2.29	-5.85
<b>Parents' social status</b>		
<i>Monthly income</i> (reference "up to 2500 Euro")		
2501–3400 €	0.89	0.89
3401–4500 €	0.94	0.96
4501 € or more	1.01	1.08
Highest ISEI	1.00	1.00
<i>Highest educational degree</i> (reference "None up to middle sec. school level")		
Vocational training or highest secondary school level ( <i>Abitur</i> )	0.87	0.98
Tertiary education below university level	0.83	0.98
University level	0.73	0.93
PhD or higher	0.77	1.00
<i>Number of books at home</i> (reference "up to 100")		
101–500	0.92	0.98
501 or more	0.83	0.85
Position generator	1.01	1.00
<i>Support at home</i>		
Talk to child about literature that was read together	0.87**	0.92**
Buying additional learning material	1.50***	1.42***
Searching for information on the internet together with child	0.99	1.00
Support with talks and presentations	1.00	0.99
<b>Students' performance</b>		
German grade		1.54***
Mathematics grade		1.79***
Difference in German grade to previous school year		0.90*
Difference in Mathematics grade to previous school year		0.99
<b>Control variables</b>		
Recommendation for Gymnasium		1.11
<i>Type of school attended</i> (reference "Hauptschule")		
<i>Grundschule</i>		1.26
<i>Realschule</i>		1.40*
<i>Integrierte Gesamtschule</i>		1.75***
<i>Schule mit mehreren Bildungsgängen</i>		1.43*
<i>Gymnasium</i>		1.69***
Migration background		1.15
Female		1.00
<b>Random components</b>		
<i>Student level:</i> Residual Variance (Intercept)	1.10	1.33
<i>R</i> <sup>2</sup>	0.16	0.24

\*\*\* $p < .001$ . \*\* $p < .01$ . \* $p < .05$ .

ing for them. For parents who talk with their children about the literature that they read together, the chance of arranging private tutoring decreases.

In Model 2 (see Table 3), which includes additional control variables, the effects for support at home remain significant and do not change substantially. Additional significant effects can be observed for students' grades and the type of school they attend. An increase in students' grades in mathematics and German (which on the German grading scale of 1 "very good" to 6 "insufficient" is to be interpreted as a decrease in performance) increases the chances of receiving private tutoring substantially. The difference in grades compared to the previous school year shows a significant negative effect for the subject German, indicating that a negative difference in grades, which corresponds to an increase in performance, increases the chances for private tutoring. Looking at the type of school that students attend, the model shows that at every type of school, except primary school, the chances of receiving private tutoring are higher than those at the lowest type of secondary school, *Hauptschule*. The sizes of the coefficients suggest that the highest chances can be observed at *Integrierte Gesamtschule* and *Gymnasium*. Adding students' performance and the control variables to the model increases the Pseudo *R*-squared, revealing that these factors have additional explanatory power for the decision for private tutoring.

Regressing the decision for private tutoring on the elements of rational choice theory (Model 3 in Table 4), we find a significant negative effect of the probability of successful completion of the *Abitur*. The lower the probability, the more likely it is that parents arrange private tutoring for their children. The motive of status maintenance is not significantly related to private tutoring in this model. The *R*-squared of the model suggests that the probability of successful completion of the *Abitur* is a relevant factor in the decision for private tutoring.

In Model 4 (see Table 4), we introduce an interaction term for the motive of status maintenance and attending *Gymnasium* instead of any other type of school. The resulting coefficient shows a significant positive relationship. According to this, students at *Gymnasium* have higher chances of receiving private tutoring when the probability of status maintenance with the school leaving certificate of *Realschule* increases.

Adding the variables for social status and the control variables (Model 5 in Table 4), the effect of the probability of successful completion of the *Abitur* is still significant, as is the interaction effect of *Gymnasium* and status maintenance. The effects of social background correspond to those found in Model 2. The same holds true for the effects of grades, except for the difference in grades compared to the previous school year in German (see Table 3). In Model 5 (see Table 4), we find an additional significant effect of migration background: students with a migration background are more likely to receive private tutoring.

Table 4: Effects on the decision for private tutoring (Model 3–5) [odds ratios]

	Model 3	Model 4	Model 5
Intercept ( <i>logit</i> )	-0.78	-0.50	-5.11
<b>Rational Choice Elements</b>			
Probability of status maintenance with school leaving certificate of <i>Realschule</i>	1.00	0.94	0.90
Probability of successful completion of <i>Abitur</i>	0.76***	0.70***	0.90**
School attended: <i>Gymnasium</i>		0.79	1.25
<i>Gymnasium</i> * Prob. of status maintenance (interaction effect)		1.16*	1.16*
<b>Parents' social status</b>			
<i>Monthly income</i> (reference "up to 2500 Euro")			
2501–3400 €			0.91
3401–4500 €			0.98
4501 € or more			1.12
Highest ISEI			1.00
<i>Highest educational deg.</i> (ref. "None to middle sec. school")			
Voc. training or highest secondary school level ( <i>Abitur</i> )			0.93
Tertiary education below university level			0.96
University level			0.93
PhD or higher			0.98
<i>Number of books at home</i> (reference "up to 100")			
101–500			1.00
501 or more			0.88
Position generator			1.01
<i>Support at home</i>			
Talk to child about literature that was read together			0.92**
Buying additional learning material			1.43***
Searching for information on the internet together with child			1.00
Support with talks and presentations			0.99
<b>Students' performance</b>			
German grade			1.48***
Mathematics grade			1.73***
Difference in German grade to previous school year			0.91
Difference in Mathematics grade to previous school year			1.15
<b>Control variables</b>			
Recommendation for Gymnasium			1.10
Migration background			1.16*
Female			1.01
<b>Random components</b>			
<i>Student level:</i> Residual Variance (Intercept)	1.12	1.45	1.34
<i>R</i> <sup>2</sup>	0.16	0.20	0.24

\*\*\* $p < .001$ . \*\* $p < .01$ . \* $p < .05$ .

## 6. Discussion

The scope of this paper is, first, to examine the relationship of social background and the parents' decision to arrange private tutoring for their children in secondary school from a panel perspective, including aspects that vary over time, such as students' performance. The results of the statistical analyses indicate that overall, social background has only minor effects. Controlling for students' performance, the type of school they attend, their sex, and a possible migration history in their families, only parents' social capital, in the form of the support they offer at home, is relevant for the decision for private tutoring. On the one hand, parents who spend time with their children on learning activities are less likely to arrange private tutoring possibly because they can offer the necessary help themselves. On the other hand, parents that often buy additional learning materials for their children have a greater chance of arranging private tutoring, which could be an indication for a general tendency to invest financially in their children's education.

The driving factor in the decision for private tutoring is students' performance, accompanied by substantial effects of the type of school that they attend. The lower the students' grades in mathematics and German, the higher the chances that they receive private tutoring. Additionally, our analyses reveal a delayed effect of grades, indicating that an increase in students' achievement in German in comparison to the previous school year is associated with a higher chance for private tutoring in the given school year. This might be an effect of continued private tutoring from the previous school year, which had already led to an increase in students' performance. Regarding the type of school students attend, the results show the highest chances for private tutoring are at *Gymnasium* and *Integrierte Gesamtschule*, which can be seen as the more demanding types of schools in the German school system.<sup>8</sup>

Overall, our analyses on the effects of social background add to the mixed results found in other studies, revealing only more or less consistent effects of parents' support at home. The effects of parents' financial or professional situations as well as their educational attainment, could not be replicated. The effects of the type of school are in line with other studies that included this characteristic in their analyses, as are the strong effects of students' performance. That we only find minor effects of social background in this study does not necessarily imply that there are no effects. In Boudon's (1974) terminology, we only analyzed secondary effects of social status, which are direct effects of parents' social status on educational decisions. However, based on our analyses, we cannot rule out that there are primary effects included in the effects of students' performance. An indication for this might be seen in the results for parents' support at home, which is known to increase primary effects, as it is an important factor in students' school performance (Tillack & Mösko, 2013).

---

<sup>8</sup> It should be kept in mind that the student body at *Integrierte Gesamtschule* is quite heterogeneous (see footnote 2).

The second and third research questions of this study focus on the relationship between private tutoring and parents' social status against the background of rational choice theory (Esser, 1999). In our analyses, we find a significant negative effect of the probability of successfully attaining the *Abitur* on the chances to receive private tutoring, even when controlling for social background, students' performance, and the type of school that students attend. This indicates that private tutoring is employed, as suggested by the rational choice theory, as a means to increase the probability of realizing the benefit of enhancing students' grades and thus eventually attain better school graduation certificates. Examining the probability of status maintenance with the school certificate of *Realschule*, we find no main effect, but there is an interaction effect with the attendance of *Gymnasium*. This effect is contrary to what we expected: it is not a decreasing probability of status maintenance with the school certificate of *Realschule* but an increasing one that also increases the chances that parents, whose children attend *Gymnasium*, arrange private tutoring for them. According to this finding, our research question on the role of the motive of status maintenance in the decision for private tutoring can be answered as follows. Overall, the motive of status maintenance is not a driving factor in parents' decision to arrange private tutoring for their children, but it is relevant for parents whose children attend *Gymnasium*. However, for them, it is not the risk of not maintaining the family's social status that drives the decision as reasoned by the rational choice theory (Esser 1999), but rather the desire to achieve a higher social status for their children than their own. This motive might have also influenced their decision to send their children to *Gymnasium* in the first place. With our data, however, this has to remain an assumption that should be analyzed in further studies.

All in all, our results indicate that the decision for private tutoring does not depend on parents' economic or cultural capital and is influenced by one aspect of social capital: parents' support at home. As expected, we find that parents employ private tutoring when they perceive a reduced probability that their children will successfully attain the *Abitur*. The motive of status maintenance is only relevant for students at *Gymnasium*. However, it is not the risk of not maintaining the family's social status that affects the decision for private tutoring, but rather parents' desire to raise their children's social status beyond their own. The strong effects of students' performance and type of school in our analyses indicate that first and foremost, investments in the form of private tutoring are made when students' performance indicates the necessity to do so and when students attend a more demanding type of school.

For future research, it is recommended to analyze the decision for private tutoring in more detail. In this study, we used a rather broad definition of private tutoring as "external, regular learning support" (see footnote 1). More substantial effects of social background may be found if private tutoring is viewed as the usage of institutionalized, costly, educational support, or if it is analyzed in a subject-specific manner. Future studies should also closely examine the – from the viewpoint of the rational choice theory unexpected – relationship of attending *Gymnasium* and

using private tutoring as a means to enhance a family's social status. The measure for the motive of status maintenance we employed here might be seen as a limiting factor. However, the effort NEPS put into the development of the empirical representation of the concept, is, to our knowledge, unprecedented and not found in other quantitative studies (Steinberg & Hoenig, 2018). Therefore, it might be useful to employ qualitative methods to gain a more detailed understanding of the underlying mechanisms of status maintenance. Finally, there is, of course, a chance that parents do not employ private tutoring as a means for status maintenance. They might rely on other means such as sending their children to private schools, choosing classes with special profiles or arranging stays abroad, to name a few (e.g., Gerhards & Hans, 2012; Jungbauer-Gans, Lohmann, & Spieß, 2012; Nonte et al., 2017). These means have not received much attention in educational research and deserve further analysis.

## Acknowledgments

This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Grade 5, doi:10.5157/NEPS:SC3:8.0.1. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

## References

- Aßmann, C., Steinhauer, H. W., Kiesl, H., Koch, S., Schönberger, B., Müller-Kuller, A., Rohwer, G., Rässler, S., & Blossfeld, H.-P. (2011). Sampling designs of the National Educational Panel Study: challenges and solutions. In H.-P. Blossfeld, H. G. Roßbach, & J. von Maurice (Eds.), *Education as a lifelong process. The German National Educational Panel Study (NEPS)* (pp. 51–65). Wiesbaden, Germany: VS.
- Autorengruppe Bildungsberichtersstattung. (2018). *Bildung in Deutschland, 2018. Ein indikatoren gestützter Bericht mit einer Analyse zu Wirkungen und Erträgen von Bildung*. Retrieved from <https://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2018/pdf-bildungsbericht-2018/bildungsbericht-2018.pdf>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Birkelbach, K., Dobischat, R., & Dobischat, B. (2017). *Außerschulische Nachhilfe. Ein prosperierender Bildungsmarkt im Spannungsfeld zwischen kommerziellen und öffentlichen Interessen*. Retrieved from: [https://www.boeckler.de/pdf/p\\_study\\_hbs\\_348.pdf](https://www.boeckler.de/pdf/p_study_hbs_348.pdf)
- Blossfeld, H.-P., Roßbach, H.-G. & Maurice, J. von (Eds.). (2011). *Education as a lifelong process. The German National Educational Panel Study (NEPS)*. Wiesbaden, Germany: VS.

- Boudon, R. (1974). *Education, opportunity, and social inequality. Changing prospects in western societies*. New York, NY: Wiley.
- Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York, NY: Greenwood Press.
- Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials Towards a formal rational action theory. *Rationality and Society*, 9, 275–305.
- Coleman, J. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94 (Supplement), S94–S120.
- Erikson, R., & Jonsson, J. O. (1996). Explaining class inequality in education: the Swedish test case. In R. Erikson, & J. O. Jonsson (Eds.), *Can education be equalized? The Swedish case in a comparative perspective* (pp. 1–63). Boulder: Westview Press.
- Esser, H. (1999). Soziologie: *Spezielle Grundlagen. Band 1: Situationslogik und Handeln*. Frankfurt a. M., Germany: Campus.
- FDZ-LIfBi. (2019). *Codebook NEPS starting cohort 3 – grade 5 Paths through lower secondary school Scientific use file version 8.0.1*. Retrieved from: [https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/8-0-1/SC3\\_8-0-1\\_Codebook\\_en.pdf](https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/8-0-1/SC3_8-0-1_Codebook_en.pdf)
- Ganzeboom, H. B. G. (2010). *International standard classification of occupations ISCO-08 with ISEI-08 scores*. Retrieved from [http://www.harryganzeboom.nl/isco08/isco08\\_with\\_isei.pdf](http://www.harryganzeboom.nl/isco08/isco08_with_isei.pdf)
- Geißler, R. (2014). *Bildungsexpansion und Bildungschancen*. Retrieved from <https://www.bpb.de/izpb/198031/bildungsexpansion-und-bildungschancen?p=all>
- Gerhards, J., & Hans, S. (2012). *Transnationales Bildungskapital und soziale Ungleichheit*. BSSE-Arbeitspapier Nr. 25. Berlin, Germany: Freie Universität Berlin.
- Guill, K. (2012). Nachhilfeunterricht: Individuelle, familiäre und schulische Prädiktoren. Münster, Germany: Waxmann.
- Guill, K., & Bonsen, M. (2011). Prädiktoren der Inanspruchnahme von Nachhilfeunterricht am Beginn der Sekundarstufe I, *Empirische Pädagogik*, 25(3), 307–330.
- Guill, K., Köller, O., & Lüdtke, O. (2019). Assessing the instructional quality of private tutoring and its effects on student outcomes: Analyses from the German National Educational Panel Study. *British Journal of Psychology*, 2019, 1–19.
- Hass, C., & Jäger, R. S. (2011). Nachhilfe – Versuch zur Dimensionierung eines Praxis- und Forschungsgebiets. *Empirische Pädagogik*, 25(3), 266–279.
- Hille, A., Spieß, C. K., & Staneva, M. (2016). Immer mehr Schülerinnen und Schüler nehmen Nachhilfe, besonders in Haushalten mit mittleren Einkommen. *DIW Wochenbericht*, 6/2016, 111–121.
- Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E.-M., McElvany, N., Stubbe, T. C., & Valtin, R. (Eds.). (2016). *Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster, Germany: Waxmann.
- Ireson, J., & Rushforth, K. (2014). Why do parents employ private tutors for their children? *Journal for Educational Research Online*, 6(1), 12–33.
- Jungbauer-Gans, M., Lohmann, H., & Spieß, K. C. (2012): Bildungsungleichheiten und Privatschulen in Deutschland. In R. Becker und H. Solga (Eds.), *Soziologische Bildungsforschung*. Kölner Zeitschrift für Soziologie und Sozialpsychologie Sonderheft (pp. 64–85). Wiesbaden, Germany: VS.
- KMK (2017). The education system in the Federal Republic of Germany, 2015/2016. A description of the responsibilities, structures and developments in education policy for the exchange of information in Europe. Retrieved from [https://www.kmk.org/fileadmin/Dateien/pdf/Eurydice/Bildungswesen-engl-pdfs/dossier\\_en\\_ebook.pdf](https://www.kmk.org/fileadmin/Dateien/pdf/Eurydice/Bildungswesen-engl-pdfs/dossier_en_ebook.pdf)

- Knowles, J. E., Frederick, C., & Whitworth, A. (2019). *merTools. Tools for analyzing mixed effect regression models. R package version 3.0.2*. Retrieved from: <https://CRAN.R-project.org/package=merTools>
- Kuznetsova A., Brockhoff P. B., & Christensen R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Lin, N., & Dumin, M. (1986). Access to occupations through social ties. *Social Networks*, 8, 365–385.
- Lorenz, J., & Stubbe, T. C. (2020). *Families' social capital and students' achievement in mathematics and science. Findings from TIMSS 2015, Germany*. Manuscript in preparation.
- Lucas, S. R. (2001). Effectively maintained inequality: education transitions, track mobility, and social background effects. *American Journal of Sociology*, 106(6), 1642–1690.
- Luplow, N., & Schneider, T. (2014). Nutzung und Effektivität privat bezahlter Nachhilfe im Primarbereich. *Zeitschrift für Soziologie*, 43(1), 31–49.
- Luplow, N., & Schneider, T. (2016). Unterstützung durch die Familie. Hausaufgaben und Nachhilfe. In A. Lange, C. Steiner, S. Schutter, & H. Reiter (Eds.), *Handbuch Kindheits- und Jugendsoziologie* (pp. 1–12) Wiesbaden, Germany: VS.
- Nonte, S., Haas, M., Krieg, M., & Stubbe, T. C. (2017, August). *The impact of specific class profiles on within-school segregation in Germany*. Paper presented at the European Conference on Educational Research (ECER), Copenhagen.
- R Development Core Team (2019). *R: A language and environment for statistical computing [Computer software]*. Vienna: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- Reimer D., & Pollack, R. (2010). Educational expansion and its consequences for vertical and horizontal inequalities in access to higher education in West Germany. *European Sociological Review*, 26(4), 415–430.
- Robitzsch, A., & Grund, S. (2019). *miceadds: Some additional multiple imputation functions, especially for 'mice'*. R package version 3.5-14. Retrieved from <https://CRAN.R-project.org/package=miceadds>
- Robitzsch, A., & Oberwimmer, K. (2019). *BIFIEsurvey: Tools for survey statistics in educational assessment* R package version 3.3-12. Retrieved from <https://CRAN.R-project.org/package=BIFIEsurvey>
- Rubin D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.
- Schneider, T. (2005). Nachhilfe als Strategie zur Verwirklichung von Bildungszielen. Eine empirische Untersuchung mit Daten des Sozio-oekonomischen Panels (SOEP). *Zeitschrift für Pädagogik*, 51(3), 363–379.
- Schneider, T. (2006). Die Inanspruchnahme privat bezahlter Nachhilfe. Ein kaum beachtetes Thema in der Bildungsforschung. In C. J. Tully (Ed.), *Lernen in flexibilisierten Welten. Wie sich das Lernen der Jugend verändert* (pp. 131–144). Weinheim, Germany: Juventa.
- Schulte Nordholt, E., & Hooft Van Huijsduijnen, J. (1997). The treatment of item non-response during the editing of survey results. In Eurostat (Ed.), *New techniques and technologies for statistics II: proceedings of the second Bonn Seminar* (pp. 55–62). Amsterdam, Netherlands: IOS Press.
- Stanat, P. (2003). Schulleistungen von Jugendlichen mit Migrationshintergrund: Differenzierung deskriptiver Befunde aus PISA und PISA-E. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann, & M. Weiß (Eds.), *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (pp. 243–260). Opladen, Germany: Leske+Budrich.

- Steinberg, H., & Hoenig, K. (2018). *Measuring rational educational choices in NEPS (NEPS Survey Paper No. 49)*. Retrieved from [https://www.neps-data.de/Portals/0/Survey%20Papers/SP\\_XLIX.pdf](https://www.neps-data.de/Portals/0/Survey%20Papers/SP_XLIX.pdf)
- Tillack, C., & Mösko, E. (2013). Der Einfluss familiärer Prozessmerkmale auf die Entwicklung der Mathematikleistung der Kinder. In F. Lipowsky, G. Faust, & C. Kastens (Eds.) *Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen. Ergebnisse der PERLE-Studie zu den ersten beiden Schuljahren* (pp. 129–149). Münster, Germany: Waxmann.
- UNESCO. (1997). *International Standard Classification of Education ISCED 1997*. Retrieved from [http://www.unesco.org/education/information/nfsunesco/doc/isced\\_1997.htm](http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm)
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67.
- Weis, M., Müller, K., Mang, J., Heine, J.-H., Mahler, N., & Reiss, K. (2018). Soziale Herkunft, Zuwanderungshintergrund und Lesekompetenz. In K. Reiss, M. Weis, E. Klieme, & O. Köller (Eds.). *PISA 2018. Grundbildung im internationalen Vergleich*. Münster, Germany: Waxmann.
- Wickham, H., François, R., Henry, L., & Müller, K. (2018). *dplyr: a grammar of data manipulation. R package version 0.7.6*. Retrieved from <https://CRAN.R-project.org/package=dplyr>

Ruben Kleinkorres, Justine Stang & Nele McElvany

## A longitudinal analysis of reciprocal relations between students' well-being and academic achievement

### Abstract

Besides acquisition of academic competencies, well-being is an important educational goal and it has been shown that both outcomes are mutually dependent. However, until now, most studies used cross-sectional designs so that the direction of the relation is not yet fully understood. In the present study we used longitudinal data from students attending grades 5 to 9 who participated in the National Educational Panel Study (NEPS). We analyzed reciprocal relations between different facets of well-being (i.e., physical, cognitive, emotional) and academic achievement. Furthermore, we examined differences between gender and type of school in these reciprocal relations. Results revealed mainly positive reciprocal relations between academic achievement and different indicators of well-being. However, multi-group models did not show differences between gender and type of school. Implications for future research and educational practice are discussed.

### Keywords

Academic achievement; Cross-lagged-panel analysis; Gender; Students' well-being; Type of school

## Eine Längsschnittanalyse der wechselseitigen Beziehungen zwischen schulischem Wohlbefinden und akademischer Leistung

### Zusammenfassung

Neben dem Kompetenzerwerb stellt Wohlbefinden ein wichtiges Bildungsziel dar und es konnte gezeigt werden, dass beide Variablen voneinander abhän-

---

Ruben Kleinkorres, M.Sc. (corresponding author) · Dr. Justine Stang · Prof. Dr. Nele McElvany, Center for Research on Education and School Development, TU Dortmund University, Vogelpothsweg 78, 44227, Dortmund, Germany  
e-mail: ruben.kleinkorres@tu-dortmund.de  
justine.stang@tu-dortmund.de  
nele.mcelvany@tu-dortmund.de

gen. Bis jetzt liegen allerdings im Wesentlichen Befunde aus Querschnittsstudien vor, sodass die Richtung des Zusammenhangs noch nicht vollständig geklärt ist. In der vorliegenden Studie wurden längsschnittliche Daten von Schülerinnen und Schülern der 5. bis zur 9. Klasse genutzt, die an der nationalen Bildungspanelstudie (NEPS) teilgenommen hatten. Untersucht wurden reziproke Zusammenhänge zwischen verschiedenen Facetten von Wohlbefinden (körperlich, kognitiv, emotional) und schulischer Leistung. Weiterhin wurden in Bezug auf diese reziproken Zusammenhänge Unterschiede zwischen Geschlechtern und Schulformen analysiert. Die Ergebnisse zeigten vor allem positive wechselseitige Beziehungen zwischen schulischer Leistung und verschiedenen Indikatoren des Wohlbefindens. Mehrgruppenmodelle ergaben jedoch keine Hinweise auf Unterschiede zwischen Geschlechtern oder Schulformen. Implikationen für Forschung und Bildungspraxis werden diskutiert.

### **Schlagworte**

*Cross-Lagged-Panel Analyse; Geschlecht; Leistung; Schulform; Wohlbefinden von Schülerinnen und Schülern*

## **1. Introduction**

The Organisation for Economic Co-operation and Development (OECD, 2017) stated that students should feel well at school so that they are motivated to learn and perform well. This demand reflects the importance of students' well-being for student achievement and stresses that there is a relation between well-being and achievement. The importance of students' well-being is also emphasized by being understood as an important educational goal, beside acquisition of academic competencies (van Petegem, Aelterman, Rossel, & Creemers, 2006). Furthermore, these two educational outcomes reflect central indicators of students' positive functioning and are most likely mutually dependent (e.g., Bücker, Nuraydin, Simonsmeier, Schneider, & Luhmann, 2018; Suldo, Riley, & Shaffner, 2006). Seligman, Ernst, Gillham, Reivich, and Linkins (2009) emphasized the importance of school for well-being due to the relation between learning, emotions, and well-being. According to the worldwide prevalence rate for schoolchildrens' depression, Seligman and colleagues (2009, p. 293) even advocate to teach "skills for happiness" at school so that students can handle demanding situations better, feel and learn well.

Although acquisition of academic competencies and well-being are both important educational goals, not much is known on their relation and mutual impact on their development. On grounds of the importance of well-being for learning, more research on student well-being was done in the last years. However, studies were mainly cross-sectional and often focused rather on single aspects of students' well-being instead of considering multiple facets of this complex construct

(Heffner, & Antaramian, 2016). Furthermore, the results of these studies are heterogeneous, whereby a current meta-analysis reported a small to medium-sized positive relation between well-being and academic achievement (Bücker et al., 2018). Nevertheless, the question of the direction between these constructs is still unanswered. Theoretically, the direction can be postulated from academic achievement to well-being as well as vice versa (see e.g., broaden-and-build-theory, Fredrickson, 2001; self-determination-theory, Ryan, & Deci, 2000). A third option could be that there is no causal relation between both constructs, although they are associated with each other. Moreover, the direction might vary for different groups of students: For example, boys and girls differ in their evaluation of school and satisfaction with life (e.g., Hascher, & Hagenauer, 2011; Palsdottir, Asgeirsdottir, & Sigfusdottir, 2012). Furthermore, type of school is an important factor for academic achievement and for well-being (Chang, McBride-Chang, Stewart, & Au, 2003). Differences between these groups of students could have an impact on the relation between well-being and academic achievement.

Altogether, there is still a need for further research in order to shed light on the direction in the relation between academic achievement and well-being and to identify important influencing factors and moderators, especially in a long-term perspective.

## 2. Theoretical background

### 2.1 Students' well-being

Whereas well-being has formerly been defined by means of objective measures like wealth or the fact that a person is married (Wilson, 1967), nowadays research often concentrates on subjective well-being (SWB). SWB is a multi-dimensional and broad construct (Seligman, 2011). However, a clear and widely accepted definition is still missing. Taken as a whole, SWB is referred to as how a person feels and thinks about his or her life in general, as well as concerning a certain domain (Diener, 1984; Diener, Suh, Lucas, & Smith, 1999).

There are several theoretical approaches of SWB. One of the most used approaches is the hedonic one (Eid, & Larsen, 2008; Ryan, & Deci, 2001). The hedonic approach distinguishes cognitive and affective components (Diener, 1984; Diener et al., 1999). The cognitive component comprises cognitive evaluations in the form of global and domain-specific satisfaction (Diener, Inglehart, & Tay, 2013). Global life satisfaction includes all life-evaluative attitudes and beliefs, whereas the domain-specific can be directed and refers to a particular area such as school (Schimmack, 2008). Concerning the affective component Bradburn (1969) already differentiated between positive and negative affect (e.g., joy vs. learned helplessness). Even though it has initially been stated that both components are independent from each other (Bradburn, 1969) the relation between positive and

negative affect has been debated controversially (e.g., Diener et al., 1999; Diener, & Emmons, 1984; Russell, & Carroll, 1999). Furthermore, there are two other important aspects which are of relevance for SWB (WHO, 2014): physical and social well-being. The physical component of SWB comprises measures like absent days as well as subjective evaluations of one's health (Richter, & Hurrelmann, 2009). Social well-being encompasses the evaluation of, for example, social relationships.

For a long time, students' SWB was not in focus of psychological research (Hascher, 2008). However, recently there has been more research on students' SWB. In general, empirical studies in European countries such as Czech Republic, Germany or Switzerland showed that students feel well in school, are satisfied with school and do not indicate physical complaints on a high level (e.g., Hascher, 2007; Urhahne, & Zhu, 2015). However, students also reported school-related worries, especially concerning grading and achievement (Hascher, 2007). Furthermore, it has been shown that SWB decreases with age (e.g., Casas, & González-Carrasco, 2019; Rohlf, 2011), whereby for most countries the decline starts around age of 10 (Casas, & González-Carrasco, 2019). Thus, the examination of the relation between SWB and academic achievement seems to be of special importance in the group of adolescents.

## 2.2 Students' SWB and its relation to academic achievement

Students' SWB and academic achievement are important indicators of their positive psychological functioning (e.g., Suldo et al., 2006). The direction of the relation between SWB and academic achievement can be explained from both sides: Academic achievement could cause SWB as well as SWB could influence academic achievement. The first mechanism is explainable for example by means of the self-determination theory (SDT; Ryan, & Deci, 2000). The SDT assumes that certain psychological needs (autonomy, competence, and relatedness) are of importance for personality growth, social development, intrinsic motivation, and also for SWB. According to this, students' academic achievement as a reflection of the fulfillment of the psychological need of competence may influence students' SWB. The second possible path direction from SWB to academic achievement might be explainable for example by the broaden-and-build theory of positive affect (Fredrickson, 2001). This theory claims that the experience of positive affect as an important component of SWB broadens people's momentary thought-action repertoires. This enables a person to think more flexible and, therefore, might lead to higher achievement. Beneficial for the relation between positive affect and achievement might be that positive affect is also associated with mastery goals (Linnenbrink, & Pintrich, 2002). Additionally, negative affect is negatively related to academic achievement (Gumora, & Arsenio, 2002). Apart from these two theoretical approaches that could explain a causal relation it is also possible that the constructs are associated in a non-causal manner.

While some studies did not find a significant correlation between these two outcomes (Huebner, 1991; Huebner, & Adlerma, 1993), numerous studies have given evidence for an interrelation between SWB and academic achievement (Bird, & Markle, 2012; Bücker et al., 2018; Crede, Wirthwein, McElvany, & Steinmayr, 2015; Pietarinen, Soini, & Pyhältö, 2014; Suldo, Shaffer, & Riley, 2008). A meta-analysis from Bücker et al. (2018) revealed that there is a medium-sized positive relation between the two constructs.<sup>1</sup> Additionally, there is evidence that higher achievement leads to lower levels of psychopathology (Suldo, & Shaffer, 2008). However, most of the studies were cross-sectional. Only few studies analyzed the causal relation between SWB and academic achievement. For example, Steinmayr, Crede, McElvany, and Wirthwein (2016) found that students' great point average in grade 11 predicted changes in life satisfaction in grade 12 positively. Because the evidence from longitudinal data is rare and only certain age groups have been considered, the causal direction of the relation between students' SWB and academic achievement is not yet fully understood.

Beside the interplay between SWB in general and academic achievement, academic achievement is also associated with a positive health behavior and perception of one's own health (Eide, Showalter, & Goldhaber, 2010; Lavy, & Sand, 2012; Sigfusdóttir, Kristjánsson, & Allegrante, 2007; Véronneau, & Dishion, 2012; Zajacova, Lynch, & Espenshade, 2005). Furthermore, studies which mainly focused on adolescents found a positive correlation between life satisfaction and academic achievement (e.g., Heffner, & Antaramian, 2016; Proctor et al., 2010). Concerning school, it was reported that life satisfaction is positively related to school grades (Gilman, & Huebner, 2006; Verkuyten, & Thijs, 2002). Moreover, students' perception of school satisfaction predicted positively students' perception of their academic achievement (e.g., Samdal, Wold, & Bronis, 1999). Concerning negative aspects of students' SWB, studies showed a negative relation to academic achievement. Studies have shown that *learned helplessness*<sup>2</sup> is negatively related to academic achievement: In particular, learned helplessness in the third grade predicted achievement test scores in grade 5 negatively (Fincham, Hokoda, & Sanders, 1989).

### **2.3 Gender and type of school as possible influencing factors on the relation between students' SWB and academic achievement**

One factor which might explain differences in the relation between SWB and academic achievement is gender. There are some studies which found sex differences in certain areas of SWB: In contrast to boys, girls stated more often physical issues and experience of stress; whereas boys reported a less positive attitude to-

<sup>1</sup> SWB was operationalized in the meta-analysis according to Diener's (1984) definition.

<sup>2</sup> The concept of *learned helplessness* (Seligman, 1972) describes the subjective conviction that one has lost the ability to change one's own life situation, because of repeated negative experiences.

wards school (Hascher, & Hagenauer, 2011; Palsdottir et al., 2012). Corresponding to this, it has been found repeatedly that girls in contrast to boys are in general more satisfied with school (Czerwenka et al., 1990; Haecker, & Werres, 1983; Hascher, & Winkler-Ebner, 2010). However, it seems that girls are more prone to learned helplessness compared to boys, especially in the domain of mathematics (Dweck, 1986; Farmer, & Vispoel, 1990). Concerning other important facets of SWB, like social problems or self-esteem, no differences between boys and girls were discovered (Hascher, & Hagenauer, 2011). In the overall picture, some differences become apparent between sexes concerning SWB whereby there are more facets in which girls have lower values than facets in which boys have lower values. So far, however, there is an ambiguous results pattern concerning gender differences in SWB (Gysin, 2017, p. 107). Given that girls tend to have better grades than boys (e.g., Berger, Alcalay, Torretti, & Milicic, 2011), it seems possible that there are gender-specific differences regarding the relation between SWB and academic achievement in that way that in girls these constructs are not as high and positively associated as in boys. A possible explanation could be that girls compensate lower well-being with a higher readiness to perform well in school. In other domains this phenomenon has been referred to by a higher "conformity of girls towards school requirements" (e.g., Sparfeldt, Buch, Schwarz, Jachmann, & Rost, 2009). Up to now, the evidence for a moderating effect of gender is sparse and inconclusive: Herman, Lambert, Reinke, and Ialongo (2008) found a moderating effect of gender in a mediation model from academic competence in the first grade over perceived control in grade six on depression in seventh grade. In contrast, there is also evidence that gender does not influence the path coefficients between academic achievement and SWB. For example, a current meta-analysis found no moderating effect of gender on the relation between SWB and academic achievement (Bücker et al., 2018).

Another factor which might lead to differences in the relation between SWB and academic achievement is type of school. One study found differences in SWB between students attending different types of school (Fend, Knörzer, Nagl, Specht, & Väth-Szusdziara, 1976). Additionally, Chang et al. (2003) assume that type of school is an important factor for well-being on the one hand, but also for achievement on the other hand. A relation between school type and well-being might be explainable by differential learning environments. Students from different school types differ in achievement as well as in other student characteristics like motivation. For instance, motivational characteristics of high school students (Gymnasium) were significantly more distinct than of students of other school types. These differences may be a result of student compositional effects, different curricula or also different teacher preparation programs (e.g., Baumert, Maaz, Stanat, & Watermann, 2009; Diedrich et al. 2019). The evidence for a moderating effect of type of school on the relation between SWB and academic achievement is still scarce. Opdenakker and van Damme (2000) found differential relations between achievement motivation and well-being dependent on school characteristics like an orderly learning environment. Assuming that the learning environments

differ between schools (Baumert et al., 2009) a moderating effect of type of school on the relation between SWB and academic achievement could also be plausible. However, Bücker et al. (2018) did not find an impact on the relation between student SWB and academic achievement through type of school.

### 3. Research questions

Due to the vital importance of SWB for students and its relation to academic achievement, we investigated the question of reciprocal relations between these constructs. Previous studies mainly used cross-sectional data and focused on only few aspects of SWB. Therefore, the direction of the relation between students' SWB and academic achievement is not yet fully understood. Because theoretical arguments support mutual effects in both directions, we pursued an explorative approach in the present study. Furthermore, several facets of SWB were taken into account and the critical age period of secondary school was focused. Additionally, we examined factors that might influence the mutual effects like gender and type of school because research on these moderating variables is sparse. The following research questions were analyzed:

1. Do reciprocal relations exist between academic achievement and a) physical well-being, b) cognitive well-being, and c) emotional well-being?
2. Do the reciprocal patterns differ between a) gender and b) type of school?

## 4. Method

### 4.1 Participants

Longitudinal data analysis with three points of measurement ( $t_1$  = grade 5,  $t_2$  = grade 7,  $t_3$  = grade 9) was based on data of the Starting Cohort 3 of the National Educational Panel Study (NEPS; Blossfeld, Roßbach, & Maurice, 2011). The original NEPS sample consisted of 4,335 students on  $t_1$ , 6,012 students on  $t_2$  and 5,779 students on  $t_3$  who attended one of the traditional academic tracks of the German secondary school system ("Hauptschule", "Realschule", "Gymnasium")<sup>3</sup>. We concentrated on students who attended the same school over the time intervals of interest ( $N_{1-3} = 2,993$ ,  $N_{2-3} = 4,303$ ) and excluded individuals who skipped or repeated classes between the corresponding intervals. The latter step led to an exclusion of 89 of the individuals who participated from  $t_1$  on and 118 of the individuals who participated on  $t_2$  and  $t_3$ . Finally, individuals with untypically high age were excluded from data: This concerned two of the individuals who participated from

---

<sup>3</sup> The sample size was increased in the NEPS between grades 6 and 7. Therefore, the number of students who participated on  $t_2$  and  $t_3$  was higher than those who participated on  $t_1$ .

$t_1$  on and five of the individuals who participated on  $t_2$  and  $t_3$ , respectively.<sup>4</sup> Our final sample comprised 2,902 students (49.7 % female)<sup>5</sup> who participated constantly within the interval from  $t_1$  to  $t_3$  and 4,180 students (49.5 % female) who participated constantly within the interval from  $t_2$  to  $t_3$ . On the first point of measurement, in grade 5, students were on average 10.75 years old ( $SD = 0.48$ ). On the second point of measurement, in grade 7, students were on average 12.77 years old ( $SD = 0.50$ )<sup>6</sup>. Sample characteristics for the different types of school are displayed in Table 1.

Table 1: Sample characteristics for different types of school

	$t_1$			$t_2-t_3$		
	Hauptschule	Realschule	Gymnasium	Hauptschule	Realschule	Gymnasium
$N$	363	789	1750	583	1176	2431
% female	43.0	48.3	51.7	44.1	45.6	52.6
$M_{age}$ ( $SD_{age}$ )	11.10 (0.59)	10.85 (0.47)	10.63 (0.40)	13.12 (0.62)	12.88 (0.50)	12.63 (0.41)

## 4.2 Instruments

### 4.2.1 Academic achievement

For the measurement of reading competence (RC) as well as mathematical competence (MC) we used tests from the NEPS (Blossfeld et al., 2011). The reading comprehension test examined students' RC in grades 5, 7, and 9. The test consisted of a total of 25 tasks which differed in difficulty. The weighted likelihood estimates (WLE) revealed good reliabilities for all points of measurement ( $t_1$ : WLE reliability = .77;  $t_2$ : WLE reliability = .79;  $t_3$ : WLE reliability = .79; Gehrer, Zimmermann, Artelt, & Weinert, 2012). MC was examined in grades 5, 7, and 9 and consisted of 25 tasks, as well. Again, reliabilities were good for all points of measurement ( $t_1$ : WLE reliability = .78;  $t_2$ : WLE reliability = .72;  $t_3$ : WLE reliability = .81; Neumann et al., 2013).

In our analyses, the corrected weighted likelihood estimates (for a detailed description see Scharl, Fischer, Gnambs, & Rohm, 2017) of these competence tests were used as global test scores.

- 
- 4 The age of these individuals differed by more than four standard deviations from sample mean age.
  - 5 In most waves there were missing values concerning sex of the students. We replaced missing values on this dichotomous variable with the median indication of sex over all other waves.
  - 6 Due to different numbers of individuals who participated within the interval from  $t_1-t_3$  and  $t_2-t_3$ , respectively, the mean age differed between these subsamples.

#### 4.2.2 SWB

Questionnaires were administered in the classroom in order to obtain data of students. Because data for academic achievement was available for grades 5, 7, and 9, we focused on constructs of SWB that had been measured longitudinally in at least two of these grades. In order to measure physical well-being we used two indicators: students' self-estimated health and number of days of absence from school. Self-estimated health served as subjective indicator for general physical well-being and was measured by a single item ("How would you describe your health status in general?") on a 5-point scale (1 = *very poor*, 2 = *poor*, 3 = *fair*, 4 = *good*, 5 = *very good*)<sup>7</sup>. Days of absence from school served as objective indicator for school-related physical well-being. Students responded on the item by indicating a single number from 0 to 99.

For the measurement of cognitive well-being we used the *satisfaction scale* that was included in the questionnaire of the NEPS. The scale comprised six items that ask respondents about their satisfaction concerning different life-domains (overall, wealth, health, family, friends, and school). Students responded to each item on an 11-point scale ranging from 0 (= *completely dissatisfied*) to 10 (= *completely satisfied*). We separated the scale into the variable satisfaction with school which was assessed with one item ("How satisfied are you with your school situation?") and the variable satisfaction with life which comprised the remaining five items (e.g., "How satisfied are you, currently, with your life as a whole?") in order to oppose these two domains of SWB to each other. Measurement invariance testing showed that at least configural invariance could be assumed for the variable satisfaction with life (see Appendix D, Tables D3 and D4).

To measure emotional well-being, we used the helplessness scale. The instrument that is used in NEPS is based on a scale by Jerusalem and Schwarzer (1993) who derived items to measure the construct of learned helplessness (Seligman, 1972). The scale originally comprised five items that measure school-related helplessness in general. In the NEPS, the scale was duplicated to measure school-related helplessness for the subjects "German" and "Mathematics" (In the following we will refer to these scales by using the terms "helplessness German" and "helplessness Math"). Participants responded to the items (e.g., "No matter how much I try in German, my grades won't get better") on a 4-point scale ranging from 1 (= *completely disagree*) to 4 (= *completely agree*). Measurement invariance testing revealed that configural invariance was only given if items 4 and 5 of the subscale helplessness German, as well as items 1 and 2 of the subscale helplessness Math were excluded in a two-factor model (see Appendix D, Tables D1 and D2). Therefore, we ran our analyses with two latent factors measured by 3 items each.

---

<sup>7</sup> In order to simplify interpretation of scores and relations with other variables, we reversed the self-estimated health scale. It was originally ranked as follows: 1 = *very good*, 2 = *good*, 3 = *fair*, 4 = *poor*, 5 = *very poor*.

### 4.2.3 Moderating variables: Gender and type of school

For the examination of our second research question, we included gender (female vs. male) and type of school in our analysis. The latter was a three-staged factor differentiated into the traditional academic tracks of the German secondary school system: The lowest track (“Hauptschule”), the intermediate track (“Realschule”), and the highest track (“Gymnasium”).

## 4.3 Data analysis

All analyses were conducted in *R* (Version 3.6.3; R Core Team, 2020). Descriptive statistics were calculated for all variables. Furthermore, differences between sexes, school types and measurement occasions were analyzed by means of ANOVAs and *t*-tests. We computed partial eta-squared and *d*-values for dependent *t*-tests (see Morris, & DeShon, 2002) to quantify the effect sizes of the different test-statistics. Post-hoc tests were corrected with the Bonferroni-Holm method.

Prior to the examination of our research questions, we tested measurement invariance for those variables that were measured by multiple-item scales (see section 4.2). We followed the suggestions of other authors, which described four main steps: configural, metric, scalar, and residual invariance (e.g., Putnick, & Bornstein, 2016; Widaman, & Reise, 1997). Indicator-specific covariances were included within measurement invariance analyses over points of measurement.

In order to answer our first research question three different longitudinal cross-lagged panel models were specified. In each model we investigated the reciprocal relations between academic achievement and one facet of SWB. While the relations from academic achievement with physical and emotional well-being, respectively, were calculated over two points of measurement, in the model investigating the relation between academic achievement and cognitive well-being three points of measurement were considered. Therefore, we were able to include random intercepts in the latter model to separate the within-person development from stable between-person differences (see Hamaker, Kuiper, & Grasman, 2015). To take into account that constructs only measured by a single item (i.e., days of absence, self-estimated health, satisfaction with school) were not measured perfectly reliable, we modeled each of the respective variables as single indicators of a latent variable and specified a value of 0.85 for the reliability (e.g., Jöreskog, & Sörbom, 1982; Petrescu, 2013). The global test scores of MC and RC were handled in the same way, whereby fallibility was taken into account by using the reliabilities of the single measurement occasions (see section 4.1). All latent variables that were included in a single model were allowed to correlate on each of the different occasions. However, in the random intercepts cross-lagged panel model we constrained covariances between the factors capturing the individuals’ stable scores over all waves (random intercepts) and exogenous within-person factors to zero. For the investigation of our second research question, multi-group models with gender and

type of school, respectively, were established to examine the moderating function of these variables. For each of the moderating variables, a model with free estimation of regression parameters for all factor levels was compared to a model that restricted the regression parameters for all factor levels to be equal.

For testing measurement invariance and all analyses, the package “lavaan” (Rosseel, 2012) was used. The parameters of the models were calculated by means of maximum likelihood estimation with robust standard errors and scaled test statistic that is asymptotically equal to the Yuan-Bentler test statistic (Rosseel, 2012). To evaluate model fit, we used robust estimators of the different fit measures (see Brosseau-Liard, Savalei, & Li, 2012). Because the  $\chi^2$ -difference-test is vulnerable in case of large samples, we used  $\Delta\text{CFI}$  for the comparison of nested models instead. We followed the suggestions by Cheung and Rensvold (2002) to prefer the restricted model if  $\Delta\text{CFI}$  is equal to or greater than -.01. The hierarchical structure of the data was taken into account (ICCs are displayed in Tables C1 and C2; see Appendix C). For this purpose the identification number of the students' schools (Blossfeld, & Roßbach, 2019) was considered as cluster variable in our analyses to compute robust standard errors of the parameters estimated in the structural equation models. Finally, missing data was handled within the structural equation models through full information maximum likelihood method.

## 5. Results

### 5.1 Descriptive results

In Table A1 (see Appendix A) means and standard deviations of the variables we used in our analyses are summarized. As can be seen in the Table, means of satisfaction with life decreased over time ( $F(2,4412) = 185.78, p < .001, \eta_p^2 = .08$ ), whereby scores differed significantly between classes 5 and 7 ( $t(2487) = 13.99, p < .001, d = -0.28$ ) and between classes 7 and 9 ( $t(2485) = 7.72, p < .001, d = -0.16$ ). Means of satisfaction with school also decreased over time ( $F(2,4752) = 226.29, p < .001, \eta_p^2 = .09$ ). Scores differed significantly between classes 5 and 7 ( $t(2556) = 17.82, p < .001, d = -0.36$ ) and between classes 7 and 9 ( $t(2556) = 3.57, p < .001, d = -0.06$ ). The same statistics are displayed in Appendix A separately for gender (Table A2) and type of school (Table A3). Satisfaction with life did not differ between gender ( $F(1,2205) = 0.52, p = .47, \eta_p^2 = .00$ ), whereas girls showed on average a higher satisfaction with school ( $F(1,2375) = 14.37, p < .001, \eta_p^2 = .01$ ). A further trend became apparent insofar that boys compared to girls reported a higher helplessness in school subject “German” ( $F(1,3334) = 96.19, p < .001, \eta_p^2 = .03$ ), whereby girls reported a higher helplessness in school subject “Math” than boys ( $F(1,3313) = 73.20, p < .001, \eta_p^2 = .02$ ). Descriptive results for type of school showed that scores in days of absence ( $F(2,2937) = 13.79, p < .001, \eta_p^2 = .01$ ), helplessness German ( $F(2,3333) = 39.31, p < .001, \eta_p^2 = .02$ ),

and helplessness Math ( $F(2,3312) = 30.73, p < .001, \eta_p^2 = .02$ ) differed significantly between students attending different types of school. Table A3 shows that scores on these measures were highest in students attending “Hauptschule” and lowest in students attending “Gymnasium”. Results reversed for the variables satisfaction with school ( $F(2,2374) = 28.11, p < .001, \eta_p^2 = .02$ ) and satisfaction with life ( $F(2,2204) = 15.45, p < .001, \eta_p^2 = .01$ ), meaning that these scores were highest in students attending “Gymnasium” and lowest in students attending “Hauptschule”. In general, it has to be noted that the means of most of the measures were not close to the center of the scale.

Correlations between all variables of interest except for type of school can be seen in Appendix B (Table B1). The table contains mostly statistically significant relations that followed theory-based expectations. It has to be noted, however, that the relations of competence measures with self-estimated health and with satisfaction with life were smaller and reached a less high level of significance compared to relations between other variables.

## 5.2 Reciprocal relations between SWB and academic achievement

### 5.2.1 Reciprocal relations between physical well-being and academic achievement

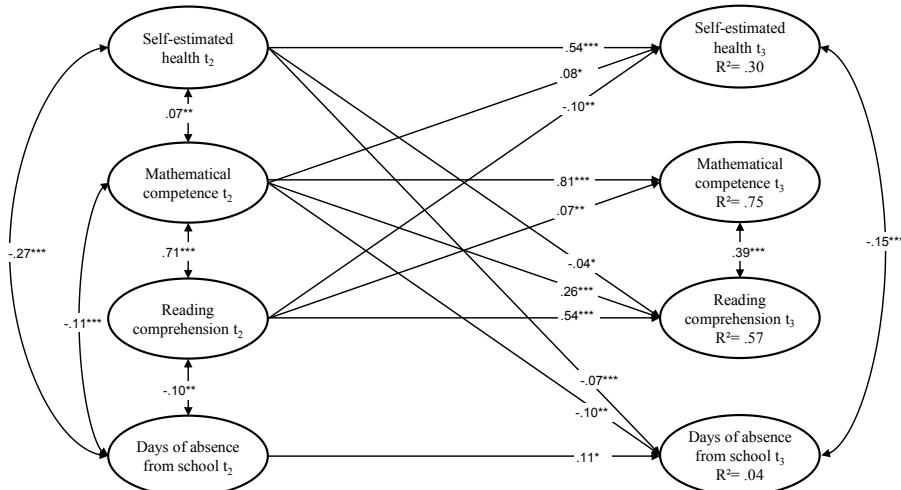
The results for our first research question can be seen in Figure 1. The path from MC on  $t_2$  on days of absence from school on  $t_3$  was significant ( $\beta = -.10, p = .001$ ), whereas the contrary path from days of absence on  $t_2$  on MC on  $t_3$  was not. The restriction of the corresponding cross-lagged regression parameters (from days of absence on mathematical competence and vice versa) to being equal was not detrimental to model fit ( $\Delta\text{CFI} = -.001$ )<sup>8</sup>. Hence, we assumed that the longitudinal association from MC on  $t_2$  on days of absence from school on  $t_3$  was not stronger than vice versa. We did not find reciprocal relations between days of absence from school and RC.

The analysis of reciprocal relations between self-estimated health and competence measures revealed a positive longitudinal association between MC on  $t_2$  and self-estimated health on  $t_3$  ( $\beta = .08, p = .020$ ), but not vice versa. Restricting the corresponding regression parameters to being equal was not detrimental to model fit ( $\Delta\text{CFI} = .000$ ). Conversely, the relation between RC and self-estimated health was reciprocal: The paths from RC on  $t_2$  on self-estimated health on  $t_3$  ( $\beta = -.10, p = .001$ ), as well as vice versa ( $\beta = -.04, p = .032$ ) were both significantly negative. But again, there was no evidence that the regression parameters differed in strength ( $\Delta\text{CFI} = .000$ ).

---

<sup>8</sup> Comparison by means of the  $\chi^2$ -difference test revealed a significant difference ( $\Delta\chi^2(1) = 13.69, p < .001$ ).

**Figure 1:** Reciprocal relations between physical well-being (measured by self-estimated health and days of absence from school) and academic achievement (MC and RC) from grade 7 to grade 9.



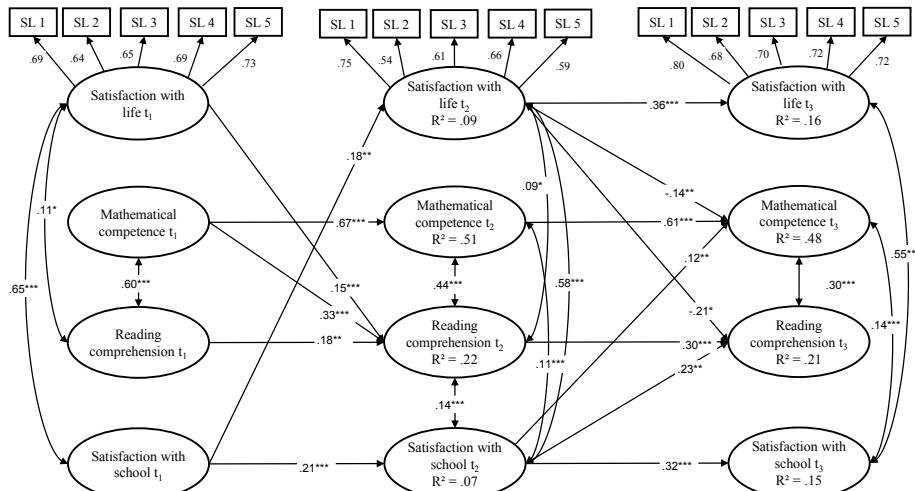
*Notes.* N = 4159. Fit measures are not reported because the model is saturated. All latent variables are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

\*p < .05. \*\*p < .01. \*\*\*p < .001.

### 5.2.2 Reciprocal relations between cognitive well-being and academic achievement

Figure 2 displays reciprocal relations between cognitive well-being and academic achievement: While there was no reciprocal relation between satisfaction with life and MC between t<sub>1</sub> and t<sub>2</sub>, the path from satisfaction with life on t<sub>2</sub> on MC on t<sub>3</sub> got significant ( $\beta = -.14$ ,  $p = .005$ ). However, a comparison with the corresponding path from MC on t<sub>2</sub> on satisfaction with life on t<sub>3</sub> did not reveal a significant difference in strength ( $\Delta\text{CFI} = .000$ ). Higher satisfaction with life on t<sub>1</sub> was associated with higher RC on t<sub>2</sub> ( $\beta = .15$ ,  $p < .001$ ) whereas the corresponding path between t<sub>2</sub> and t<sub>3</sub> was significantly negative ( $\beta = -.21$ ,  $p = .013$ ). In both cases, however, the restriction of paths from satisfaction with life on RC and vice versa to being equal was not detrimental to model fit ( $\Delta\text{CFI}_{1,2} = .000$ ;  $\Delta\text{CFI}_{2,3} = .000$ ). Concerning satisfaction with school we identified a unidirectional association of satisfaction with school on t<sub>2</sub> on MC on t<sub>3</sub> ( $\beta = .12$ ,  $p = .004$ ). Restriction of cross-lagged path from satisfaction with school on t<sub>2</sub> on MC on t<sub>3</sub> and vice versa to being equal was not detrimental to model fit ( $\Delta\text{CFI} = .000$ ). We obtained similar results for the reciprocal relation between satisfaction with school and RC: Higher satisfaction with school on t<sub>2</sub> was associated with higher scores in RC on t<sub>3</sub> ( $\beta = .23$ ,  $p = .002$ ). This association was, however, not stronger than from RC on t<sub>2</sub> on satisfaction with school on t<sub>3</sub> ( $\Delta\text{CFI} = .000$ ).

**Figure 2:** Reciprocal relations between cognitive well-being (measured by satisfaction with life and satisfaction with school) and academic achievement (MC and RC) from grade 5 to grade 9.



Notes.  $N = 2899$ .  $\chi^2(187) = 479.51$ , AIC = 232205.43, BIC = 233023.61, CFI = .989, TLI = 0.979, RMSEA = .028, SRMR = .025. Variables "Mathematical competence", "Reading comprehension" and "Satisfaction with school" are measured by single indicators which have not been included in the figure. Latent variables that model random intercepts have not been included in the figure for clarity purposes. Only statistically significant paths are shown.

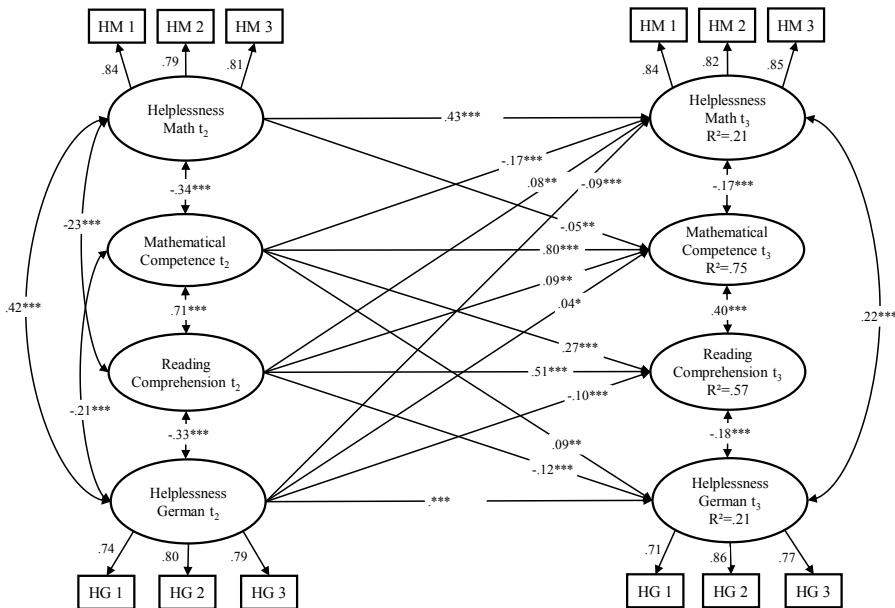
\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

### 5.2.3 Reciprocal relations between emotional well-being and academic achievement

Figure 3 displays the pattern of reciprocal relations between emotional well-being and academic achievement. We observed a negative reciprocal relation between helplessness Math and MC: Higher helplessness Math on t<sub>2</sub> was associated with lower MC on t<sub>3</sub> ( $\beta = -.05$ ,  $p = .002$ ) and vice versa ( $\beta = -.17$ ,  $p < .001$ ). However, the longitudinal relation between helplessness Math and RC was unidirectional, whereby higher scores in RC on t<sub>2</sub> were associated with higher helplessness Math on t<sub>3</sub> ( $\beta = .08$ ,  $p = .008$ ). Helplessness German showed reciprocal relations with both competence measures: Higher helplessness German on t<sub>2</sub> was associated with higher scores in MC on t<sub>3</sub> ( $\beta = .04$ ,  $p = .030$ ) and the path from MC on t<sub>2</sub> on helplessness German on t<sub>3</sub> was significantly positive as well ( $\beta = .09$ ,  $p = .006$ ). Moreover, helplessness German on t<sub>2</sub> was negatively associated with RC on t<sub>3</sub> ( $\beta = -.10$ ,  $p < .001$ ) and vice versa ( $\beta = -.12$ ,  $p < .001$ ). When the opposite regression parameters from a competence measure on a measure of emotional well-being and vice versa were compared to each other we did not find differences in the strength of paths for any of the reciprocal relations (all  $\Delta\text{CFIs} = .000$ )<sup>9</sup>.

9 The  $\chi^2$ -difference test revealed that the associations from RC on t<sub>2</sub> on helplessness German on t<sub>3</sub> and vice versa differed significantly from each other ( $\Delta\chi^2(1) = 7.14$ ,  $p < .01$ ).

**Figure 3:** Reciprocal relations between emotional well-being (measured by helplessness German and helplessness Math, respectively) and academic achievement (MC and RC) from grade 7 to grade 9.



Notes. N = 4159,  $\chi^2(80) = 354.51$ , AIC = 130731.60, BIC = 131187.58, CFI = .989, TLI = 0.983, RMSEA = .031, SRMR = .022. Variables "Mathematical competence" and "Reading comprehension" are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

\*p < .05. \*\*p < .01. \*\*\*p < .001.

### 5.3 Effects of moderating variables on the reciprocal relations between SWB and academic achievement

Model comparisons examining potential moderating effects of gender and type of school on the reciprocal relations between well-being (physical, cognitive, emotional) and academic achievement are displayed in Table 2. As can be seen,  $\Delta\text{CFI}$  values suggested that the restriction of regression parameters to being equal between sexes and types of school, respectively, was not detrimental to model fit in any of the models. This implies that the moderating variables did not have an effect on the reciprocal relations between SWB and academic achievement. Nevertheless, in the following some descriptive differences in the patterns of reciprocal relations between academic achievement and different facets of SWB are described.

### **5.3.1 Moderating effects on the reciprocal relations between physical well-being and academic achievement**

Reciprocal relations between physical well-being and academic achievement are displayed in Figures E1 and E2 (see Appendix E) for boys and girls, respectively. As can be seen, significant associations between MC on  $t_2$  and self-estimated health ( $\beta = .12, p = .024$ ) and days of absence ( $\beta = -.16, p < .001$ ), respectively, on  $t_3$  could only be observed for female students. However, the restriction of the corresponding paths to being equal for both sexes, did reveal a significant difference neither for the association between MC and self-estimated health ( $\Delta\text{CFI} = -.001$ )<sup>10</sup> nor for the association between MC and days of absence ( $\Delta\text{CFI} = .000$ ).

Patterns of reciprocal relations between SWB and academic achievement for different types of schools can be obtained from figures F1, F2, and F3 (see Appendix F). They show only small differences in the pattern of reciprocal relations: The most prominent differences were that MC on  $t_2$  was not associated with days of absence on  $t_3$  for students attending "Realschule" and that days of absence on  $t_2$  were negatively associated with MC on  $t_3$  only in students attending "Gymnasium". The restriction of the corresponding paths to being equal for all types of school did not affect model fit (all  $\Delta\text{CFIs} = -.001$ )<sup>11</sup>.

### **5.3.2 Moderating effects on the reciprocal relations between cognitive well-being and academic achievement**

Concerning gender, the comparison of the patterns of reciprocal relations (see Figures G1 and G2; Appendix G) did not reveal any striking differences between female and male students.

With a view on figures H1, H2, and H3 (see Appendix H) it becomes apparent that there were also no noteworthy differences in the pattern of reciprocal relations between cognitive well-being and academic achievement dependent on different types of schools.

### **5.3.3 Moderating effects on the reciprocal relations between emotional well-being and academic achievement**

The patterns of reciprocal relations between emotional well-being and academic achievement separately for male and female students are displayed in Figures I1 and I2 (see Appendix I). We did not obtain significant paths from helplessness

<sup>10</sup> The  $\chi^2$ -difference test revealed a significant difference ( $\Delta\chi^2(1) = 14.67, p < .001$ ).

<sup>11</sup> The  $\chi^2$ -difference test revealed a significant difference between types of school concerning the association between days of absence on  $t_2$  and MC on  $t_3$  ( $\Delta\chi^2(2) = 21.03, p < .001$ ).

**Table 2:** Model comparisons examining the potential moderating effects of gender and type of school on the reciprocal relations between well-being (physical, cognitive, emotional) and academic achievement

Facet	Moderator	Model	$\chi^2(\text{df})$	$\Delta\chi^2(\text{df})$	<i>p</i>	AIC	BIC	CFI	$\Delta\text{CFI}$
Physical well-being	Gender	Unrestricted	0 (0)			95083.31	95640.61	1.000	
		Restricted	62.68 (16)	62.68 (16)	< .001	95117.16	95573.14	.994	-.006
	Type of school	Unrestricted	0 (0)			93228.11	94064.06	1.000	
		Restricted	57.21 (32)	57.21 (32)	< .01	93235.81	93869.11	.994	-.006
Cognitive well-being	Gender	Unrestricted	706.75 (374)			231535.25	233171.61	.984	
		Restricted	767.15 (406)	60.40 (32)	< .01	231540.39	232985.65	.983	-.001
	Type of school	Unrestricted	913.82 (561)			229162.78	231617.33	.981	
		Restricted	1014.70 (625)	100.16 (64)	< .01	229139.58	231211.91	.980	-.001
Emotional well-being	Gender	Unrestricted	451.97 (160)			130008.33	130920.29	.988	
		Restricted	474.64 (176)	23.85 (16)	.093	130004.92	130815.54	.988	.000
	Type of school	Unrestricted	492.83 (240)			128021.95	129389.89	.989	
		Restricted	552.10 (272)	59.81 (32)	< .01	128032.00	129197.27	.987	-.002

Math on  $t_2$  on MC on  $t_3$  and from RC on  $t_2$  on helplessness German on  $t_3$  for female students. However, the restriction of the corresponding paths to being equal for both sexes was not detrimental to model fit (all  $\Delta\text{CFIs} = .000$ ).

Figures J1, J2, and J3 (see Appendix J) show the reciprocal relations between emotional well-being and academic achievement separately for different types of school. It can be seen that the pattern of reciprocal associations of students attending “Hauptschule” did show nearly no significant paths. The restriction of single paths to being equal for all types of school, however, was not detrimental to model fit in any of the cases (all  $\Delta\text{CFIs} = .000$ ).

## 6. Discussion

Due to the importance of students’ SWB and academic achievement for educational success, this study investigated reciprocal relations between several facets of SWB and academic achievement in adolescents. Additionally, it was analyzed whether these reciprocal relations differ for gender and for type of school.

### 6.1 Reciprocal relations between SWB and academic achievement

In accordance with previous research (Bird, & Markle, 2012; Bücker et al., 2018; Crede et al., 2015; Pietarinen et al., 2014; Suldo et al., 2008), we found mainly positive relations between indicators of students’ SWB and academic achievement. Regarding physical well-being, results revealed that MC played a crucial role for SWB even though the effect sizes were relatively small: MC in grade 7 was significantly associated with both indicators of physical well-being in grade 9, whereas we did not find associations between measures of physical well-being in grade 7 and MC in grade 9. A possible explanation could be that higher MC constitutes the fulfillment of the psychological need of competence (see Ryan, & Deci, 2000) which in turn has a positive impact on the physical well-being of students. However, we also found that RC and self-estimated health had a negative longitudinal relation that was of reciprocal nature. This is a surprising finding, even though these relations were very small and we did not find significant bivariate correlations between both measures. The examination of reciprocal relations between cognitive well-being and academic achievement revealed that satisfaction with life in grade 5 was positively associated with changes in RC in grade 7 while paths from satisfaction with life in grade 7 on MC and RC in grade 9 indicated a negative longitudinal relation. Whereas the former result coincides with prior research that reports positive longitudinal associations between satisfaction with life and school grades (e.g., Ng, Huebner, & Hills, 2015), the latter result seems to be in conflict to these earlier findings. Additionally, higher satisfaction with school in grade 7 was associated

with higher scores in MC and RC in grade 9. Finally, we found mainly negative reciprocal relations between the negative facet of emotional well-being (i.e., helplessness) and academic achievement. Cross-lagged paths between helplessness Math and MC, as well as between helplessness German and RC were negative which fits to theory as well as prior research (e.g., Fincham et al., 1989). Furthermore, we observed a positive reciprocal relation between helplessness German and MC and a positive association between RC in grade 7 and helplessness Math in grade 9. This result can be interpreted analogously to earlier findings that found negative associations from verbal and math achievement on self-concept in the respective other domain (Möller, Pohlmann, Köller, & Marsh, 2009). The underlying internal/external frame of reference model which states that students evaluate their achievement in a given subject not only in relation to the achievement of others but in relation to their own achievement in other subjects, as well (e.g., Möller et al., 2009), could also be an explanation for the positive effects we found between academic achievement in one subject and helplessness in another subject.

Overall, we found evidence that there are reciprocal relations between SWB and academic achievement, whereby it was dependent on the considered facet of well-being if more paths from measures of competence on measures of SWB were statistically significant or vice versa. It has to be noted, however, that we did not find differences in the strength of reciprocal paths. With regard to the direction of the relation, the results do not provide an explicit answer. Therefore, our results neither support only the SDT (Ryan, & Deci, 2000) nor only the broaden-and-build theory (Fredrickson, 2001).

## **6.2 Effects of moderating variables on the reciprocal relations between SWB and academic achievement**

In general, there was no evidence for moderating effects of the variables gender and type of school. As mentioned earlier, findings in relation to moderating effects of these variables on the relation between SWB and academic achievement are sparse and heterogenous: Single studies suggest moderating effects of gender (Herman et al., 2008) and type of school (Opdenakker, & van Damme, 2000). In contrast, our results support earlier findings that did report neither moderating effects of gender nor of type of school on the relation between SWB and academic achievement (cf., Bücker et al., 2018). However, it has to be considered that we used  $\Delta\text{CFI}$  in our analyses to examine moderating effects of the variables gender and type of school. This measure is relatively conservative in comparison with  $\Delta\chi^2$ .

## **6.3 Strengths and Limitations**

A clear strength of the present study is the approach to analyze reciprocal relations between academic achievement and SWB. For this purpose, in our analyses,

we used longitudinal data with two or even three measurement time points. In the research field of students' well-being, the usage of longitudinal data is an important extension to the mainly cross-sectional studies which have been conducted in the past. Another quality of our study is the large sample size which promises a high external validity of our results. A further strength of the present investigation is that the multidimensionality of SWB was met by including several facets of the construct and taking into account at least one important factor for each of these. The consideration of more than one or two SWB constructs helped to get a fine grained picture on the different relations between SWB and academic achievement. Additionally, the study offers important insights for two different domains of academic achievement which are both of high importance because mathematics and reading are two main key competencies. Last but not least, we analyzed differential effects for gender and school types.

Nevertheless, it is important to be aware of some limitations when interpreting the results of the present study. First, even though we used longitudinal data to investigate the reciprocal relations between academic achievement and well-being, we could only use data of two different points of measurement in two of our models. In these models we could, therefore, not distinguish stable variability between persons from within person variance (Hamaker et al., 2015) which may lead to spurious results. Furthermore, the minimum interval between two points of measurement was two years in each model. Because of this relatively large time lag, short term effects might not be detected and longitudinal associations should be interpreted with caution. Second, even though we included several aspects of SWB, the operationalization of SWB was not completely satisfying: One of the reasons was that some constructs had been surveyed rarely and not parallel with other important constructs in the NEPS. For example, indicators for the domain-specific social component of SWB (e.g., social inclusion in class) had only been measured in grades 6–8, whereas indicators for academic achievement like RC and MC had been measured in grades 5, 7, and 9. Therefore, a longitudinal analysis including this facet was not possible. Another problem concerning operationalization of SWB was that some constructs were measured with single items. Even though we corrected for reliability by modeling them as single indicators of latent variables, single items are prone to contain random and nonrandom errors (Bollen, 1989, p. 151). Moreover, constructs measured by multiple indicators did partly show low levels of measurement invariance. Especially, the satisfaction scale did rarely reach levels higher than metric invariance. Concerning the results of the present study it should be considered that latent mean differences can only be interpreted reliably when high levels of measurement invariance are given (e.g., Putnick, & Bornstein, 2016). Furthermore, we considered the nesting of students in different schools but we were neither able to account for variability between German states nor to consider the nesting of students in classes. This may have affected the estimation of standard errors. Another factor that could have had an impact on the results is that the sample sizes of the different types of school differed a lot.

Finally, we did not consider possible controls and third variables. On the one hand we did not include controls for pre-existing differences in achievement like socioeconomic status and ethnic heritage (e.g., Dicke et al., 2018). On the other hand we did not include variables that may mediate the relation between SWB and academic achievement. Following the SDT (Ryan, & Deci, 2000), one mechanism explaining the relation from academic achievement on SWB could be that a good performance in school causes a higher self-concept and that this in turn could enhance well-being. Furthermore, effects from SWB on academic achievement could be explained by a mediation model, too: positive emotions could lead to an increase in interest and creativity which in turn could improve performance in school (e.g., Fredrickson, 2001).

#### **6.4 Implications for future research and educational practice**

The present study revealed new insights concerning the relation between SWB and academic achievement. However, further research is needed to support the presented findings. As has been mentioned above, the construct of SWB was not operationalized satisfactorily. To be able to draw conclusions which are more content-related, future studies should consider a greater number of facets of SWB to satisfy the complexity of the construct and should analyze students' SWB by means of well-validated instruments. Furthermore, it would be desirable to investigate the reciprocal relation between SWB and academic achievement on basis of at least three points of measurement and shorter time lags between each of them. This would enable researchers to distinguish stable variability between persons from within person variance (Hamaker et al., 2015).

Moreover, future studies should examine effects of additional variables. In this regard, variables like socioeconomic status and ethnic heritage should be included to control for pre-existing differences in achievement. The role of variables like self-concept, interest and divergent thinking that possibly mediate the relation between SWB and academic achievement should be investigated, as well. Because we found evidence that – especially – higher scores in MC were associated with higher scores of several aspects of SWB on a later time, it would be interesting to investigate a possible mediation through self-concept in mathematics on this relation. Furthermore, recent findings suggest that the association between positive emotions and academic achievement might be mediated through motivation (Mega, Ronconi, & De Beni, 2014). Therefore, it would be interesting to investigate if these results apply on other facets of well-being as well. Another desideratum in this context might be to investigate influences of variables on higher hierarchical levels on students' SWB or academic achievement. For example, the role of compositional effects of classes or the big-fish-little-pond-effect (Marsh, & Parker, 1984), which explains differences in academic self-concepts could be related to SWB of students. Big-fish-little-pond-effects are a result of academic achievement groupage which is very prominent in Germany (Köller, 2004). Additionally, teachers' SWB

should also be investigated more deeply, because teachers' SWB might have an influence on quality of teacher-student relationships, which, in turn, could influence the emotional-motivational characteristics of students. In order to study such research questions that require consideration of variables on the class (or the school) level, multi-level models should be specified to determine, for example, cross-level effects.

After a replication of the results, also implications for educational practice are possible. Due to the importance of students' SWB (e.g., satisfaction with school) for academic achievement, it would be thinkable to emphasize its relevance in training programs for teachers or to involve it in university curricula of teacher students. Furthermore, the influence of mathematical competence on SWB has to be investigated more deeply to derive implications for educational practice: If, for example, self-concept would mediate the relation between mathematical competence and well-being, intervention programs might be developed in order to strengthen mathematical self-concept of students, which also may have positive influences on their SWB.

## Acknowledgements

This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Grade 5, doi:10.5157/NEPS:SC3:8.0.1. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

## References

- Baumert, J., Maaz, K., Stanat, P., & Watermann, R. (2009). Schulkomposition oder Institution – was zählt? Schulstrukturen und die Entstehung schulformspezifischer Entwicklungsverläufe. *Die Deutsche Schule*, 101, 33–46.
- Berger, C., Alcalay, L., Torretti, A., & Milicic, N. (2011). Socio-emotional well-being and academic achievement: Evidence from a multilevel approach. *Psicología: Reflexão e Crítica*, 24, 344–351.
- Bird, J. M., & Markle, R. S. (2012). Subjective well-being in school environments: Promoting positive youth development through evidence-based assessment and intervention. *American Journal of Orthopsychiatry*, 82, 61–66.
- Blossfeld, H.-P. & Roßbach, H.-G. (Eds.). (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS)* (2nd ed.). Wiesbaden, Germany: VS.
- Blossfeld, H.-P., Roßbach, H.-G. & Maurice, J. von (Eds.). (2011). *Education as a life-long process. The German National Educational Panel Study (NEPS)*. Wiesbaden, Germany: VS.
- Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research*, 17(3), 303–316.

- Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago: Aldine.
- Brosseau-Liard, P. E., Savalei, V., & Li, L. (2012). An investigation of the sample performance of two nonnormality corrections for RMSEA. *Multivariate behavioral research*, *47*, 904–930.
- Bücker, S., Nuraydin, S., Simonsmeier, B. A., Schneider, M., & Luhmann, M. (2018). Subjective well-being and academic achievement: A meta-analysis. *Journal of Research in Personality*, *74*, 83–94.
- Casas, F., & González-Carrasco, M. (2019). Subjective well-being decreasing with age: New research on children over 8. *Child Development*, *90*, 375–394.
- Chang, L., McBride-Chang, C., Stewart, S. M., & Au, E. (2003). Life satisfaction, self-concept, and family relations in Chinese adolescents and children. *International Journal of Behavioral Development*, *27*, 182–189.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.
- Crede, J., Wirthwein, L., McElvany, N., & Steinmayr, R. (2015). Adolescents' academic achievement and life satisfaction: The role of parents' education. *Frontiers in Psychology*, *6*, 1–8.
- Czerwenka, K., Nölle, K., Pause, G., Schlotthaus, W., Schmidt, H. J., & Tesslo, J. (1990). *Schülerurteile über die Schule: Bericht über eine internationale Untersuchung*. Frankfurt am Main, Germany: Lang.
- Dicke, T., Marsh, H. W., Parker, P. D., Pekrun, R., Guo, J., & Televantou, I. (2018). Effects of school-average achievement on individual self-concept and achievement: Unmasking phantom effects masquerading as true compositional effects. *Journal of Educational Psychology*, *110*, 1112–1126.
- Diedrich, J., Schiepe-Tiska, A., Ziernwald, L., Tupac-Yupanqui, A., McElvany, N., ... (2019). Lesebezogene Schülermerkmale in PISA 2018: Motivation, Leseverhalten, Selbstkonzept und Lesestrategiewissen. In K. Reiss, M. Weis, E. Klieme, & O. Köller (Eds.), *PISA 2018. Grundbildung im internationalen Vergleich* (pp. 81–109). Münster, Germany: Waxmann.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*, 542–575.
- Diener, E., & Emmons, R. A. (1984). The independence of positive and negative affect. *Journal of personality and social psychology*, *47*(5), 1105–1117.
- Diener, E., Inglehart, R., & Tay, L. (2013). Theory and validity of life satisfaction scales. *Social Indicators Research*, *112*, 497–527.
- Diener, E., Suh, E. M., Lucas, R. R., & Smith, H. L. (1999). Subjective well-being: three decades of progress. *Psychological Bulletin*, *125*, 276–302.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*, 1040–1048.
- Eid, M., & Larsen, R. J. (2008). *The Science of subjective well-being*. New York, NY: Guilford Press.
- Eide, E. R., Showalter, M. H., & Goldhaber, D. D. (2010). The relation between children's health and academic achievement. *Children and Youth Services Review*, *32*, 231–238.
- Farmer, H. S., & Vispoel, W. P. (1990). Attributions of female and male adolescents for real-life failure experiences. *Journal of Experimental Education*, *58*, 127–140.
- Fend, H., Knörzer, W., Nagl, W., Specht, W., & Väth-Szusdziara, R. (1976). *Sozialisationseffekte der Schule*. Weinheim, Germany: Beltz.
- Fincham, F. D., Hokoda, A., & Sanders, R. (1989). Learned helplessness, test anxiety, and academic achievement: A longitudinal analysis. *Child Development*, *60*, 1461–1474.
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*, 218–226.

- Gehrer, K., Zimmermann, S., Artelt, C., & Weinert, S. (2012). *The assessment of reading competence (including sample items for grade 5 and 9): Scientific Use File 2012, Version 1.0.0.* Bamberg: University of Bamberg, National Educational Panel Study (NEPS).
- Gilman, R., & Huebner, E. S. (2006). Characteristics of adolescents who report very high life satisfaction. *Journal of Youth and Adolescence*, 35, 311–319.
- Gumora, G., & Arsenio, W. F. (2002). Emotionality, emotion regulation, and school performance in middle school children. *Journal of School Psychology*, 40, 395–413.
- Gysin, S. H. (2017). *Subjektives Wohlbefinden von Schülerinnen und Schülern.* Weinheim, Germany: Beltz Juventa.
- Haecker, H., & Werres, W. (1983). *Schule und Unterricht im Urteil der Schüler: Bericht einer Schülerbefragung in der Sekundarstufe I.* Frankfurt am Main, Germany: Lang.
- Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, 20(1), 102–116.
- Hascher, T. (2007). Exploring students' well-being by taking a variety of looks in the classroom. *Hellenic Journal of Psychology*, 4, 331–349.
- Hascher, T. (2008). Quantitative and qualitative research approaches to assess student well-being. *International Journal of Educational Research*, 47, 84–96.
- Hascher, T., & Hagenauer, G. (2011). Wohlbefinden und Emotionen in der Schule als zentrale Elemente des Schulerfolgs unter der Perspektive geschlechtsspezifischer Ungleichheiten. In A. Hadjar (Eds.), *Geschlechtsspezifische Bildungsungleichheiten* (pp. 285–308). Wiesbaden, Germany: VS.
- Hascher, T., & Winkler-Ebner, C. (2010). Gesundheit und Bildung von Kindern und Jugendlichen. In P. Paulus (Ed.), *Bildungsförderung durch Gesundheit. Bestandsaufnahme und Perspektiven für eine gute gesunde Schule* (pp. 31–56). Weinheim, Germany: Juventa.
- Heffner, A. L., & Antaramian, S. P. (2016). The role of life satisfaction in predicting student engagement and achievement. *Journal of Happiness*, 17, 1681–1701.
- Herman, K. C., Lambert, S. F., Reinke, W. M., & Ialongo, N. S. (2008). Low academic competence in first grade as a risk factor for depressive cognitions and symptoms in middle school. *Journal of Counseling Psychology*, 55, 400–410.
- Huebner, E. S. (1991). Correlates of life satisfaction in children. *School Psychology Quarterly*, 6, 103–111.
- Huebner, E. S., & Alderman, G. L. (1993). Convergent and discriminant validation of a children's life satisfaction scale: Its relationship to self and teacher-reported psychological problems and school functioning. *Social Indicators Research*, 30, 71–82.
- Jerusalem, M., & Schwarzer, R. (1993). Dimensionen der Hilflosigkeit. In G. Westhoff (Ed.), *Handbuch psychosozialer Meßinstrumente: Ein Kompendium für epidemiologische und klinische Forschung zu chronischer Krankheit* (pp. 402–404). Göttingen, Germany: Hogrefe.
- Jöreskog, K. G., & Sörbom, D. (1982). Recent developments in structural equation modeling. *Journal of marketing research*, 19(4), 404–416.
- Köller, O. (2004). *Konsequenzen von Leistungsgruppierungen [Consequences of achievement groupage].* Münster, Germany: Waxmann.
- Lavy, V., & Sand, E. (2012). *The friends factor: How students' social networks affect their academic achievement and well-being?* Cambridge, MA: National Bureau of Economic Research.
- Linnenbrink, E. A., & Pintrich, P. R. (2002). Achievement goal theory and affect: An asymmetrical bidirectional model. *Educational Psychologist*, 37, 69–78.
- Marsh, H. W., & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of personality and social psychology*, 47(1), 213–231.

- Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology, 106*, 121–131.
- Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research, 79*, 1129–1167.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological methods, 7*(1), 105–125.
- Neumann, I., Duchhardt, C., Grüßing, M., Heinze, A., Knopp, E., & Ehmke, T. (2013). Modeling and assessing Mathematical competence over the lifespan. *Journal for Educational Research Online, 5*(2), 80–109.
- Ng, Z. J., Huebner, S. E., & Hills, K. J. (2015). Life satisfaction and academic performance in early adolescents: Evidence for reciprocal association. *Journal of school psychology, 53*, 479–491.
- Opdenakker, M. C., & Van Damme, J. (2000). Effects of schools, teaching staff and classes on achievement and well-being in secondary education: Similarities and differences between school outcomes. *School effectiveness and school improvement, 11*, 165–196.
- OECD – Organisation for Economic Cooperation Development. (2017). *PISA 2015. Results (Volume III): Students' Well-Being*. Paris, France: OECD.
- Palsdottir, A., Asgeirsottir, B. B., & Sigfusdottir, I. D. (2012). Gender difference in wellbeing during school lessons among 10–12-year-old children: The importance of school subjects and student-teacher relationships. *Scandinavian Journal of Public Health, 40*, 605–613.
- Petrescu, M. (2013). Marketing research using single-item indicators in structural equation models. *Journal of Marketing Analytics, 1*(2), 99–117.
- Pietarinen, J., Soini, T., & Pyhältö, K. (2014) Students' emotional and cognitive engagement as the determinants of well-being and achievement in school. *International Journal of Educational Research, 67*, 40–51.
- Proctor, C., Linley, P. A., & Maltby, J. (2010). Very happy youths: Benefits of very high life satisfaction among adolescents. *Social Indicators Research, 98*, 519–532.
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental review, 41*, 71–90.
- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.R-project.org/>.
- Richter, M., & Hurrelmann, K. (Eds.). (2009). *Gesundheitliche Ungleichheit: Grundlagen, Probleme, Konzepte* (2nd ed.). Wiesbaden, Germany: VS.
- Rohlf, C. (2011). Bildungseinstellungen. Schule und formale Bildung aus der Perspektive von Schülerinnen und Schülern. Wiesbaden, Germany: VS.
- Rosseel Y (2012). “lavaan: An R Package for Structural Equation Modeling.” *Journal of Statistical Software, 48*(2), 1–36. <http://www.jstatsoft.org/v48/i02/>.
- Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological bulletin, 125*(1), 3–30.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68–78.
- Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology, 52*, 141–166.
- Samdal, O., Wold, B., & Bronis, M. (1999). Relationship between Students' Perceptions of School Environment, Their Satisfaction with School and Perceived Academic

- Achievement: An International Study. *School Effectiveness and School Improvement*, 10, 296–320.
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.
- Scharl, A., Fischer, L., Gnambs, T., & Rohm, T. (2017). NEPS Technical Report for Reading—Scaling Results of Starting Cohort 3 for Grade 9.
- Schimmack, U. (2008). Structure of subjective well-being. In M. Eid & R. J. Larsen (Eds.), *The science of subjective well-being* (pp. 97–123). New York, NY: Guilford Press.
- Seligman, M. E. (1972). Learned helplessness. *Annual review of medicine*, 23, 407–412.
- Seligman, M. E. (2011). *Flourish: A visionary new understanding of happiness and well-being*. New York, NY: Free Press.
- Seligman, M. E., Ernst, R. M., Gillham, J., Reivich, K., & Linkins, M. (2009). Positive education: Positive psychology and classroom interventions. *Oxford Review of Education*, 35, 293–311.
- Sigfusdóttir, I. D., Kristjánsson, A. L., & Allegrante, J. P. (2007). Health behaviour and academic achievement in Icelandic school children. *Health Education Research*, 22, 70–80.
- Sparfeldt, J. R., Buch, S. R., Schwarz, F., Jachmann, J., & Rost, D. H. (2009). “Rechnen ist langweilig” – Langeweile in Mathematik bei Grundschülern. *Psychologie in Erziehung und Unterricht*, 56, 16–26.
- Steinmayr, R., Crede, J., McElvany, N., & Wirthwein, L. (2016). Subjective wellbeing, test anxiety, academic achievement: Testing for reciprocal effects. *Frontiers in Psychology*, 6, 1–13.
- Suldo, S. M., Riley, K. N., & Shaffner, E. J. (2006). Academic correlates of children and adolescents’ life satisfaction. *School Psychology International*, 27, 567–582.
- Suldo, S. M., & Shaffer, E. J. (2008). Looking beyond psychopathology: The dual factor model of mental health in youth. *School Psychology Review*, 37, 52–68.
- Suldo, S. M., Shaffer, E. J., & Riley, K. N. (2008). A social-cognitive-behavioral model of academic predictors of adolescents’ life satisfaction. *School Psychology Quarterly*, 23, 56–69.
- Urhahne, D., & Zhu, M. (2015). Accuracy of teachers’ judgments of students’ subjective well-being. *Learning and Individual Differences*, 43, 226–232.
- van Petegem, K., Aelterman, A., Rossell, Y., & Creemers, B. (2006). Student perception as moderator for student well-being. *Social Indicators Research*, 83, 447–463.
- Verkuyten, M., & Thijs, J. (2002). School satisfaction of elementary school children: the role of performance, peer relations, ethnicity, and gender. *Social Indicators Research*, 59, 203–228.
- Véronneau, M. H., & Dishion, T. J. (2012). Middle school friendships and academic achievement in early adolescence: A longitudinal analysis. *Journal Early Adolescence*, 31, 99–124.
- Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). Washington, DC: American Psychological Association.
- Wilson, W. R. (1967). Correlates of avowed happiness. *Psychological bulletin*, 67, 294–306.
- WHO – World Health Organization. (2014). Basic documents: Forty-eighth edition. Geneva, Switzerland: World Health Organization.
- Zajacova, A., Lynch, S. M., & Espenshade, T. J. (2005). Self-efficacy, stress, and academic success in college. *Research in Higher Education*, 46, 677–706.

## Appendix

### Appendix A

**Table A1:** Means, standard deviations and skewness for the variables used in the analyses of the present study

Variable	Grade 5 <sup>a</sup>		Grade 7		Grade 9	
	M (SD)	Skewness	M (SD)	Skewness	M (SD)	Skewness
DA	-		1.44 (4.09)	10.79	1.36 (2.52)	3.95
SH	-		4.25 (0.78)	-0.94	4.14 (0.80)	-0.77
SL	8.89 (1.51)	-3.14	8.39 (1.43) <sup>a</sup>	-1.62 <sup>a</sup>	8.13 (1.66) <sup>a</sup>	-1.73 <sup>a</sup>
SS	8.06 (2.25)	-1.59	7.10 (2.31) <sup>a</sup>	-1.00 <sup>a</sup>	6.94 (2.26) <sup>a</sup>	-0.97 <sup>a</sup>
HG	-		1.77 (0.69)	0.87	1.80 (0.69)	0.68
HM	-		1.68 (0.71)	0.96	1.74 (0.75)	0.85
MC	0.24 (1.14)	0.00	0.89 (1.22)	-0.01	0.15 (1.20)	0.20
RC	0.24 (1.23)	0.17	0.84 (1.37)	0.20	0.10 (1.13)	0.20

*Notes.* n = 4180. DA = Days of absence from school, SH = Self-estimated health, SL = Satisfaction with life, SS = Satisfaction with school, HG = Helplessness German, HM = Helplessness Math, MC = Mathematical competence, RC = Reading competence. Values for MC and RC are WLEs.

<sup>a</sup>n = 2902 students.

Table A2: Means, standard deviations and skewness for the variables used in the analyses of the present study separately for both sexes

Variable	Male <sup>a</sup>						Female <sup>b</sup>					
	Grade 5 <sup>c</sup>		Grade 7		Grade 9		Grade 5 <sup>d</sup>		Grade 7		Grade 9	
		M (SD)	Skewness	M (SD)	Skewness	M (SD)	Skewness	M (SD)	Skewness	M (SD)	Skewness	M (SD)
DA	-	1.63 (4.96)	10.57	1.33 (2.70)	4.51	-	-	1.24 (2.94)	6.59	1.39 (2.34)	3.05	
SH	-	4.26 (0.78)	-1.02	4.22 (0.79)	-0.90	-	-	4.25 (0.77)	-0.87	4.06 (0.79)	-0.67	
SL	8.83 (1.54)	-3.08	8.37 (1.48) <sup>c</sup>	-1.85 <sup>c</sup>	8.19 (1.73) <sup>c</sup>	-1.87 <sup>c</sup>	8.94 (1.48)	-3.21	8.41 (1.39) <sup>d</sup>	-1.33 <sup>d</sup>	8.08 (1.58) <sup>d</sup>	-1.56 <sup>d</sup>
SS	7.90 (2.36)	-1.46	6.95 (2.38) <sup>c</sup>	-0.94 <sup>c</sup>	6.88 (2.33) <sup>c</sup>	-0.92 <sup>c</sup>	8.22 (2.12)	-1.73	7.25 (2.23) <sup>d</sup>	-1.06 <sup>d</sup>	7.01 (2.19) <sup>d</sup>	-1.03 <sup>d</sup>
HG	-	1.86 (0.72)	0.71	1.90 (0.73)	0.56	-	-	1.68 (0.65)	1.04	1.70 (0.64)	0.77	
HM	-	1.60 (0.69)	1.07	1.64 (0.72)	1.00	-	-	1.75 (0.71)	0.87	1.84 (0.77)	0.71	
MC	0.37 (1.14)	1.21	1.06 (1.27)	-0.02	0.29 (1.25)	0.16	0.11 (1.12)	-0.02	0.73 (1.14)	-0.10	0.00 (1.12)	0.17
RC	0.17 (1.27)	0.16	0.69 (1.39)	0.23	0.01 (1.17)	0.26	0.32 (1.19)	0.21	0.98 (1.32)	0.21	0.21 (1.06)	0.20

Notes. DA = Days of absence from school, SH = Self-estimated health, SL = Satisfaction with life, SS = Satisfaction with school, HG = Helplessness German, HM = Helplessness Math, MC = Mathematical competence, RC = Reading competence. Values for MC and RC are weighted likelihood estimates.  
<sup>a</sup>n = 2109, <sup>b</sup>n = 2071, <sup>c</sup>n = 1461, <sup>d</sup>n = 1441.

Table A3: Means and standard deviations for the variables used in the analyses of the present study separately for different types of school

Variable	Hauptschule <sup>a</sup>						Realschule <sup>b</sup>						Gymnasium <sup>c</sup>					
	Grade 5 <sup>d</sup>		Grade 7		Grade 9		Grade 5 <sup>e</sup>		Grade 7		Grade 9		Grade 5 <sup>f</sup>		Grade 7		Grade 9	
	M (SD)	SK	M (SD)	SK	M (SD)	SK	M (SD)	SK	M (SD)	SK	M (SD)	SK	M (SD)	SK	M (SD)	SK	M (SD)	SK
DA	-	2.93 (8.67)	6.31	1.84 (3.21)	3.77	-	1.58 (3.86)	6.64	1.53 (2.92)	4.42	-	-	1.01 (2.33)	5.14 (2.23)	1.21 (2.23)	3.51		
SH	-	4.25 (0.85)	-1.03	4.11 (0.80)	-0.46	-	4.22 (0.77)	-0.85	4.08 (0.80)	-0.78	-	-	4.29 (0.77)	-1.01 (0.79)	4.17 (0.79)	-0.83		
SL	8.48 (1.89)	-2.31 (1.68) <sup>d</sup>	8.23	-1.21 <sup>d</sup>	8.15 (1.84) <sup>d</sup>	-1.64 <sup>d</sup> (1.91)	8.63 (1.91)	-2.72 (1.42) <sup>e</sup>	8.34 (1.75) <sup>e</sup>	-1.42	8.03 (1.75) <sup>e</sup>	-1.55 (1.15)	9.08 (1.15)	-3.24 (1.38) <sup>f</sup>	8.44 (1.38) <sup>f</sup>	-1.80 <sup>f</sup> (1.57) <sup>f</sup>	-1.84 <sup>f</sup>	
SS	7.33 (2.84)	-1.02 (2.64) <sup>d</sup>	6.76 (2.38) <sup>d</sup>	-0.79 <sup>d</sup>	7.02 (2.38) <sup>d</sup>	-0.91 <sup>d</sup>	7.68 (2.54)	-1.29 (2.44) <sup>e</sup>	6.82 (2.28) <sup>e</sup>	-0.90 (2.28) <sup>e</sup>	6.78 (1.91)	-0.90 (1.91)	8.37 (1.91)	-1.80 (2.16) <sup>f</sup>	7.29 (2.16) <sup>f</sup>	-1.06 <sup>f</sup> (2.22) <sup>f</sup>	7.00 (2.22) <sup>f</sup>	
HG	-	2.03 (0.81)	0.46	1.89 (0.76)	0.66	-	1.85 (0.73)	0.73 (0.69)	1.83 (0.69)	0.62	-	-	1.65 (0.64)	1.05 (0.68)	1.73 (0.68)	0.85		
HM	-	1.96 (0.84)	0.53	1.89 (0.84)	0.73	-	1.78 (0.75)	0.79 (0.72)	1.71 (0.72)	0.81	-	-	1.58 (0.66)	1.17 (0.73)	1.69 (0.73)	0.91		
MC	-1.17 (0.91)	0.19 (0.97)	-0.64 (0.91)	0.09 (0.91)	-1.09 (0.84)	0.44 (0.84)	-0.21 (0.97)	-0.00 (0.97)	0.41 (0.93)	0.19 (0.93)	-0.34 (0.96)	0.29 (0.96)	0.73 (0.96)	0.29 (1.02)	1.42 (1.02)	0.17 (1.06)	0.67 (1.06)	
RC	-1.11 (0.98)	0.51 (1.10)	-0.62 (0.83)	0.14 (1.04)	-1.04 (1.04)	0.32 (1.04)	-0.15 (1.12)	0.37 (1.12)	0.30 (0.90)	0.36 (0.90)	-0.26 (0.90)	0.38 (0.90)	0.69 (0.90)	0.37 (0.90)	1.30 (1.23)	0.26 (1.23)	0.55 (1.03)	0.18

Notes. <sup>a</sup>n = 4180. DA = Days of absence from school, SH = Self-estimated health, SL = Satisfaction with life, SS = Satisfaction with school, HG = Helplessness German, HM = Helplessness Math, MC = Mathematical competence, RC = Reading competence, SK = Skewness. Values for MC and RC are weighted likelihood estimates.  
<sup>b</sup>n = 583, <sup>c</sup>n = 1176, <sup>d</sup>n = 2431, <sup>e</sup>n = 363, <sup>f</sup>n = 789, <sup>†</sup>n = 1750.

## Appendix B

Table B1: Correlations between all variables included in the present study

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.
1. DA <sub>t<sub>2</sub></sub>																				
2. DA <sub>t<sub>3</sub></sub>	.13***																			
3. SH <sub>t<sub>2</sub></sub>	-.22***	-.08***																		
4. SH <sub>t<sub>3</sub></sub>	-.11***	-.15***	-.40***																	
5. SL <sub>t<sub>1</sub></sub>	-.07***	-.03	.16***	.13***																
6. SL <sub>t<sub>2</sub></sub>	-.11**	-.08**	.45***	.26***	.30***															
7. SL <sub>t<sub>3</sub></sub>	-.07**	-.11***	.27***	.43***	.20***	.40***														
8. SS <sub>t<sub>1</sub></sub>	-.06**	-.06**	.13***	.11***	.57***	.25***	.15***													
9. SS <sub>t<sub>2</sub></sub>	-.11***	-.09***	.27***	.17***	.19***	.50***	.27***	.29***												
10. SS <sub>t<sub>3</sub></sub>	-.08***	-.12***	.19***	.25***	.11***	.25***	.52***	.52***	.16***	.38***										
11. HG <sub>t<sub>2</sub></sub>	.06**	.09***	-.13***	-.09***	-.15***	-.27***	-.13***	-.13***	-.21***	-.36***	-.21***									
12. HG <sub>t<sub>3</sub></sub>	.04	.06**	-.08***	-.06**	-.08***	-.14***	-.12***	-.12***	-.16***	-.22***	-.26***	.37***								
13. HM <sub>t<sub>2</sub></sub>	.10***	.08***	-.12***	-.08***	-.13***	-.21***	-.15***	-.15***	-.17***	-.32***	-.22***	.34***	.13***							
14. HM <sub>t<sub>3</sub></sub>	.05*	.12***	-.09***	-.11***	-.04	-.13***	-.16***	-.16***	-.09***	-.21***	-.30***	.08***	.16***	.38***						
15. MC <sub>t<sub>1</sub></sub>	-.12***	-.12***	.02	-.00	.06***	.00	-.00	.11***	.14***	.10***	.13***	.08***	.25***	.21***						
16. MC <sub>t<sub>2</sub></sub>	-.10***	-.11***	.05***	.03	.07***	.03	.02	.11***	.13***	.14***	.17***	.06***	.28***	.22***	.73***					
17. MC <sub>t<sub>3</sub></sub>	-.10***	-.13***	.05***	.02	.06***	.01	-.00	.10***	.14***	.16***	.15***	.07***	.28***	.27***	.70***	.73***				
18. RC <sub>t<sub>1</sub></sub>	-.10***	-.07***	-.00	-.05*	.07***	-.01	-.05*	.11***	.13***	.07***	.23***	-.17***	-.16***	-.07***	.62***	.55***				
19. RC <sub>t<sub>2</sub></sub>	-.09***	-.09***	.03	-.02	.11***	.06**	.02	.11***	.19***	.14***	.27***	-.16***	-.19***	-.11***	.54***	.61***	.57***	.61***		
20. RC <sub>t<sub>3</sub></sub>	-.07***	-.08***	.01	-.02	.06***	.01	-.02	.12***	.17***	.12***	.25***	-.21***	-.20***	-.14***	.53***	.53***	.60***	.61***	.63***	
21. Gender	-.04***	.01	-.01	-.10***	.04	.01**	-.03	.07***	.03	-.13***	-.15***	.10***	.13***	-.12***	-.14***	-.12***	-.12***	.06***	.10***	.10***

Notes. DA = Days of absence from school, SH = Self-estimated health, SL = Satisfaction with life, SS = Satisfaction with school, HG = Helplessness German, HM = Helplessness Math, MC = Mathematical competence, RC = Reading competence, t<sub>1</sub> = measurement in grade 5, t<sub>2</sub> = measurement in grade 7, t<sub>3</sub> = measurement in grade 9.

\*p < .05. \*\*p < .01. \*\*\*p < .001.

## Appendix C

In the following you can see the intraclass correlation coefficients (ICC) for the variables we used in our analyses. Sample sizes differed depending on the number of measurement time points used in the respective analysis. Therefore, ICCs are displayed once for variables included in the analysis over three points of measurement (see Table C1) and once for the variables included in the analyses over two points of measurement (see Table C2). As can be seen from the tables the ICC values of the variables mathematical competence and reading comprehension are very high. This is not a surprising finding because the cluster variable contains schools from different types of school and those should vary highly on academic achievement.

**Table C1:** ICC values of variables included in the analysis over  $t_1$  to  $t_3$

	$t_1$	$t_2$	$t_3$
Satisfaction with life	.04	.05	.03
Satisfaction with school	.06	.03	.03
Mathematical competence	.41	.41	.38
Reading comprehension	.31	.31	.33

**Table C2:** ICC values of variables included in the analyses over  $t_2$  to  $t_3$

	$t_2$	$t_3$
Days of absence from school	.03	.02
Self-estimated health	.03	.02
Helplessness German	.07	.03
Helplessness Math	.06	.05
Mathematical competence	.41	.38
Reading comprehension	.32	.33

## Appendix D

Measurement invariance of the helplessness scale had to be tested between different points of measurement, as well as between sexes and types of school. Therefore, we started by testing measurement invariance between different occasions. More precisely, we first checked measurement invariance between points of measurement considering all individuals and afterwards by focusing on different groups of gender and type of school (see Table D1). Subsequently, we tested measurement invariance between different sexes and types of school for single points of measurement (see Table D2). Configural invariance over points of measurement was achieved through excluding items 4 (“When my teacher calls me surprisingly in German I cannot answer even the simplest questions”) and 5 (“No matter if I try to do my homework in German, I always make many mistakes”) of the subscale *helplessness German* and 1 (“No matter how much I try in Math, my grades won’t get better”) and 2 (“It’s not worth practicing mathematics for a class test, I will be bad again”) of the subscale *helplessness Math*. Tables D1 and D2 show that – according to  $\Delta\text{CFI}$  – metric invariance could be assumed for the helplessness scale.

Table D1: Examination of measurement invariance of the helplessness scale over points of measurement

Model	$\chi^2$	Df	p	CFI/TLI	RMSEA/SRMR	Comparison	$\Delta\text{CFI}$	$\Delta\chi^2$	$\Delta df$	$p(\Delta\chi^2)$
Overall:										
1. configural	216.66	42	<.001	.989/.982	.039/.025					
2. metric	237.89	46	<.001	.987/.982	.039/.027	2 vs. 1	.002	21.23	4	<.001
3. scalar	295.90	52	<.001	.984/.980	.041/.028	3 vs. 2	.003	62.63	6	<.001
4. residual	302.44	58	<.001	.983/.981	.040/.028	4 vs. 2	.004	64.23	12	<.001
Gender = male <sup>a</sup> :										
1. configural	114.68	42	<.001	.990/.984	.036/.025					
2. metric	129.82	46	<.001	.989/.984	.037/.028	2 vs. 1	.001	15.04	4	<.01
3. scalar	160.80	52	<.001	.985/.981	.040/.029	3 vs. 2	.004	33.87	6	<.001
4. residual	162.31	58	<.001	.985/.983	.038/.028	4 vs. 2	.004	32.68	12	<.01
Gender = female <sup>b</sup> :										
1. configural	147.27	42	<.001	.987/.979	.042/.027					
2. metric	155.09	46	<.001	.986/.980	.041/.029	2 vs. 1	.001	8.02	4	.091
3. scalar	187.60	52	<.001	.983/.978	.042/.031	3 vs. 2	.003	34.08	6	<.001
4. residual	193.56	58	<.001	.982/.979	.041/.032	4 vs. 2	.004	38.75	12	<.001
Type of school = "Hauptschule" <sup>c</sup> :										
1. configural	85.71	42	<.001	.978/.966	.056/.039					
2. metric	92.93	46	<.001	.977/.967	.055/.041	2 vs. 1	.001	7.05	4	.134
3. scalar	108.87	52	<.001	.973/.965	.057/.047	3 vs. 2	.004	16.52	6	<.05
4. residual	118.15	58	<.001	.969/.965	.057/.046	4 vs. 2	.008	25.09	12	<.05
Type of school = "Realschule" <sup>d</sup> :										
1. configural	105.21	42	<.001	.985/.977	.044/.029					
2. metric	111.60	46	<.001	.985/.978	.043/.030	2 vs. 1	.000	6.59	4	.159
3. scalar	127.40	52	<.001	.983/.978	.043/.030	3 vs. 2	.002	16.01	6	<.05
4. residual	144.96	58	<.001	.979/.976	.045/.031	4 vs. 2	.006	32.83	12	<.01
Type of school = "Gymnasium" <sup>e</sup> :										
1. configural	134.45	42	<.001	.990/.984	.036/.025					
2. metric	156.61	46	<.001	.988/.982	.038/.029	2 vs. 1	.002	21.76	4	<.001
3. scalar	258.28	52	<.001	.978/.972	.048/.038	3 vs. 2	.010	113.14	6	<.001
4. residual	263.65	58	<.001	.976/.973	.047/.039	4 vs. 2	.012	97.69	12	<.001

Notes. N = 4180,  $\chi^2$  = Chi-Square, df = degrees of freedom. CFI = Comparative Fit Index. TLI = Tucker-Lewis index. RMSEA = Root mean square error of approximation. SRMR = Standardized root mean square residual.  $p(\Delta\chi^2)$ -values are calculated based on the Satorra-Bentler scaled chi-square difference test (Satorra & Bentler, 1994).

<sup>a</sup>n = 2109, <sup>b</sup>n = 2071, <sup>c</sup>n = 583, <sup>d</sup>n = 1166, <sup>e</sup>n = 2431.

**Table D2:** Examination of measurement invariance of the helplessness scale over gender and types of school on single occasions

Model	$\chi^2$	Df	p	CFI/TLI	RMSEA/SRMR	Comparison	$\Delta CFI$	$\Delta \chi^2$	$\Delta df$	$p(\Delta \chi^2)$
Over gender on $t_2$ :										
1. configural	121.01	16	<.001	.985/.971	.070/.030					
2. metric	134.76	20	<.001	.984/.976	.064/.032	2 vs. 1	.001	8.19	4	.085
3. scalar	277.86	26	<.001	.967/.962	.081/.059	3 vs. 2	.017	171.49	6	<.001
4. residual	286.95	32	<.001	.963/.965	.077/.060	4 vs. 2	.021	147.21	12	<.001
Over gender on $t_3$ :										
1. configural	68.59	16	<.001	.993/.987	.048/.023					
2. metric	76.04	20	<.001	.993/.989	.044/.026	2 vs. 1	.000	7.48	4	.112
3. scalar	224.53	26	<.001	.975/.971	.071/.064	3 vs. 2	.018	172.34	6	<.001
4. residual	229.79	32	<.001	.972/.974	.068/.065	4 vs. 2	.021	143.07	12	<.001
Over types of school on $t_2$ :										
1. configural	111.16	24	<.001	.987/.976	.063/.029					
2. metric	127.45	32	<.001	.986/.981	.056/.031	2 vs. 1	.001	13.46	8	.097
3. scalar	289.95	44	<.001	.968/.967	.074/.074	3 vs. 2	.018	200.59	12	<.001
4. residual	383.42	56	<.001	.952/.962	.080/.077	4 vs. 2	.034	251.47	24	<.001
Over types of school on $t_3$ :										
1. configural	73.26	24	<.001	.993/.988	.047/.023					
2. metric	87.51	32	<.001	.993/.990	.043/.027	2 vs. 1	.000	13.59	8	.093
3. scalar	161.57	44	<.001	.986/.985	.051/.040	3 vs. 2	.007	82.95	12	<.001
4. residual	175.97	56	<.001	.983/.986	.049/.044	4 vs. 2	.010	78.45	24	<.001

Notes. N = 4180.  $\chi^2$  = Chi-Square, df = degrees of freedom. CFI = Comparative Fit Index. TLI = Tucker-Lewis index. RMSEA = Root mean square error of approximation. SRMR = Standardized root mean square residual.  $p(\Delta \chi^2)$ -values are calculated based on the Satorra-Bentler scaled chi-square difference test (Satorra & Bentler, 1994).

Similar to the procedure described above, measurement invariance of the satisfaction scale had to be tested between different points of measurement, as well as between sexes and types of school. Therefore, we started by testing measurement invariance between different occasions. More precisely, we first checked measurement invariance between points of measurement considering all individuals and afterwards by focusing on different groups of gender and type of school (see Table D3). Subsequently, we tested measurement invariance between different sexes and types of school for single points of measurement (see Table D4). Because we were interested in differences between general life satisfaction and satisfaction with school, we excluded one item (“How satisfied are you with your school situation?”) and calculated measurement invariance for the remaining 5 items. As can be seen from Tables D3 and D4 – according to  $\Delta\text{CFI}$  – at least configural invariance was given for the satisfaction scale. Furthermore, metric invariance was given for most measurement invariance analyses, as well. Only the measurement invariance analyses of female students over points of measurement and students attending “Hauptschule” over points of measurement represented an exception of this.

Table D3: Examination of measurement invariance of the satisfaction scale over points of measurement

Model	$\chi^2$	Df	p	CFI/TLI	RMSSEA/SRMR	Comparison	$\Delta CFI$	$\Delta \chi^2$	$\Delta df$	$p(\Delta \chi^2)$
Overall:										
1. configural	151.58	72	<.001	.991/.986	.028/.026					<.001
2. metric	204.21	80	<.001	.985/.980	.033/.039	2 vs. 1	.006	47.27	8	<.001
3. scalar	484.99	90	<.001	.953/.946	.055/.068	3 vs. 2	.032	355.20	10	<.001
4. residual	602.09	100	<.001	.936/.933	.061/.073	4 vs. 2	.049	339.37	20	<.001
Gender = male <sup>a</sup> :										
1. configural	89.07	72	<.001	.996/.994	.018/.028					
2. metric	98.89	80	<.001	.996/.994	.018/.033	2 vs. 1	.000	9.82	8	.278
3. scalar	216.67	90	<.001	.971/.967	.044/.056	3 vs. 2	.025	152.49	10	<.001
4. residual	288.24	100	<.001	.954/.952	.052/.061	4 vs. 2	.042	164.12	20	<.001
Gender = female <sup>b</sup> :										
1. configural	132.99	72	<.001	.986/.980	.034/.031					
2. metric	205.88	80	<.001	.970/.961	.047/.055	2 vs. 1	.016	60.74	8	<.001
3. scalar	401.39	90	<.001	.928/.916	.069/.088	3 vs. 2	.042	231.15	10	<.001
4. residual	511.66	100	<.001	.895/.890	.079/.096	4 vs. 2	.075	252.25	20	<.001
Type of school = "Hauptschule" <sup>c</sup> :										
1. configural	80.16	72	<.001	.992/.989	.024/.049					
2. metric	100.66	80	<.001	.981/.975	.037/.065	2 vs. 1	.011	20.61	8	<.01
3. scalar	125.96	90	<.001	.968/.962	.045/.065	3 vs. 2	.013	27.87	10	<.01
4. residual	170.53	100	<.001	.931/.928	.063/.076	4 vs. 2	.050	61.86	20	<.001
Type of school = "Realschule" <sup>d</sup> :										
1. configural	96.38	72	<.001	.991/.986	.028/.032					
2. metric	125.22	80	<.001	.982/.977	.037/.049	2 vs. 1	.009	27.36	8	<.001
3. scalar	187.96	90	<.001	.963/.957	.051/.061	3 vs. 2	.019	78.09	10	<.001
4. residual	252.07	100	<.001	.937/.934	.063/.069	4 vs. 2	.045	112.81	20	<.001
Type of school = "Gymnasium" <sup>e</sup> :										
1. configural	123.11	72	<.001	.990/.985	.028/.031					
2. metric	153.34	80	<.001	.985/.980	.033/.042	2 vs. 1	.005	26.04	8	<.01
3. scalar	452.76	90	<.001	.925/.913	.068/.097	3 vs. 2	.060	356.85	10	<.001
4. residual	569.01	100	<.001	.895/.890	.076/.105	4 vs. 2	.090	343.28	20	<.001

Notes. N = 2902.  $\chi^2$  = Chi-Square, df = degrees of freedom. CFI = Comparative Fit Index. TLI = Tucker-Lewis index. RMSEA = Root mean square error of approximation. SRMR = Standardized root mean square residual.  $p(\Delta \chi^2)$ -values are calculated based on the Satorra-Bentler scaled chi-square difference test (Satorra & Bentler, 1994).

<sup>a</sup>n = 2109, <sup>b</sup>n = 2071, <sup>c</sup>n = 583, <sup>d</sup>n = 1166, <sup>e</sup>n = 2431.

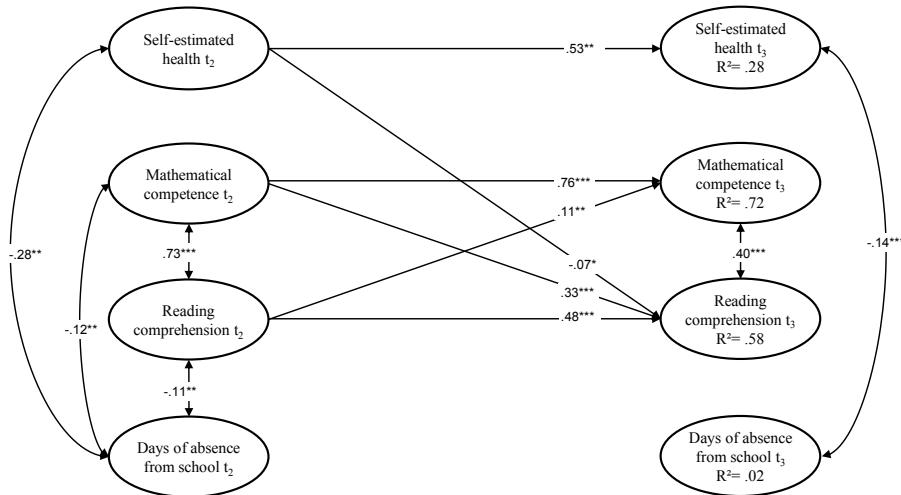
Table D4: Examination of measurement invariance of the satisfaction scale over gender and types of school on single occasions

Model	$\chi^2$	Df	p	CFI/TLI	RMSSEA/SRMR	Comparison	$\Delta CFI$	$\Delta \chi^2$	$\Delta df$	$p(\Delta \chi^2)$
Over gender on $t_1$ :										
1. configural	30.05	10	<.001	.991/0.982	.056/.018					
2. metric	34.71	14	<.001	.991/0.987	.048/.025	2 vs. 1	.000	4.16	4	.385
3. scalar	46.33	19	<.001	.989/0.989	.044/.030	3 vs. 2	.002	11.07	5	.050
4. residual	66.78	24	<.001	.978/0.982	.055/.043	4 vs. 2	.013	31.07	10	<.001
Over gender on $t_2$ :										
1. configural	22.47	10	<.001	.993/0.986	.041/.016					
2. metric	41.56	14	<.001	.985/0.979	.051/.037	2 vs. 1	.008	19.55	4	<.001
3. scalar	72.77	19	<.001	.974/0.973	.058/.042	3 vs. 2	.011	41.53	5	<.001
4. residual	79.09	24	<.001	.971/0.976	.054/.049	4 vs. 2	.014	38.04	10	<.001
Over gender on $t_3$ :										
1. configural	22.19	10	<.001	.995/0.990	.045/.014					
2. metric	28.77	14	<.001	.994/0.992	.041/.024	2 vs. 1	.001	6.19	4	.186
3. scalar	68.45	19	<.001	.984/0.983	.059/.035	3 vs. 2	.010	62.28	5	<.001
4. residual	84.50	24	<.001	.979/0.982	.061/.040	4 vs. 2	.015	59.79	10	<.001
Over types of school on $t_1$ :										
1. configural	31.32	15	<.001	.993/0.985	.048/.015					
2. metric	53.21	23	<.001	.987/0.983	.051/.035	2 vs. 1	.006	22.34	8	<.01
3. scalar	113.52	33	<.001	.971/0.974	.064/.065	3 vs. 2	.016	89.17	10	<.001
4. residual	278.24	43	<.001	.842/0.890	.131/.098	4 vs. 2	.145	298.12	20	<.001
Over types of school on $t_2$ :										
1. configural	26.50	15	<.001	.994/0.988	.039/.017					
2. metric	34.44	23	<.001	.994/0.992	.031/.025	2 vs. 1	.000	7.88	8	.445
3. scalar	54.27	33	<.001	.990/0.991	.033/.032	3 vs. 2	.004	22.62	10	<.05
4. residual	117.66	43	<.001	.963/0.974	.056/.046	4 vs. 2	.031	87.79	20	<.001
Over types of school on $t_3$ :										
1. configural	25.08	15	<.001	.996/0.992	.040/.014					
2. metric	33.02	23	<.001	.997/0.995	.031/.022	2 vs. 1	-.001	6.84	8	.554
3. scalar	55.92	33	<.001	.993/0.994	.036/.030	3 vs. 2	.004	28.18	10	<.01
4. residual	80.01	43	<.001	.988/0.991	.042/.039	4 vs. 2	.109	47.90	20	<.001

Notes. N = 2902.  $\chi^2$  = Chi-Square, df = degrees of freedom. CFI = Comparative Fit Index. TLI = Tucker-Lewis index. RMSSEA = Root mean square error of approximation. SRMR = Standardized root mean square residual.  $p(\Delta \chi^2)$ -values are calculated based on the Satorra-Bentler scaled chi-square difference test (Satorra & Bentler, 1994).

## Appendix E

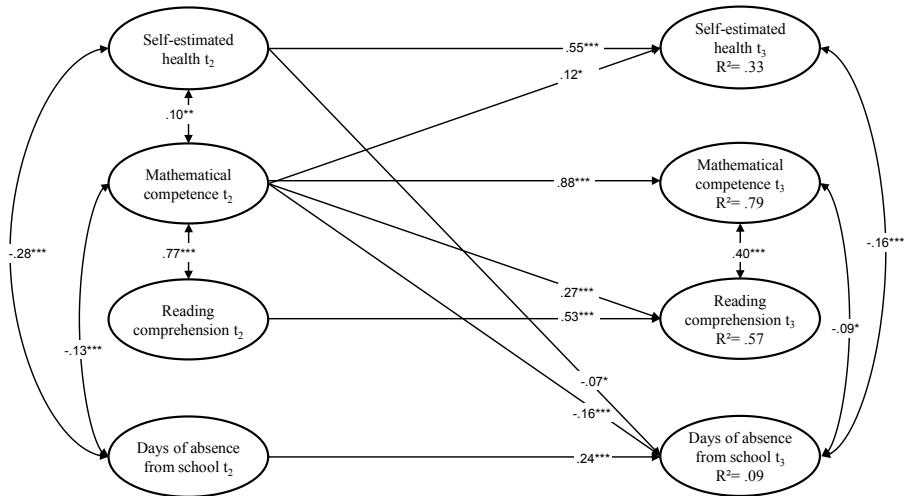
**Figure E1:** Reciprocal relations between physical well-being (measured by self-estimated health and days of absence from school) and academic achievement (MC and RC) from grade 7 to grade 9.



*Notes.* Sex = male. N = 2100. Fit measures are not reported because the model is saturated. All latent variables are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

**Figure E2:** Reciprocal relations between physical well-being (measured by self-estimated health and days of absence from school) and academic achievement (MC and RC) from grade 7 to grade 9.

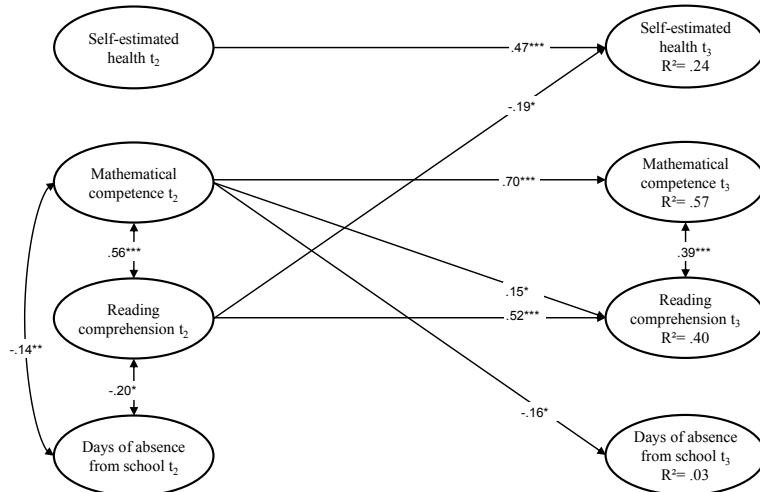


*Notes.* Sex = female. N = 2059. Fit measures are not reported because the model is saturated. All latent variables are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

## Appendix F

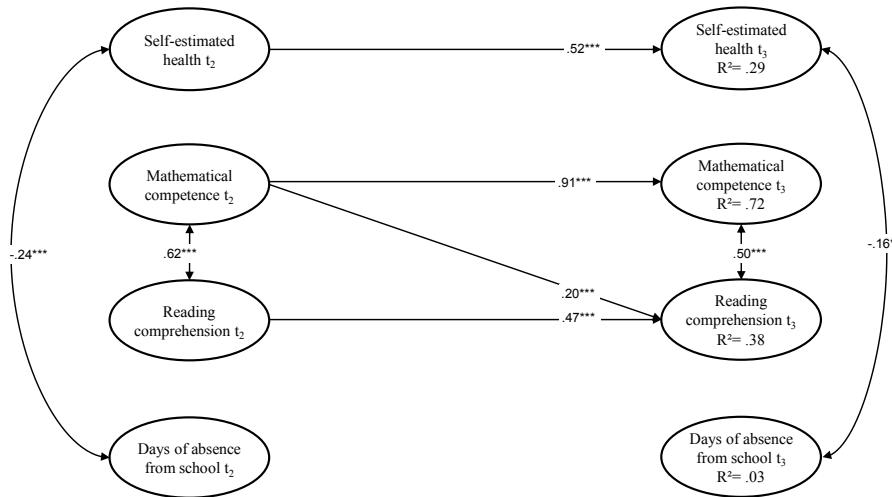
**Figure F1:** Reciprocal relations between physical well-being (measured by self-estimated health and days of absence from school) and academic achievement (MC and RC) from grade 7 to grade 9.



*Notes.* Type of School = “Hauptschule” (lowest track of German secondary school system).  $N = 577$ . Fit measures are not reported because the model is saturated. All latent variables are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

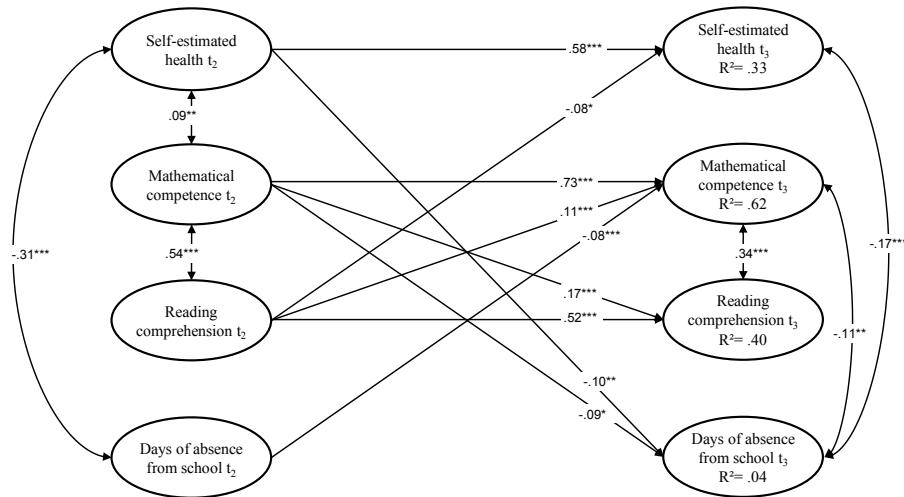
Figure F2: Reciprocal relations between physical well-being (measured by self-estimated health and days of absence from school) and academic achievement (MC and RC) from grade 7 to grade 9.



Notes. Type of School = “Realschule” (intermediate track of German secondary school system). N = 1162. Fit measures are not reported because the model is saturated. All latent variables are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

\*p < .05. \*\*p < .01. \*\*\*p < .001.

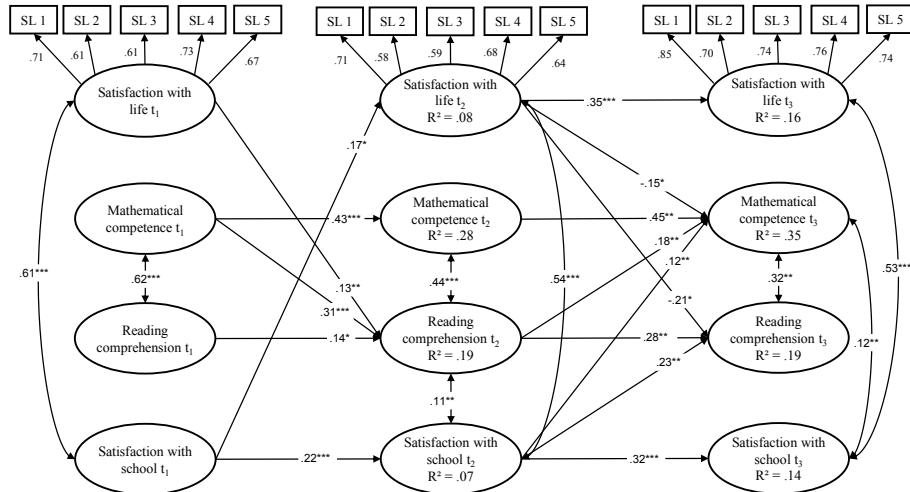
Figure F3: Reciprocal relations between physical well-being (measured by self-estimated health and days of absence from school) and academic achievement (MC and RC) from grade 7 to grade 9.



Notes. Type of School = "Gymnasium" (highest track of German secondary school system). N = 2420. Fit measures are not reported because the model is saturated. All latent variables are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.  
 \*p < .05. \*\*p < .01. \*\*\*p < .001.

## Appendix G

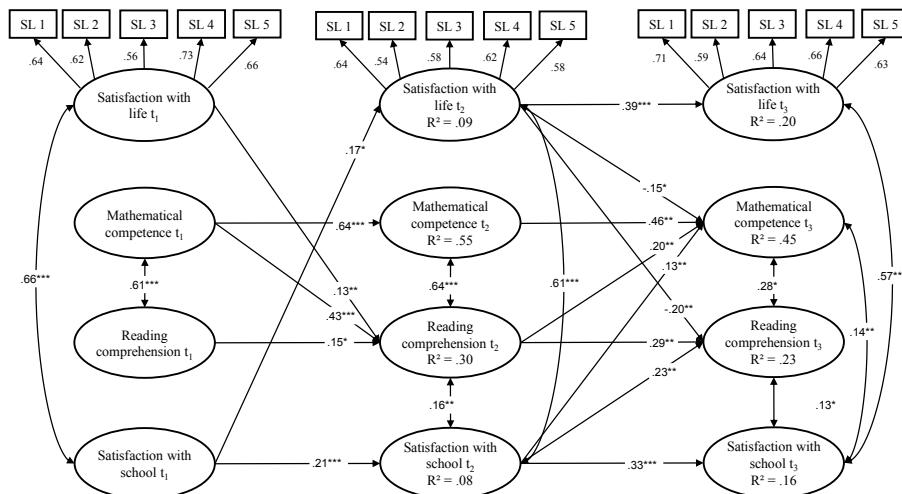
**Figure G1:** Reciprocal relations between cognitive well-being (measured by satisfaction with life and satisfaction with school) and academic achievement (MC and RC) from grade 5 to grade 9.



Notes. Sex = male. N = 1458.  $\chi^2(187) = 279.73$ , AIC = 117852.80, BIC = 118576.82, CFI = .991, TLI = 0.987, RMSEA = .022 SRMR = .024. Variables "Mathematical competence", "Reading comprehension" and "Satisfaction with school" are measured by single indicators which have not been included in the figure. Latent variables that model random intercepts have not been included in the figure for clarity purposes. Only statistically significant paths are shown.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

**Figure G2:** Reciprocal relations between cognitive well-being (measured by satisfaction with life and satisfaction with school) and academic achievement (MC and RC) from grade 5 to grade 9.

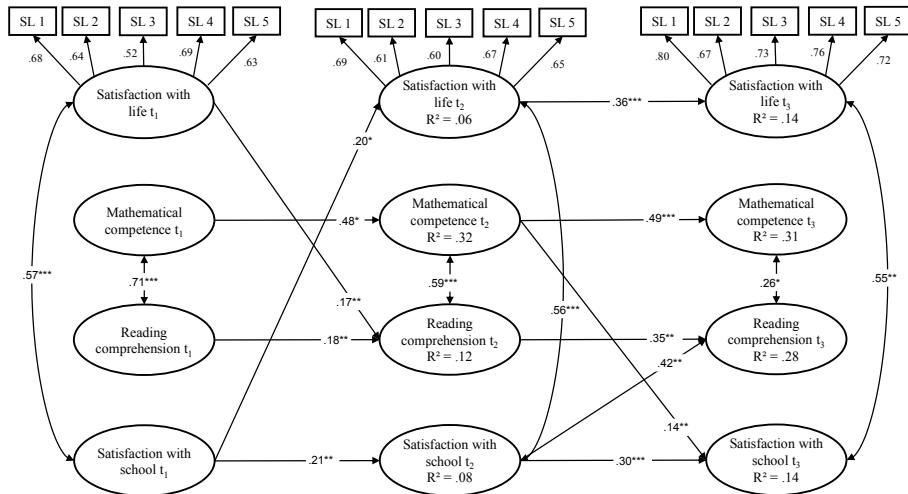


Notes. Sex = female. N = 1441.  $\chi^2(187) = 429.22$ , AIC = 113682.45, BIC = 114404.87, CFI = .977, TLI = 0.967, RMSEA = .035 SRMR = .035. Variables "Mathematical competence", "Reading comprehension" and "Satisfaction with school" are measured by single indicators which have not been included in the figure. Latent variables that model random intercepts have not been included in the figure for clarity purposes. Only statistically significant paths are shown.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

## Appendix H

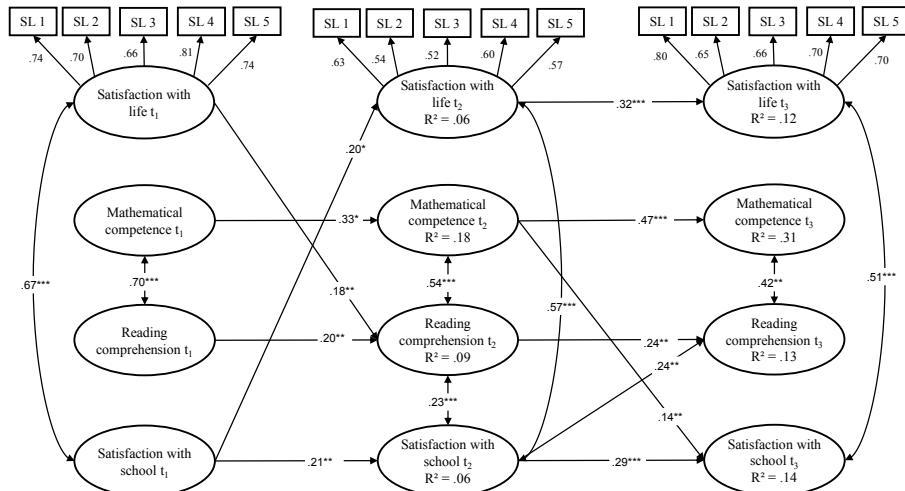
**Figure H1:** Reciprocal relations between cognitive well-being (measured by satisfaction with life and satisfaction with school) and academic achievement (MC and RC) from grade 5 to grade 9.



**Notes.** Type of school = “Hauptschule” (lowest track of German secondary school system). N = 363.  $\chi^2(187) = 235.65$ , AIC = 29320.07, BIC = 29853.60, CFI = .981, TLI = 0.971, RMSEA = .029. SRMR = .038. Variables “Mathematical competence”, “Reading comprehension” and “Satisfaction with school” are measured by single indicators which have not been included in the figure. Latent variables that model random intercepts have not been included in the figure for clarity purposes. Only statistically significant paths are shown.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

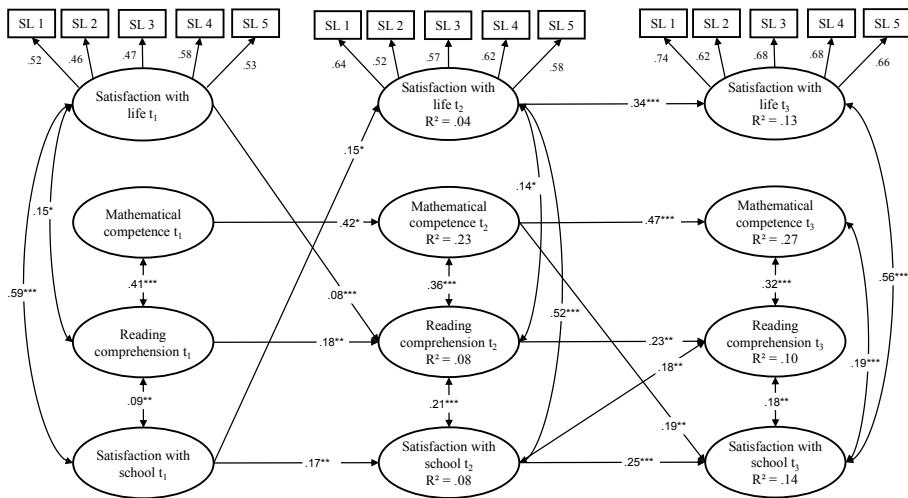
**Figure H2:** Reciprocal relations between cognitive well-being (measured by satisfaction with life and satisfaction with school) and academic achievement (MC and RC) from grade 5 to grade 9.



**Notes.** Type of school = “Realschule” (intermediate track of German secondary school system). N = 788.  $\chi^2(187) = 279.53$ , AIC = 63884.19, BIC = 64523.92, CFI = .984, TLI = 0.976, RMSEA = .028 SRMR = .032. Variables “Mathematical competence”, “Reading comprehension” and “Satisfaction with school” are measured by single indicators which have not been included in the figure. Latent variables that model random intercepts have not been included in the figure for clarity purposes. Only statistically significant paths are shown.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

**Figure H3:** Reciprocal relations between cognitive well-being (measured by satisfaction with life and satisfaction with school) and academic achievement (MC and RC) from grade 5 to grade 9.

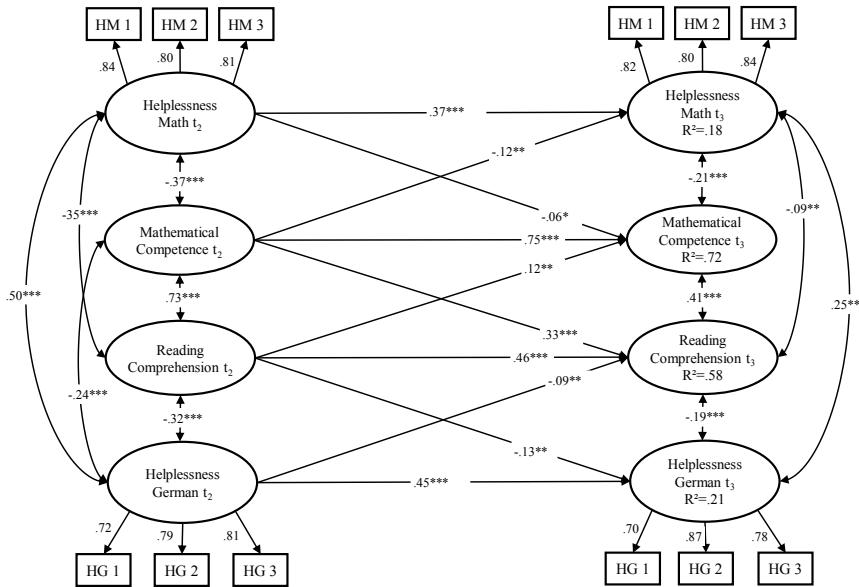


**Notes.** Type of school = “Gymnasium” (highest track of German secondary school system).  $N = 1748$ .  
 $\chi^2(187) = 382.40$ , AIC = 135958.52, BIC = 136707.40, CFI = .980, TLI = 0.972, RMSEA = .029  
SRMR = .030. Variables “Mathematical competence”, “Reading comprehension” and “Satisfaction with school” are measured by single indicators which have not been included in the figure. Latent variables that model random intercepts have not been included in the figure for clarity purposes. Only statistically significant paths are shown.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

## Appendix I

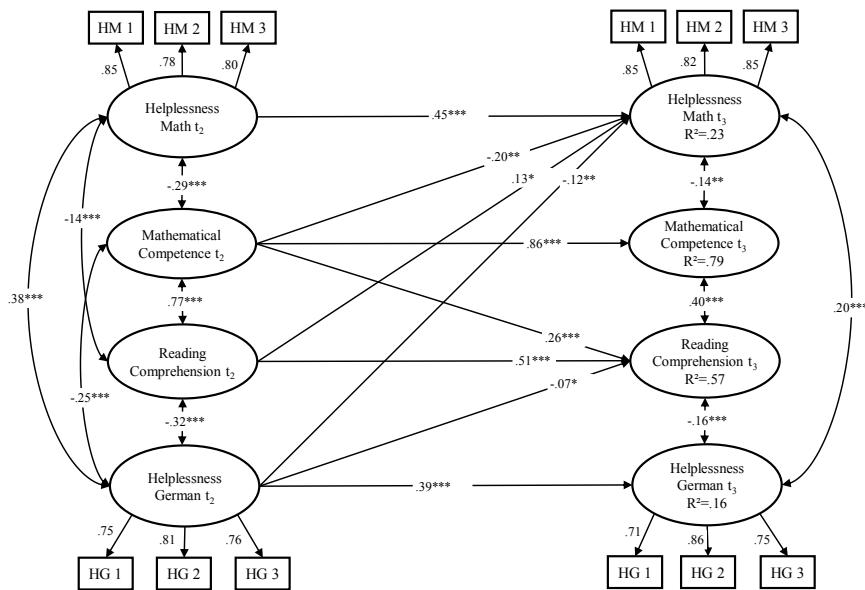
**Figure I1:** Reciprocal relations between emotional well-being (measured by helplessness German and helplessness Math, respectively) and academic achievement (MC and RC) from grade 7 to grade 9.



Notes. Sex = male. N = 2100.  $\chi^2(80) = 223.86$ , AIC = 66547.67, BIC = 66954.46, CFI = .988, TLI = 0.983, RMSEA = .032, SRMR = .026. Variables "Mathematical competence" and "Reading comprehension" are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

**Figure I2:** Reciprocal relations between emotional well-being (measured by helplessness German and helplessness Math, respectively) and academic achievement (MC and RC) from grade 7 to grade 9.

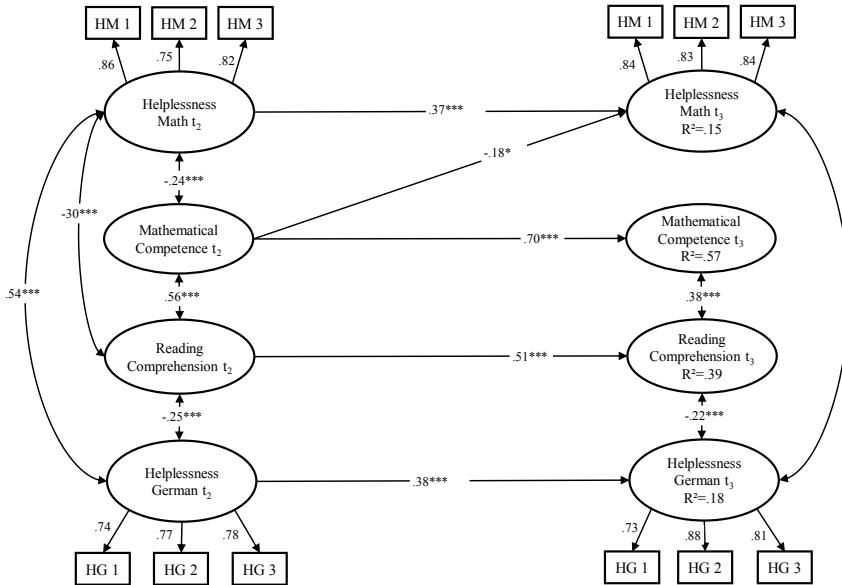


*Notes.* Sex = female.  $N = 2059$ .  $\chi^2(80) = 228.23$ , AIC = 63460.65, BIC = 63866.00, CFI = .988, TLI = 0.983, RMSEA = .031, SRMR = .022. Variables "Mathematical competence" and "Reading comprehension" are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

$p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

## Appendix J

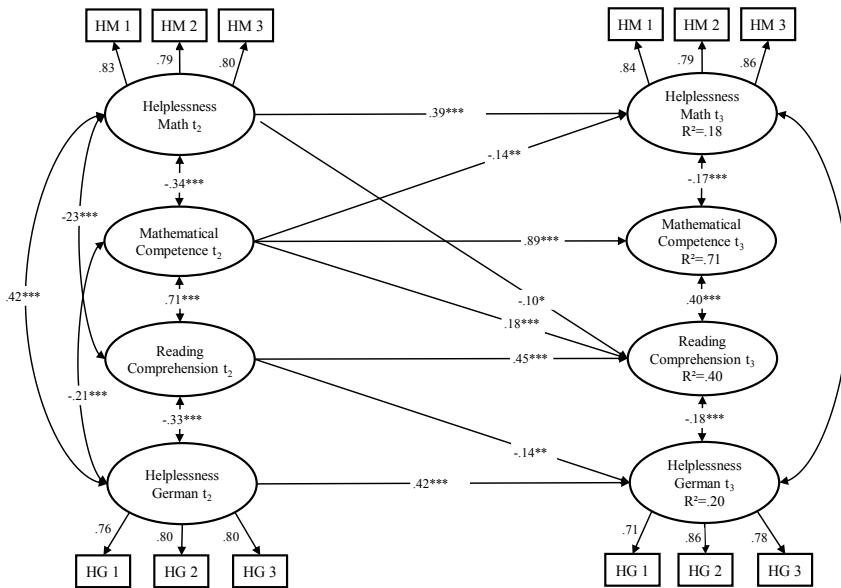
**Figure J1:** Reciprocal relations between emotional well-being (measured by helplessness German and helplessness Math, respectively) and academic achievement (MC and RC) from grade 7 to grade 9.



Notes. Type of school = “Hauptschule” (lowest track of German secondary school system).  $N = 577$ .  
 $\chi^2(80) = 104.68$ , AIC = 17665.57, BIC = 17979.33, CFI = .991, TLI = 0.987, RMSEA = .024, SRMR = .030. Variables “Mathematical competence” and “Reading comprehension” are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

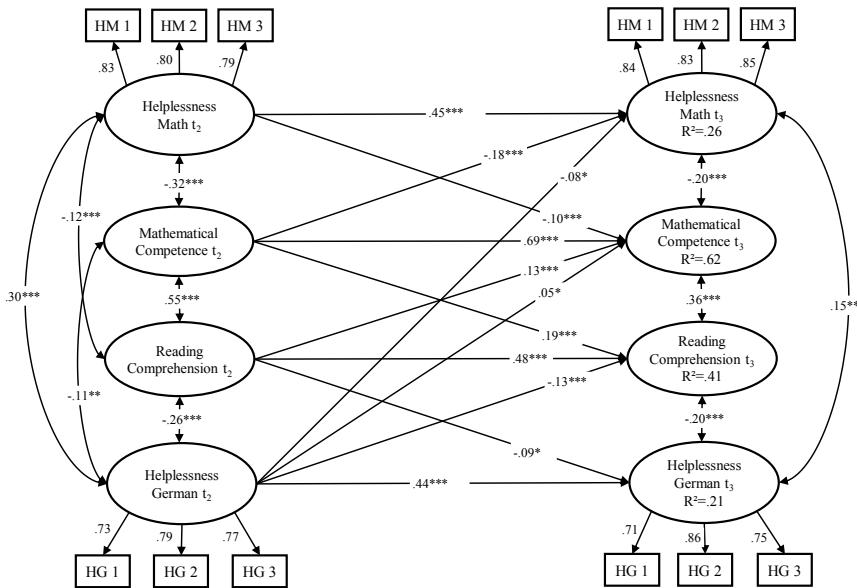
**Figure J2:** Reciprocal relations between emotional well-being (measured by helplessness German and helplessness Math, respectively) and academic achievement (MC and RC) from grade 7 to grade 9.



*Notes.* Type of school = “Realschule” (intermediate track of German secondary school system). N = 1162.  $\chi^2(80) = 159.17$ , AIC = 36065.55, BIC = 36429.72, CFI = .988, TLI = 0.981, RMSEA = .031, SRMR = .023. Variables “Mathematical competence” and “Reading comprehension” are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

\*p < .05. \*\*p < .01. \*\*\*p < .001.

Figure J3: Reciprocal relations between emotional well-being (measured by helplessness German and helplessness Math, respectively) and academic achievement (MC and RC) from grade 7 to grade 9.



Notes. Type of school = “Gymnasium” (highest track of German secondary school system). N = 2420.  $\chi^2(80) = 228.462$ , AIC = 74479.07, BIC = 63866.00, CFI = .988, TLI = 0.983, RMSEA = .029, SRMR = .022. Variables “Mathematical competence” and “Reading comprehension” are measured by single indicators which have not been included in the figure. Only statistically significant paths are shown.

\*p < .05. \*\*p < .01. \*\*\*p < .001.

Karin Gehrer & Lena Nusser

## Binnendifferenzierender Deutschunterricht und dessen Einfluss auf die Lesekompetenzentwicklung in der Sekundarstufe I

### Zusammenfassung

*Binnendifferenzierung im Unterricht ist in zunehmend heterogenen Klassen eine große didaktische Herausforderung. Doch obwohl Binnendifferenzierung für heterogene Gruppen schon seit den 70er-Jahren eine pädagogische Forderung ist, zeigte sich bisher in Deutschland immer noch ein Übergewicht des lehrergelenkten Klassenunterrichts. In Anbetracht knapper Zeitressourcen scheint eine Umsetzung in der Praxis erschwert. In diesem Beitrag widmen wir uns den Fragen, a) inwiefern Binnendifferenzierung im Deutschunterricht der Sekundarstufe aktuell umgesetzt wird und b) inwiefern das Lehrer\*innenhandeln im Deutschunterricht im Sinne binnendifferenzierender Unterrichtsmethoden einen Einfluss auf die Entwicklung der Lesekompetenz der Schüler\*innen zeigt und c) inwiefern die Heterogenität der Klassen einen Einfluss auf die Anwendung und Wirkung binnendifferenzierender Maßnahmen hat.*

*Anhand von Daten der Startkohorte 3 des Nationalen Bildungspanels wird die längsschnittliche Lesekompetenzentwicklung von Klasse 5 bis 9 ( $N_{\text{Klasse } 5} = 5283$ ) mit Latent-Change-Score-Modellen modelliert und der Einfluss des binnendifferenzierenden Lehrer\*innenhandelns ( $N_{LpK5} = 226$ ) im Deutschunterricht untersucht. Die Ergebnisse zeigen, dass a) die Binnendifferenzierung der Deutschlehrkräfte im Verlauf der Sekundarstufe abnimmt, b) das binnendifferenzierende Lehrer\*innenhandeln sich nur zu Beginn der Sekundarstufe positiv auf die Leistungsveränderung auswirkt und c) die Heterogenität der Klassen keinen Einfluss auf den Einsatz von Binnendifferenzierung hat.*

### Schlagworte

*Binnendifferenzierung; Leistungsstreuung; Lesekompetenz; Sekundarstufe; Längsschnitt*

---

Dr. Karin Gehrer (corresponding author) · Dr. Lena Nusser, Leibniz-Institut für Bildungsverläufe e.V. (LIfBi), Wilhelmsplatz 3, 96047 Bamberg, Deutschland  
E-Mail: karin.gehrer@lifbi.de  
lena.nusser@lifbi.de

# Adaptive instruction in German classes and its effect on reading competence development in secondary school

## Abstract

*Adaptive teaching in class is a didactical challenge in increasingly heterogeneous learning groups. However, although adaptive teaching has been a pedagogical demand for heterogeneous groups since the 1970s, there still has been a predominance of teacher-directed instruction in Germany. In the light of limited time resources, implementation in practice seems difficult. This paper examines a) the extent to which adaptive instruction is implemented at secondary level and b) the extent to which teacher's practices in German class has an influence on the development of reading competence and c) how heterogeneity regarding reading competence effects the implementation and the influence of adaptive instruction.*

*Using data from Starting Cohort 3 of the National Educational Panel Study, the longitudinal development of reading competence from Grades 5 to 9 ( $N_{grade\ 5} = 5283$ ) is modelled and the influence of adaptive teaching within German classes ( $N_{GT\_grade5} = 226$ ) is examined within the framework of latent-change-score-models. The results show that a) adaptive teaching on average declines in the course of lower secondary school, b) a positive impact of adaptive instruction is found at the beginning of secondary school and c) class heterogeneity does not influence the implementation of adaptive instruction.*

## Keywords

*Adaptive teaching; Performance variance; Reading competence; Secondary school; Longitudinal*

## 1. Einleitung

Die Heterogenität in deutschen Schulklassen nimmt zu, u.a. durch strukturelle Maßnahmen wie die Wandlung von einem drei- zu einem zweigliedrigen Schulsystem oder die steigende Inklusion von Schüler\*innen mit einem sonderpädagogischen Förderbedarf oder der Integration von Geflüchteten (Vock & Gronostaj, 2017). Damit einher gehen Herausforderungen hinsichtlich der Gestaltung des Unterrichts und der adäquaten Förderung aller Schüler\*innen. Als gute Lehrkräfte werden Personen beschrieben, die „in professioneller Weise verschiedene Varianten des Lehrens [beherrschen], um ihren Unterricht in mehrkriterialer Weise erfolgreich zu machen“ (Weinert & Helmke, 1996, S. 148). Eine dieser Varianten, die in Bezug auf heterogene Klassenkompositionen immer wieder genannt wird, ist die der Binnendifferenzierung (Messner, 2009). Die Anwendung adaptiver Unterrichtsformen, die Übungen mit verschiedenen Schwierigkeitsgraden enthalten, soll leistungsbezogener Heterogenität angemessen begegnen. Wie aber sind die Effekte der postulierten Binnendifferenzierung? Lohnt der erhöhte Aufwand

und bringt er bessere Schulleistungen der damit geförderten Schüler\*innen hervor? Ziel der vorliegenden Studie ist es, die Kompetenzentwicklung im Lesen im Verlauf der Sekundarstufe unter Berücksichtigung des binnendifferenzierenden Angebots der Deutschlehrkräfte zu untersuchen. Die Fokussierung auf die Lesekompetenz und damit den Deutschunterricht liegt darin begründet, dass Lesekompetenz nicht nur eine zentrale kulturelle Schlüsselqualifikation darstellt, die gesellschaftliche, soziale und politische Teilhabe ermöglicht (z.B. Artelt et al., 2001), sondern sie bildet auch innerhalb der Schule die Basis für eine gelingende Wissensaneignung in anderen Fächern. In der Sekundarstufe besteht die Aufgabe der Förderung der Lesekompetenz darin, den jugendlichen Schüler\*innen ein vertieftes Leseverständnis im Sinne einer effektiven Informationsverarbeitung zu ermöglichen, leseförderliche Gewohnheiten zu etablieren und Wege zu trainieren, um verschiedenartige Texte besser erschließen und interpretieren zu können (Artelt et al., 2001, S. 76–78, S. 131–134). Bereits die PISA-Lesenergebnisse 2000 zeigten, dass in der deutschen Sekundarstufe ein großer Bedarf zur Förderung der Lesekompetenz besteht.

## 2. Binnendifferenzierung

### 2.1 Was meint Binnendifferenzierung im Unterricht

Binnendifferenzierung als Lehrerhandeln ist die methodische Möglichkeit, verschiedene Aspekte des Lernprozesses innerhalb einer (heterogenen) Lerngruppe bzw. eines Klassenverbandes unterschiedlich zu gestalten (vgl. Bönsch, 2012; Eisenmann & Grimm, 2016). Wenngleich in Theorie und Empirie, national und international, die Termini nicht einheitlich gefasst sind (vgl. Bohl, Batzel & Richey, 2012; Schneider & Ludwig, 2012), sondern sich in dem Bereich unterschiedlichste Ausprägungen finden, gilt es, zumindest die grundsätzlichen Konzepte, wenn möglich zu unterscheiden. *Äußere* oder auch *institutionelle Differenzierung* (engl. *tracking*) meint eine Gliederung der Schülerschaft in möglichst homogene Leistungsgruppen: In unserem historischen Bildungswesen sind dies in einfacherster Form die Jahrgangsgruppen, sowie das gegliederte Schulsystem mit einer intendiert leistungsorientierten Schulformdifferenzierung (z.B. Hauptschule, Realschule, Gymnasium), aber auch die intraschulische Differenzierung zwischen Grundkursen neben Leistungskursen (Roeder, 1997; Vock & Gronostaj, 2017) oder räumlich getrennter Unterricht mit unterschiedlichen Lehrkräften (Saalfrank, 2012). Demgegenüber sprechen wir von *Binnendifferenzierung* oder *Innerer Differenzierung*, wenn innerhalb der heterogenen Klasse verschiedene Niveaus berücksichtigt werden, der Unterricht möglichst flexibel an die unterschiedlichen Lernbedürfnisse angepasst wird, wenn also in einer Klasse parallele Lernangebote bestehen, in denen die Schüler\*innen gemäß ihrem unterschiedlichen Lerntempo, ihren eigenen Interessen, kognitiven Fähigkeiten oder Leistungen

unterschiedliche Arten und Mengen von Lernmaterialien, Aufgabenstellungen und Zwischenlernzielen erhalten (Helmke, 2007; Vock & Gronostaj, 2017; vgl. für adaptive teaching: Corno & Snow, 1986; für adapted teaching: z.B. Schulz-Heidorf, 2016). Dabei kann von zwei Basisformen Innerer Differenzierung gesprochen werden: einerseits von der Differenzierung von Methoden und Medien innerhalb der Lerngruppe, andererseits wenn eine Differenzierung der Lernziele und Lerninhalte innerhalb einer Klasse vorgenommen wird (Klafki & Stöcker, 1986, S. 504). Letzteres ist auch als „goal-differentiated teaching“ bekannt (Schulz-Heidorf, 2016, S. 8). Die Unterrichtsdifferenzierung erfolgt je nachdem nach Lerntempo, Leistung, Umfang der Aufgaben, der Lern-Ziele oder auch thematisch nach Interesse durch das Anbieten von Wahlmöglichkeiten. Die Operationalisierung von Binnendifferenzierung in der Klasse geschieht oft über die Organistation von parallelen Kleingruppen, die für einen bestimmten Zeitraum zusammenarbeiten, es gibt aber auch Modelle mit unterschiedlichen Lernpfadangeboten (wahl differenzierter Unterricht), oder in der Annäherung an einen offenen Unterricht die Form von Lernstationen oder andere Methoden (vgl. z.B. Bohl, Batzel & Richey, 2012; Bönsch, 2012). Eingebaut in den Ablauf des Unterrichts über mehrere Wochen hinweg, könnte die Umsetzung von Binnendifferenzierung so aussehen, dass zuerst die Grundlernziele im heterogenen Klassenverband vermittelt werden, überprüft mit z.B. einem Test zur Ermittlung des Lernstandes und der Lernlücken. Darauf aufbauend wird der Klassenverband in drei Gruppen geteilt: Neben dem mittleren Niveau werden auf vertieftem Niveau die Schüler\*innen ohne Lücken gefördert. Mit den Schüler\*innen, welche die Grundlernziele nicht erreicht haben, werden diese Ziele in einem Wiederholungskurs nochmals, bestenfalls mit anderen Techniken, erarbeitet (Paradies & Linser, 2017, S. 12). Die Höchstform der inneren Differenzierung ist eine konsequente Individualisierung im Unterricht (Eberle, Kuch & Track, 2016; Schulz-Heidorf, 2016, 2018).

Binnendifferenzierung zielt darauf ab, dass der Unterricht passend zur Bandbreite der Lernenden zugeschnitten wird, er ist weder über- noch unterfordernd und setzt die passende Diagnose des jeweiligen Ist-Standes der Lernstände, Lernergebnisse und Lernprozesse voraus.

Durch Binnendifferenzierung wird eine positive Beeinflussung der Motivation, des Selbstkonzepts und der Schulleistung der Lernenden erwartet. Das Ziel ist hierbei eine gezielte Förderung von unterschiedlich leistungsfähigen Subgruppen innerhalb der Klasse, um allen geeignete Lernchancen zu bieten (Bönsch, 2011; Eberle et al., 2016; Schneider & Ludwig, 2012).

Ähnliche Ziele verfolgt auch das Konzept des *adaptiven Unterrichtens* bzw. der Adaptivität von Unterricht, welches von Helmke und Weinert (1997) als Prozessmerkmal für guten Unterricht als aussichtsreiches didaktisches Modell in Bezug auf interindividuelle Unterschiede bewertet wird. Der Terminus adaptiver Unterricht wird nicht selten wenig trennscharf oder auch synonym verwendet mit Binnendifferenzierung. Augenscheinlich wird Adaptivität eher von der Ergebnisperspektive her definiert, indem eine Leistungshomogenisierung oder ein

vergleichbarer Lernfortschritt durch eine optimale Passung der Lernumgebung und den individuellen Lernvoraussetzungen erreicht werden soll (Hertel, 2014).

Validiert ist die Binnendifferenzierung auch im Framework von Praetorius, Klieme, Herbert & Pinger (2018), im Rahmen der drei grundlegenden Dimensionen des guten Unterrichts (Klassenraummanagement, Unterstützung der Schüler\*innen und kognitive Aktivierung). In der Dimension *Student Support* findet sich die Subdimension *Differenzierung und adaptive Unterstützung*, die in 21 Publikationen mit unterschiedlichen Stichproben und verschiedenen Domänen (u.a. Mathematik, Naturwissenschaften, Lesen, Schreiben) operationalisiert wurde. Das differenzierende Lehrerhandeln wird hier übergreifend wie folgt beschrieben: Die Lehrperson bietet Übungen mit verschiedenen Schwierigkeitsgraden und Tempi sowie Zusatzaufgaben für Schnelle an und nimmt sich zusätzliche Zeit zum wiederholten Erklären bei Verständnischwierigkeiten einzelner Schüler\*innen oder Gruppen.

## **2.2 Welche Herausforderungen sind mit der Umsetzung von Binnendifferenzierung verbunden**

Die Beschreibung der Konzepte verdeutlicht implizit den Mehraufwand der Lehrkräfte, wenn sie binnendifferenzierende Maßnahmen in ihren Unterricht einbauen wollen. So lautet auch die Begründung der Lehrerschaft, wenn sie entsprechende Ansätze nicht umsetzt (Roeder, 1997). Daher ist es nicht verwunderlich, dass in Deutschlands Regelschulen binnendifferenzierender Unterricht eher selten angeboten wird (Vock & Gronostaj, 2017; vgl. auch Roeder, 1997; Wischer, 2007, 2008).

Vielmehr zeigt sich zu Beginn des 21. Jahrhunderts an Grundschulen immer noch ein Übergewicht des lehrergelenkten Klassenunterrichts, in dem Lehrpersonen häufig im Klassenverband Instruktionen geben, täglich das Lehrbuch, aber kein niveauangepasstes Material einsetzen und nur selten eine eigene Wahl von Büchern erlauben, also gemäß der Lehrerbefragung in der IGLU/PIRLS-Studie einen Unterrichtsstil ohne erkennbare Individualisierung oder Differenzierung ausüben (Lankes, 2004, S. 559–564). Dieser (traditionelle) Typ der Lehrperson kommt nur in Griechenland und Italien noch häufiger vor als in Deutschland. Der Typus des (lehrergelenkten) individualisierenden Gruppenunterrichts, der sich an den individuellen Leistungsvoraussetzungen der Schüler\*innen orientiert und mit niveauangepassten Materialien darauf antwortet, ist in Deutschland noch eher selten (knapp 10 % der Lehrpersonen). Auch ein von der Lehrkraft wenig gelenkter, sehr individualisierter Unterricht findet sich in Deutschland kaum (Lankes, 2004).

In der Sekundarstufe wird der traditionelle Lehrervortrag häufig von einer weiteren Sozialform mit starker Lehrerzentrierung, dem sog. Klassengespräch, einem eng geführten fragend-entwickelnden Gespräch im Klassenverband, abgelöst; andere Sozialformen wie Gruppenarbeit oder Einzel-Stillarbeit bleiben selten und

niveaudifferenzierte Aufgaben werden kaum vergeben (Blömeke & Müller, 2009; vgl. Prenzel et al., 2002).

Obwohl Binnendifferenzierung in heterogenen Gruppen schon in den 70er-Jahren eine pädagogische Forderung mit hoher Innovationserwartung war und heute wieder mit ähnlichen Reformansprüchen in die Schuldebatten aufgenommen wird (Trautmann & Wischer, 2009, 2011), stellt sie doch in Anbetracht von knappen Zeit- und anderen Ressourcen eine „Komplexitätsfalle“ dar (Wischer & Trautmann, 2010, S. 160), die eine Umsetzung in der Praxis behindert (Gruehn, 2000). Zudem ist bisher eher wenig binnendifferenzierendes Lehrmaterial auf dem deutschen Markt erhältlich (Vock & Gronostaj, 2017).

Binnendifferenzierung wird zudem als ein „Balanceproblem“ im Spannungsfeld zwischen dem Bezug auf die Gemeinschaft und den Ansprüchen der Einzelnen gesehen (Bräu, 2005). Den Grad der Differenzierung bezeichnet Saalfrank (2012) auch als den „Grad der Ungleichbehandlung“ (S. 71), was einen Widerspruch zum curricularen Bildungsanspruch deutlich macht. „Ein wahrhaft herkulisches pädagogisch-psychologisches Problem“ nannte Franz Weinert (1997, S. 50) diese Herausforderung insgesamt. Denn Individualisierung werde oft als Selbstverständlichkeit gefordert, ohne die gewünschten Lehr- und Unterrichtsmethoden zu konkretisieren, welche tatsächlich geeignet wären um praktikabel und lernwirksam die gewünschten Erfolge ohne Nebenwirkungen und Überforderung der Lehrkräfte zu erreichen. Auch Wischer (2008) betont die Notwendigkeit, die Anforderungen der Umsetzung in der Praxis aufgrund ungünstiger Rahmenbedingungen nicht zu unterschätzen, denn Binnendifferenzierung ist ein „hoch anspruchsvolles“ Konzept, welches durch die parallel ablaufenden Lernprozesse und -situationen und die damit verbundene „Vervielfältigung von Entscheidungsnotwendigkeiten die Komplexität des Handlungsfeldes Unterricht noch einmal deutlich erhöht“ (S. 717).

## 2.3 Wie kann Binnendifferenzierung im Unterricht gelingen

Für die erfolgreiche Durchführung von differenzierendem und individuell fördern- dem Unterricht werden neue spezifische Lehrerkompetenzen gefordert, wie professionelle Diversity-Kompetenz (Carle, 2005) und übergreifende metakognitive Kompetenzen zum Erfassen der Lernprozesse einer Klasse als auch der Individuen im Sinne von sogenannter „adaptive[r] Lehrkompetenz“, durch welche die vier professionellen Dimensionen Sachkompetenz, diagnostische, didaktische und Klassenführungskompetenz orchestriert werden (Beck et al., 2008, S. 47). Bei hoher adaptiver Lehrkompetenz handle es sich um ein ideales Konstrukt, das sich den Gegebenheiten von Schule, aber auch der jeweiligen Klassenstufe entsprechend flexibel anpassen und ausprägen müsse (S. 170–172). Diesen Erfordernissen sollten sich Lehramtsausbildungen zunehmend stellen (vgl. Trautmann & Wischer, 2011).

Weitere Voraussetzungen für die erfolgreiche Umsetzung einer Differenzierung im Klassenraum sind gute diagnostische bzw. auch förderdiagnostische<sup>1</sup> Kompetenzen der Lehrkräfte. Als „Diagnose-Dilemma“ bekannt wurde das Pänenomen bei PISA, als nur 11 % der schlechtesten Leser\*innen bei den PISA-Tests 2000 durch ihre Lehrkräfte erkannt wurden (Artelt et al., 2001). Im adaptiven und binnendifferenzierenden Unterrichten ist eine regelmäßige, lernbegleitende Diagnostik jedoch wesentlich, um den individuellen Lernverlauf und die „Diskrepanz zwischen Lernstand und Lernziel“ zu erfassen (Klieme & Warwas, 2011, S. 811) und basierend darauf, angemessene Lerninhalte und -materialien zur Verfügung zu stellen.

Aber auch schülerseitig braucht es für gelingenden binnendifferenzierenden Unterricht beispielweise eine Akzeptanz der angebotenen oder zugewiesenen Materialien sowie die Bereitschaft, sich in einem Kleingruppenunterricht zu engagieren und diszipliniert zu verhalten. Schülerrseitig fehle häufig eine wesentliche Voraussetzung für binnendifferenzierten Unterricht, stellt Roeder (1997) im Gespräch mit Gesamtschullehrer\*innen fest, nämlich „die Fähigkeit zum selbständigen Umgang mit Texten“ (S. 247).

All dies sind Bedingungsfaktoren, welche ein Gelingen des komplexen Konzeptes Binnendifferenzierung erschweren.

## 2.4 Effekte von Binnendifferenzierung

Mittels ländervergleichender Reviews und Metaanalysen versuchten mehrere Forscher\*innen die divergenten Resultate von Einzelstudien in einem größeren Zusammenhang zu betrachten. In der Zusammenfassung von drei Metaanalysen und der Reanalyse eines internationalen Datensets berichtet Scheerens (2000), dass Effekte von Schulfaktoren allgemein geringen Einfluss auf die sprachlichen oder mathematischen Leistungen der Schüler\*innen zeigten, währenddem Unterrichtsfaktoren, die in überwiegend qualitativen oder experimentellen Studien erfasst wurden, durchschnittliche bis große Effekte erreichten. Für *Differenzierung und adaptiven Unterricht* wurde jedoch nur eine Effektstärke von  $d = 0.22$  dokumentiert; vergleichend dazu fallen die Effekte der Lernmöglichkeiten ( $d = 0.15$ ) sowie der Mitwirkung der Eltern ( $d = 0.08$ ) noch kleiner aus (S. 57).

In einer weiteren anglo-amerikanischen Metaanalyse (Lou et al., 1996) mit 51 Studien zu leistungsdifferenzierten Subgruppen versus nicht-differenziertem Klassenunterricht wird von einem sehr kleinen Leistungsvorteil für klasseninterne Leistungsdifferenzierung ( $d = 0.17$ ) berichtet. Unwesentlich größere Effekte im Hinblick auf die Lernleistung resultieren dabei, wenn die parallelen Gruppen innerhalb der Klassen möglichst klein gehalten werden (3-4 Schüler\*innen;  $d = 0.22$ ). In weiteren 12 Studien untersuchten Lou und Kolleg\*innen (1996)

<sup>1</sup> Förderdiagnostische Kompetenzen als elementare Lehrerkompetenzen meinen nach Hanke (2005) sowohl das Erfassen von individuellen Lernvoraussetzungen und -prozessen (Diagnose) als auch das darauf basierende Fördern von Lernprozessen durch die Gestaltung individuell angemessener Lernanregungen und deren Evaluation.

Binnendifferenzierungsmaßnahmen daraufhin, ob sich leistungshomogene Subgruppen mit niveauangepassten Materialien oder heterogene Gruppen mit unterstützendem Tutorensystem besser bewähren; sie berichten eine leichte Überlegenheit der homogenen Gruppendifferenzierung ( $d = 0.12$ ) (vgl. Schneider & Ludwig, 2012).

Für einen materialdifferenzierenden Unterricht (*different instruction*) zeigt sich ein Leistungsvorteil mit der Effektstärke von  $d = 0.25$  (Lou et al., 1996, S. 440–442). Für Leseförderung im Gruppenunterricht wurde ein kleinerer lernförderlicher Effekt ( $d = 0.13$ ) als für Mathematik und Naturwissenschaften ( $d = 0.20$ ) berichtet. In der Unterscheidung zwischen homogenen versus heterogenen Leistungs(klein)-gruppen ist der Effekt für Lesen jedoch höher ( $d = 0.36$ ) als für die andern beiden Fächer und zeigt sich hier ein Vorteil von leistungshomogenen Gruppen (Lou et al., 1996, S. 445–446).

Die Erfolge von unterschiedlichen Leseförderungsprogrammen für Lese-schwäche in der Mittel- und Oberstufe in USA überprüften Slavin, Cheung, Groff und Lake (2008) in einer Metaanalyse von 33 Studien. Insbesondere die mixed-method Programme (READ 180, mit 8 Studien; Voyager Passport, mit 1 Studie) enthielten Module individualisierten Unterrichts (wie nivellierte Materialien, Zusatzbearbeitungszeit, Extrahilfe) sowie die systematische Förderung für Kleingruppen innerhalb der Klasse. Die mittlere Effektgröße für diese Programme mit sehr großen Stichproben im Vergleich zu Kontrollgruppen ohne Förderung betrug  $d = 0.23$ .

Die umfangreiche Synthese von Hattie (2009, 2018), mit einer Liste von rund 250 Faktoren, die auf Schülerleistung eine Auswirkung haben, fasst Ergebnisse aus 1500-Meta-Analysen von 90000 Studien mit 300 Mio. Schüler\*innen zusammen. Für Binnendifferenzierung *im Sinne der Schaffung leistungshomogener Gruppen (within-class grouping)* werden 129 Studien und 2 Meta-Analysen mit ca. 16000 Personen einbezogen. Hier wird im Mittel eine sehr kleine Effektstärke von  $d = 0.16$  berichtet, die gegenüber anderen Unterrichtsfaktoren vernachlässigbar zu sein scheint; wenn jedoch berücksichtigt wird, dass der Effekt des Gruppierens von der Klassengröße abhängig ist, so findet sich für die *Differenzierung i.S.v. leistungshomogenen Kleingruppen innerhalb großer Klassen* (über 35 Schüler\*innen) sogar eine nach Hatties Daumenregeln beinahe mittlere Effektstärke von  $d = 0.35$  (Hattie, 2009, S. 94). Für individualisierten Unterricht, bei dem jedem Kind *flexible Lehrmethoden* und Motivationsstrategien angeboten werden, berichtet Hattie eine overall Effektstärke von  $d = 0.23$  (Hattie, 2009, S. 199). Mit *Enrichment*, womit Zusatzangebote für Hochleistende gemeint sind, also ebenfalls eine differenzierende Maßnahme, wird ein Effekt von  $d = 0.39$  erreicht (Hattie, 2009, S. 101; vgl. Köller, 2014).

Es zeigt sich somit – ähnlich wie bereits Helmke und Weinert (1997, S. 92–93) konstatierten – länderübergreifend eine große Breite von empirischen Befunden mit signifikanten, aber eher kleinen Effekten von Individualisierung und Differenzierung im Unterricht und insgesamt nur wenigen Befunden für das Leseverständnis. International wird für Lesen in der Sekundarstufe lediglich eine

(!) Studie im Längsschnitt ausgewiesen, welche aber nicht auf Klassenebene, sondern auf der Schulebene aggregiert ist (Scharenberg, 2012).

## 2.5 Situation in Deutschland

Vor allem für den Leseunterricht zeigt sich im europäischen Vergleich, dass binnendifferenzierender Gruppenunterricht bzw. hoch individualisierender Leseunterricht bisher kaum realisiert wird, und dies trotz der Empfehlung individueller Förderung in vielen deutschen Leselehrplänen (Lankes, 2004). Dies betrifft insbesondere Deutschlands Grundschulen; im Fach Deutsch hat sich die Notwendigkeit der systematischen Leseförderung jedoch von den Grundschulen in die Sekundarstufen fortgepflanzt (Rosebrock & Nix, 2011). Die Fachdidaktik stellt hierzu unterschiedliche Lesetrainings und binnendifferenzierende Instrumente zur Verfügung (z.B. Paradies & Linser, 2017). Nicht zuletzt zeigen die aktuellen PISA-Ergebnisse, dass auch am Ende der Pflichtschulzeit etwa ein Fünftel der Schüler\*innen ein gering ausgeprägtes Leseverständnis haben; insbesondere sind leseschwache Schüler\*innen mit steigendem Anteil an nicht-gymnasialen Schulformen zu finden (Becker-Mrotzek et al., 2019).

Empirische Befunde in Deutschland zu den Auswirkungen von Binnendifferenzierung in leistungsheterogenen Klassen liegen wenige vor. Zu angewandten Formen der inneren Differenzierung bzw. zu adaptivem Unterricht berichten Weinert und Helmke (1996) sowie Schrader und Helmke (2008; Helmke, 2017) in Grundschulen am häufigsten mit Vorteilen für leistungsschwächere und langsame re Schüler\*innen. In der Sekundarstufe hingegen beziehen sich die bisher berichteten häufigsten Differenzierungsmaßnahmen auf leistungsstarke Schüler\*innen, und zwar am häufigsten an Realschulen (Helmke, 2017). Beziiglich der Verwendung von binnendifferenzierenden Maßnahmen wurde auch in Deutschstunden gefunden, dass binnendifferenzierende Aufgaben zwar selten, jedoch insbesondere von denjenigen Lehrpersonen eingesetzt werden, die Aufgaben mit hohem kognitiven Anregungsgrad vergeben (Blömeke & Müller, 2009, für 11. und 12. Klassen).

## 2.6 Ziel der Studie

Die bisherige Forschungslage zu binnendifferenzierenden Maßnahmen im Unterricht ist different, für Deutschland eher gering ausgeprägt, insbesondere für den Deutschunterricht. Daher möchte die vorliegende Studie den Fragen nachgehen, wie es um die Binnendifferenzierung in deutschen Schulen bestellt ist und inwieweit das binnendifferenzierende Handeln der Lehrkräfte einen Effekt auf die Lesekompetenzentwicklung der Schüler\*innen hat. Auch wenn in der Sekundarstufe keine systematische Leseförderung betrieben wird, so ist doch der Umgang mit Texten, ihr vertieftes Lesen und darauf beruhendes Schlussfolgern, Reflektieren und Bewerten ein zentraler Bestandteil des Deutschunterrichts. Je

besser der Inhalt und das Anforderungsniveau der bearbeiteten Texte und sich darauf beziehendes Aufgabenmaterial auf das Leistungsniveau der Schüler\*innen abgestimmt sind, desto besser können theoretisch die Einzelnen von dem zu ihnen passenden Lernangebot profitieren. Somit wird gemäß der Theorie angenommen, dass der Mehraufwand binnendifferenzierender Maßnahmen eine positive Wirkung auf die Lesekompetenzentwicklung von Schüler\*innen hat. Hierbei wird jedoch entsprechend der vorliegenden, auch fachübergreifenden Empirie nicht von einem besonders starken und langfristigen Effekt ausgegangen, sondern es werden unter Berücksichtigung verschiedener Einflüsse und wechselnder Lehrkräfte im Verlauf der Sekundarstufe eher kleine und kurzfristig wirkende Effekte erwartet. Darüber hinaus stellt sich die Frage, inwieweit die Klassenkomposition im Sinne leistungsbezogener Heterogenität einen Einfluss auf die Kompetenzentwicklung hat. Ob binnendifferenzierende Maßnahmen vor allem in heterogeneren Klassen auftreten und eine positive Wirkung zeigen, wird ebenfalls untersucht. Gemäß Theorie und allgemeiner Didaktik sollte das Angebot von Binnendifferenzierung in besonders heterogenen Klassen verstärkt erfolgen.

### 3. Methode

#### 3.1 Stichprobe

Für die folgenden Analysen wurden Daten der Startkohorte 3 (Klasse 5) des Nationalen Bildungspanels (NEPS; Blossfeld, Roßbach & von Maurice, 2011) verwendet. Eine deutschlandweite, repräsentative Stichprobe von Sekundarschüler\*innen wurde jährlich zu bildungsrelevanten Aspekten befragt und ihre schulischen Kompetenzen getestet. Die Erhebungen fanden jährlich jeweils im ersten Schulhalbjahr statt. Zu Beginn der Sekundarstufe enthielt die Stichprobe 5283 teilnahmebereite Schüler\*innen ( $M_{\text{Alter}} = 10.92$ ;  $SD = 0.52$ ). In Klasse 7 erfolgte eine Aufstockung um 2205 Schüler\*innen. Das Geschlechterverhältnis ist nahezu ausgewogen (weiblich: 48.5 %). Gut 43 % der Stichprobe besucht ein Gymnasium. Neben der jährlichen Befragung und Testung der Schüler\*innen wurden auch verschiedene Kontextpersonen befragt; hierzu gehörten auch die Klassen- und Fachlehrkräfte (Frahm et al., 2011). Die Stichprobe ist in der 5. Klasse in 469 Deutschklassen genestet bzw. in 884 Deutschklassen in der 9. Klasse (siehe Tabelle 1). Die große Mehrzahl der Schüler\*innen haben an mindestens zwei (benachbarten) Messzeitpunkten teilgenommen (70 %), gut 40 % der Stichprobe weisen Kompetenzwerte zu allen drei Messzeitpunkten auf.

**Tabelle 1:** Stichprobe

	Klasse 5	Klasse 6	Klasse 7	Klasse 8	Klasse 9
Schüler*innen (mit Lesetestteilnahme)	5208	-	6211	-	4416
Deutschklassen	469	469	648	702	884
Teilnehmende Deutschlehrkräfte	226	241	319	375	311

### 3.2 Instrumente

**Lesekompetenz:** Die Lesekompetenz wurde in den Klassen 5, 7 und 9 mit eigen entwickelten Tests (Gehrer, Zimmermann, Artelt & Weinert, 2012; 2013), mit welchen das Textverständnis unterschiedlicher kontinuierlicher Textsorten erfasst wird, standardisiert und längsschnittlich kohärent erhoben (zur Testgüte für K5: Pohl, Haberkorn, Hardt & Wiegand, 2012; für K7: Krannich et al., 2017; für K9: Scharl, Fischer, Gnambs & Rohm, 2017). Pro Messzeitpunkt bearbeitete jede\*r Schüler\*in innerhalb von 28 Minuten ein Testheft mit fünf Texten (kommentierend, literarisch, Sach- und Werbetext, Anleitung; Gehrer & Artelt, 2013) zu denen Verständnisfragen mit unterschiedlichen kognitiven Anforderungen (Informationen entnehmen, textbezogenes Schlussfolgern, Reflektieren und Bewerten) gestellt werden. Die Antwortformate sind (a) klassische Einfachwahlaufgaben (MC; eine Antwort aus vier Optionen ist als richtig anzukreuzen), (b) mehrzeilige Richtig-Falsch-Aufgaben (CMC; für jede Aussage ist zu entscheiden, ob sie stimmt oder nicht) und (c) Zuordnungsaufgaben (ZO; für jeden Textabschnitt wird ein passender Zwischenstitel ausgewählt).

Die NEPS-Lesentests decken ein breites Fähigkeitsspektrum ab. Zum ersten Messzeitpunkt in der 5. Klasse beinhaltet der Test (insgesamt 33 Items) auch viele einfache Items (Pohl et al., 2012). Ab dem zweiten Messzeitpunkt wurde basierend auf dem Kompetenzwert der Vorwelle ein leichtes und ein schweres Testheft zugewiesen (Krannich et al., 2017). Daher kann davon ausgegangen werden, dass die Tests die Kompetenzniveaus – vor allem auch im unteren Bereich – hinreichend differenzieren.

Die Tests werden im Rahmen der Item-Response-Theorie (IRT) mit Partial-Credit-Gewichtung für die komplexen Items (CMC, ZO) ausgewertet; die Personenfähigkeiten werden als Weighted-Likelihood-Personenparameterschätzer (WLE; Warm, 1989) angegeben (Pohl & Carstensen, 2013). Für den Längsschnitt sind die Tests über parallele Linkstudien verknüpft und auf der selben Metrik abgetragen (Fischer, Rohm, Gnambs & Carstensen, 2016).

**Binnendifferenzierung:** Die Fragen zur inneren Differenzierung (adaptiert nach Ditton, 2000) wurden von den Deutschlehrpersonen im jährlichen Lehrerfragebogen in den Klassenstufen 5 bis 9 beantwortet.

Die Skala der Binnendifferenzierung enthält ursprünglich sieben Items, die in einer 5-Punkte-Likert-Skala zu beantworten sind ( $1 = \text{trifft gar nicht zu}$  bis  $5 = \text{trifft völlig zu}$ ).

Die Items umfassen Aussagen zum eigenen Lehrerhandeln, beispielsweise: „Leistungsstarken Schülerinnen und Schülern gebe ich Extraaufgaben, durch die sie wirklich gefordert werden“ (vgl. auch Schulz-Heidorf, 2018). Aufgrund variiierender Teilnahmequoten der Lehrkräfte liegen nicht für alle Schüler\*innen umfangliche, jährliche Informationen über ihren jeweiligen Deutschunterricht vor.

*Leistungsbezogene Heterogenität:* Je größer die Streuung innerhalb einer Klasse, desto eher sind binnendifferenzierende Maßnahmen angezeigt. Operationalisiert wurde leistungsbezogene Heterogenität für jeden Lesekompetenzmesszeitpunkt durch die Distanz zwischen den Schüler\*innen mit dem höchsten und dem niedrigsten Lesekompetenzwert (WLE) in jeder Deutschklasse.

*Bildungshintergrund:* Als Kontrollvariable wurde der familiäre Bildungshintergrund aufgenommen. Dieser wurde bei den Eltern über den höchsten erreichten allgemeinbildenden Schulabschluss des befragten Elternteils bzw. dessen Partner\*in erfasst. Die Bildungsabschlüsse werden in der ISCED-Kodierung (International Standard Classification of Education) zur Verfügung gestellt (Pelz & Zielonka, 2017).

*Sprachhintergrund:* Weiterhin wurde für den Sprachhintergrund der Schüler\*innen kontrolliert. Während der ersten Teilnahme wurden die Schüler\*innen gefragt, welche Sprache sie als Kind in ihrer Familie gelernt haben. Nur ein sehr geringer Prozentsatz (0.06 %) gibt an, dass sie ausschließlich eine nicht deutsche Sprache als Muttersprache haben. Knapp 22 % der Schüler\*innen berichten, dass sie neben Deutsch auch mindestens eine weitere Sprache als Kind in ihrer Familie lernten.

## 4. Analysen

### 4.1 Prüfung der Faktorstruktur

Querschnittliche konfirmatorische Faktorenanalysen für alle fünf Wellen bestätigten keine Eindimensionalität des Konstruktions Binnendifferenzierung. Ein rekodiertes Item sowie zwei weitere Items mit ungenügenden Faktorladungen ( $< 0.5$ ) wurden ausgeschlossen. Mit vier Items konnte die eindimensionale Faktorstruktur der Skala für jeden der fünf Messzeitpunkte bestätigt werden.

Tabelle 2: Deskriptive Ergebnisse der Skala Binnendifferenzierung (auf Lehrkraftebene)

Binnendifferenzierung	Klasse 5	Klasse 6	Klasse 7	Klasse 8	Klasse 9
a) Von Schülerinnen und Schülern mit geringer Leistungsfähigkeit verlange ich deutlich weniger.	2.91 (1.00)	2.93 (1.00)	2.85 (1.01)	2.84 (0.98)	2.72 (0.94)
b) Ich bilde Gruppen von Schülerinnen und Schülern mit ähnlichen Fähigkeiten. (r)	2.56 (1.02)	2.52 (0.90)	2.54 (0.95)	2.53 (0.93)	2.24 (0.88)
c) Ich bilde Gruppen von Schülerinnen und Schülern mit unterschiedlichen Fähigkeiten.	2.45 (0.91)	2.39 (0.80)	2.45 (0.88)	2.40 (0.83)	2.27 (0.88)
<b>d) Ich gebe Schülerinnen und Schülern je nach Leistungsstärke unterschiedlich schwierige Hausaufgaben.</b>	2.58 (1.20)	2.55 (1.20)	2.48 (1.18)	2.38 (1.20)	2.21 (1.09)
<b>e) Ich lasse schnellere Schülerinnen und Schüler schon zum Nächsten übergehen, während ich mit den langsameren noch übe oder wiederhole.</b>	3.76 (1.00)	3.71 (1.05)	3.76 (0.94)	3.64 (1.03)	3.57 (1.02)
<b>f) Wenn Schülerinnen und Schüler etwas nicht verstanden haben, vergabe ich gezielte Zusatzaufgaben.</b>	3.39 (1.07)	3.51 (0.98)	3.38 (1.01)	3.33 (0.97)	3.26 (0.99)
<b>g) Leistungsstarken Schülerinnen und Schülern gebe ich Extraaufgaben, durch die sie wirklich gefordert werden.</b>	3.67 (1.01)	3.56 (1.04)	3.50 (0.99)	3.47 (1.00)	3.31 (1.03)
Skalenmittelwert (basierend auf den Items d-g)	3.35 (0.80)	3.33 (0.79)	3.29 (0.75)	3.21 (0.78)	3.09 (0.72)

Anmerkungen. Angegeben sind jeweils Mittelwert und Standardabweichung (in Klammern). Fett: Items, die für die weiteren Analysen genutzt wurden. (r) = rekodiert.

## 4.2 Prüfung der längsschnittlichen Messinvarianz

In mehreren Schritten wurde die Messinvarianz des Konstruktions Binnendifferenzierung über die fünf Messzeitpunkte geprüft (zum Vorgehen vgl. Nusser, Carstensen & Artelt, 2015; Schroeders & Gnambs, 2018; Schwab & Helm, 2015). Insgesamt variiert die Teilnahmequote der Lehrkräfte über die Wellen stark, so dass zu den verschiedenen Messzeitpunkten unterschiedlich vollständige Daten vorliegen. Zudem ist davon auszugehen, dass die Lehrpersonen im Verlauf der Sekundarschulzeit wechseln, so dass die jährlichen Befragungen zunächst als querschnittliche Messungen vergleichend geprüft wurden. In einem Mehrgruppenvergleich konnte über alle Wellen hinweg die konfigurale Messinvarianz auf Grund guter Fitwerte verifiziert werden, sowie auch die metrische Invarianz mit frei geschätzter Faktorladung. Die skalare Messinvarianz (gleiche Intercepts über alle Messzeitpunkte) konnte nicht uneingeschränkt bestätigt werden. Die freie Schätzung der Intercepts dreier Items zu einzelnen Messpunkten

bestätigte partielle skalare Messinvarianz mit einem guten Modellfit ( $\chi^2 = 64.19$ ; df = 39, CFI = .974; TLI = .980; RMSEA = .044; SRMR = .065; siehe Tabelle 3).

**Tabelle 3:** Messinvarianzprüfung des Konstruktions Binnendifferenzierung (fünf Befragungszeitpunkte von Deutschlehrer\*innen)

Modell	$\chi^2$	df	$\chi^2/df$	CFI	TLI	RMSEA	SRMR	$\Delta CFI$	$\Delta RMSEA$
Konfigurale Invarianz	19.162	10	1.91	.990	.971	.052	.021	-	-
Metrische Invarianz	44.976	26	1.73	.980	.977	.047	.060	.010	-.005
Skalare Invarianz	83.130	42	1.98	.957	.969	.054	.074	.023	.007
Partielle skalare Invarianz	64.190	39	1.65	.974	.980	.044	.065	.006	.003

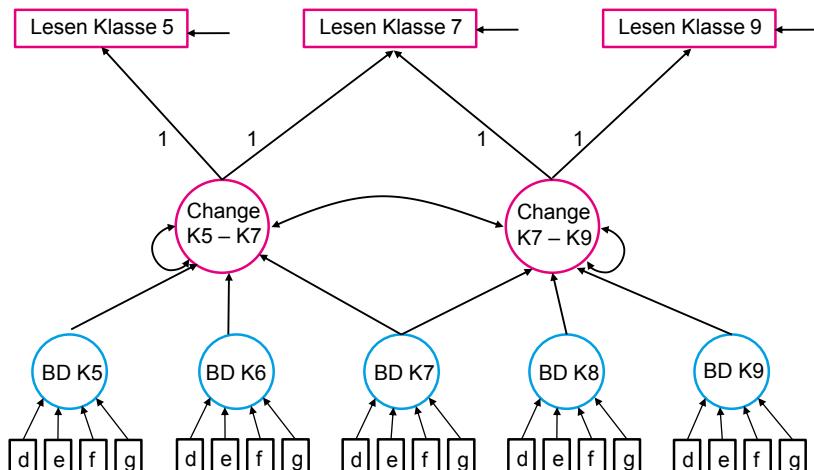
Anmerkungen. df = Freiheitsgrade. CFI = Comparative Fit Index. TLI = Tucker Lewis Index. RMSEA = Root Mean Square Error of Approximation. SRMR = Standardized Root Mean Square Residual.

### 4.3 Latent-Change-Score-Modelle

Zu unterschiedlichen Zeitpunkten wirken unterschiedliche Lehrkräfte auf die Schüler\*innen ein. Ihr Handeln bestimmt für eine gewisse Zeit, ein Schuljahr oder mehr, den Deutschunterricht. Es wird erwartet, dass der konkrete Unterricht eher kurzfristige, direkte Effekte auf die Kompetenzentwicklung der Schüler\*innen hat, jedoch nicht anzunehmen ist, dass eine spezifische Lehrkraft die Lesekompetenzentwicklung über den gesamten Verlauf der Sekundarstufe beeinflusst. Daher wird kein Wachstumskurvenmodell genutzt, sondern ein Latent-Change-Score-Modell, auch True-Individual-Change-Modell, herangezogen, welches jeweils die Veränderung der Lesekompetenz von einem zum nächsten Messzeitpunkt latent modelliert (McArdle, 2009; Steyer, 2009). Dieser Ansatz erlaubt es, die intraindividuelle Entwicklung zwischen zwei Messzeitpunkten regressionsbasiert zu untersuchen und das jeweilige Lehrkrafthandeln im Unterricht möglichst zeitnah mit der Veränderung der Lesekompetenz der Schüler\*innen in Verbindung zu bringen. Somit werden im Modell zwei latente Change-Scores für die Lesekompetenzentwicklung modelliert: Klasse 5 zu Klasse 7 sowie Klasse 7 zu Klasse 9. Die in den jeweiligen (Zwischen-)Jahren erhobenen Maße zur Binnendifferenzierung im Unterricht wurden als latente Prädiktoren für die Veränderung der Lesekompetenz aufgenommen (siehe Abbildung 1; Modell 1). Zusätzlich wurden relevante Kontrollvariablen (Geschlecht, Schulform, Muttersprache, Bildungshintergrund der Eltern) im Modell berücksichtigt (Modell 2). Das Heterogenitätsmaß der Leistungsstreuung innerhalb der Deutschklasse wurde in einem weiteren Schritt in die Regression (Modell 3) aufgenommen. Um die Frage zu untersuchen, ob die Leistungsstreuung innerhalb der Deutschklasse die Wirkung des binnendifferenzierenden Angebots der Lehrkräfte moderiert, werden Interaktionseffekte in das Modell aufgenommen (Modell 4).

Alle Analysen wurden mit *Mplus 7* (Muthén & Muthén, 2012) unter Nutzung der *full information likelihood* sowie unter Anwendung robuster Standardfehler zur Kontrolle der genesteten Datenstruktur in Klassen durchgeführt (Algorithm = Integration).

**Abbildung 1:** Latent-Change-Score-Modell für die Einwirkung von Binnendifferenzierung (BD) im Unterricht auf die Lesekompetenzentwicklung von Klasse 5 bis Klasse 9



## 5. Ergebnisse

Die Lehrkräfte der Sekundarstufe haben binnendifferenzierende Maßnahmen im Mittel weder stark noch schwach, sondern durchschnittlich ausgeprägt angewendet. Der Mittelwert der inneren Differenzierung über alle Wellen ( $M = 3.15$ ) lag leicht über dem theoretischen Mittelwert. Gegen Ende der Sekundarstufe scheint das Ausmaß der internen Differenzierung tendenziell etwas abzunehmen (siehe Tabelle 2).

Querschnittliche Berechnungen zeigen in den Klassenstufen 5, 7 und 9 negative Zusammenhänge zwischen der Unterrichtsdifferenzierung durch die Deutschlehrkraft und dem aggregierten Klassenmittelwert der Lesekompetenz ( $r_{K5} = -.30^*$ ,  $r_{K7} = -.31^*$ ,  $r_{K9} = -.19^*$ ): d.h. Schüler\*innen in Klassen mit durchschnittlich schlechteren Leseleistungen erhielten vermehrt differenzierten Unterricht als jene Klassen mit einem höheren Kompetenzmittelwert. Die Lesekompetenzstreuung innerhalb der Klassen hat keinen signifikanten Zusammenhang mit der Binnendifferenzierung ( $r_{K5} = .04$ ,  $r_{K7} = -.11$ ,  $r_{K9} = -.13$ ; ohne Tabelle).

Die Latent-Change-Score-Modellierung für längsschnittliche Analysen sind in Tabelle 4 abgetragen: Die Lesekompetenz der Schüler\*innen verbessert sich von Klasse 5 zu Klasse 7 bzw. von Klasse 7 zu Klasse 9 signifikant. Der Einsatz binnen-

Tabelle 4: Ergebnisse der Latent-Change-Score-Modelle

	Modell 1 (n = 5620)		Modell 2 (n = 5620)		Modell 3 (n = 5620)		Modell 4 (n = 1116)	
	Change K5-K7	Change K7-K9	Change K5-K7	Change K7-K9	Change K5-K7	Change K7-K9	Change K5-K7	Change K7-K9
Mittlere Lesekompetenzveränderung	<b>0.60</b> <b>(0.02)</b>	<b>0.44</b> <b>(0.02)</b>	<b>0.31</b> <b>(0.06)</b>	<b>0.53</b> <b>(0.05)</b>	<b>0.39</b> <b>(0.10)</b>	<b>0.73</b> <b>(0.07)</b>	<b>0.61</b> <b>(0.18)</b>	<b>0.86</b> <b>(0.30)</b>
Binnendifferenzierung Klasse 5	0.02 (0.05)		<b>0.15</b> <b>(0.07)</b>		0.11 (0.08)		0.07 (0.14)	
Binnendifferenzierung Klasse 6	-0.05 (0.05)		-0.09 (0.07)		-0.07 (0.08)		-0.05 (0.14)	
Binnendifferenzierung Klasse 7		0.03 (0.04)		0.01 (0.04)		0.01 (0.04)		-0.30 (0.25)
Binnendifferenzierung Klasse 8		0.02 (0.06)		0.02 (0.06)		0.01 (0.06)		0.21 (0.16)
Binnendifferenzierung Klasse 9		0.03 (0.05)		0.04 (0.05)		0.05 (0.05)		0.24 (0.17)
<i>Heterogenitätsmaße</i>								
Kompetenzstreuung Klasse 5					-0.02 (0.03)		-0.08 (0.04)	
Kompetenzstreuung Klasse 7						<b>-0.08</b> <b>(0.02)</b>		-0.12 (0.11)
<i>Interaktionseffekte</i>								
K5: Kompetenzstreuung * Binnendifferenzierung							<b>-0.01</b> <b>(0.02)</b>	
K7: Kompetenzstreuung * Binnendifferenzierung								-0.04 (0.04)
<i>Kontrollmaße</i>								
Schulform (0 = nicht Gymnasium; 1 = Gymnasium)		<b>0.26</b> <b>(0.02)</b>	<b>-0.07</b> <b>(0.02)</b>	<b>0.27</b> <b>(0.02)</b>	<b>-0.06</b> <b>(0.02)</b>	<b>0.19</b> <b>(0.04)</b>		-0.11 (0.10)
Geschlecht (0 = Junge; 1 = Mädchen)		<b>0.06</b> <b>(0.01)</b>	-0.02 (0.02)	<b>0.07</b> <b>(0.01)</b>	-0.02 (0.02)	<b>0.09</b> <b>(0.03)</b>		-0.03 (0.03)
Höchster Bildungsabschluss Eltern		<b>-0.06</b> <b>(0.02)</b>	-0.01 (0.02)	<b>-0.06</b> <b>(0.02)</b>	0.01 (0.02)	-0.05 (0.03)		-0.05 (0.06)
Muttersprache (0 = nur Deutsch; 1 = andere Sprache)		-0.03 (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)	0.03 (0.03)		0.06 (0.04)
aBIC	162476.49		184466.67		203222.26		50628.74	

Anmerkung. Standardisierte Koeffizienten. Fett = p < .01. aBIC = Sample-Size adjusted BIC.

differenzierender Maßnahmen der entsprechenden Jahrgänge wurde auf den jeweiligen Change Score regressiert (Change Score K5–K7: Binnendifferenzierung der Klassen 5 und 6; Change Score K7–K9: BD Klassen 7 bis 9).

Im Basismodell 1 ohne Kontrollvariablen hat die Binnendifferenzierung der Deutschlehrpersonen zunächst keinen signifikanten Einfluss auf die Entwicklung der Lesekompetenz der Schüler\*innen. Unter Kontrolle relevanter Hintergrundmerkmale (Geschlecht, Bildungshintergrund der Eltern, Sprachhintergrund) und insbesondere der Schulform ergibt sich ein positiver Effekt des binnendifferenzierenden Lehrkrafthandelns in Klassenstufe 5 auf die Lesekompetenzveränderung bis Jahrgang 7 (Modell 2).

Die Streuung der Lesekompetenz (Modell 3) in der 7. Jahrgangsstufe hat einen signifikanten, leicht negativen Effekt auf die Leistungsveränderung von Klasse 7 zu Klasse 9. Je größer die Breite der Kompetenzverteilung, desto geringer fällt die mittlere Lesekompetenzentwicklung in der zweiten Phase der Sekundarstufe I aus. Die erwarteten Interaktionseffekte (Modell 4) der Kompetenzstreuung innerhalb der Klassen mit den angebotenen Differenzierungsmaßnahmen als relevanter Moderator für die Lesekompetenzveränderung bestätigen sich nicht. Aufgrund der Datenstruktur reduziert sich für das letzte Modell die Stichprobe deutlich, die Koeffizienten bleiben aber grundsätzlich vergleichbar mit den anderen Modellen.

Auch bei für die Schulformen getrennte Analysen ergeben sich weder für die Gymnasien noch die anderen Schulen der Sekundarstufe veränderte Effektgrößen oder weitere signifikante Resultate. Einzig für das Heterogenitätsmaß, den Effekt der Kompetenzstreuung innerhalb der Klassenstufe 7, zeigt sich ein Unterschied: In Modell 3 bestätigt sich der negative Effekt des Heterogenitätsmaßes auf die Kompetenzentwicklung ausschließlich für die Nicht-Gymnasien, während für die Gymnasien diese Moderation nicht bedeutsam ist (ohne Tabelle).

## 6. Diskussion

Mit Blick auf zunehmend heterogene Klassen und die Frage der adäquaten Förderung einer solchen Schüler\*innenschaft werden in diesem Beitrag längsschnittliche Effekte binnendifferenzierender Maßnahmen im Deutschunterricht auf die Lesekompetenzentwicklung im Verlauf der Sekundarstufe anhand von Daten der Startkohorte 3 des NEPS untersucht. Für die Skala Binnendifferenzierung konnte mit vier Items über fünf Messzeitpunkte partielle skalare Messinvarianz von Klasse 5 bis Klasse 9 bestätigt werden. Dieses Ergebnis erlaubt die Nutzung der in der Lehrkraftbefragung jährlich erhobenen Skala über die Zeit hinweg als vergleichbares Konstrukt.

Es zeigt sich, dass Deutschlehrkräfte der Sekundarstufe häufiger als von anderen Studien berichtet (vgl. Lankes, 2004; Roeder, 1997; Vock & Gronostaj, 2017; Wischer, 2007, 2008) differenzierende Maßnahmen im Unterricht anwenden. Nach Klassenstufen getrennte Berechnungen ergeben, dass der Einsatz binnendif-

ferenziender Maßnahmen im Verlauf der Sekundarstufe I abnimmt. Insgesamt werden binnendifferenzierende Maßnahmen in Deutschklassen mit durchschnittlich niedrigen Lesekompetenzwert vermehrt eingesetzt; folglich erfolgt dies eher an nicht-gymnasialen Schulformen. Jedoch zeigt sich kein Zusammenhang zwischen der Heterogenität der Klasse und dem Lehrkrafthandeln. Dieser Befund ist insoweit erwartungswidrig, als dass Binnendifferenzierung insbesondere für leistungs-heterogene Klassen gefordert wird (vgl. z.B. Bräu & Schwerdt, 2005; Eisenmann & Grimm, 2016). Es scheint, dass Lehrkräfte ihre Unterrichtskonzeption und die Anpassung angebotener Materialien im Sinne differenzierter Aufgaben-schwierigkeiten eher nicht auf die konkret zu unterrichtende Klasse, die Leistungs-stände und Lernbedürfnisse ihrer Schüler\*innen abstimmen.

Längsschnittliche Analysen mit dem Latent-Change-Score-Modellansatz ergeben, dass die Auswirkungen der Maßnahmen zur inneren Differenzierung im Deutschunterricht auf die Entwicklung der Lesekompetenz der Schüler\*innen im Mittel vernachlässigbar sind. Dies könnte mit dem oben beschriebenen Phänomen zusammenhängen, dass Lehrer\*innen die eingesetzten Maßnahmen nicht optimal einsetzen.

Die binnendifferenzierenden Maßnahmen der Lehrkräfte wirken sich lediglich in Klasse 5 auf die Zunahme der Lesekompetenz bis zur Klasse 7 hin aus. Dies betont die Wichtigkeit dieser Eintrittsjahrgangsstufe in die Sekundarstufe, in der im Vergleich mit den anderen Klassenstufen der Sekundarstufe auch vermehrt Binnendifferenzierung umgesetzt wird. Mögliche Erklärung für diesen Befund, der allerdings nicht unter Berücksichtigung der Leistungsstreuung erhalten bleibt, könnte ein insbesondere in diesem Eintrittsjahr in die Sekundarstufe angemessenes Eingehen der Lehrkräfte auf die unterschiedlichen Lernstände, Tempi und bisherigen Gewohnheiten der Schüler\*innen sein: Nicht nur für die Schüler\*innen ist das Lernen an einer neuen Schule, in einer neuen Schulform, mit neuen Lehrer\*innen multipel herausfordernd, auch die Lehrkräfte stehen neuen Schüler\*innen gegenüber und betrachten ihre neue Klasse (hoffentlich) mit interessiertem Blick und der guten Absicht, ihr Lehrangebot bestmöglich an die neue Klassenkomposition anzupassen.

Unter Einbezug von Heterogenitätsmaßen im Sinne der Streuung von Lesekompetenz innerhalb der Deutschklasse zeigt sich der Befund, dass in den obe- ren Klassen (7.–9. Jahrgangsstufe) eine stärkere leistungsbezogene Heterogenität sich leicht negativ auf die mittlere Lesekompetenzentwicklung auswirkt. Dieses Resultat ist bei getrennten Analysen ausschließlich für die Stichprobe an nicht-gymnasialen Schulformen, aber nicht mehr bei Gymnasien zu finden. Die Befunde von Scharenberg (2012) zeigen in 7. und 8. Klassen keinen schulformbezogenen Effekt der Kompetenzstreuung an Hamburger Schulen, jedoch einen negativen Interaktionseffekt der Leistungsheterogenität für Gymnasien. Bei den vorliegenden bundesländerübergreifenden NEPS-Daten können unter Einbezug der Interaktion zwischen Kompetenzstreuung und Binnendifferenzierung keine signifikanten Effekte mehr berichtet werden.

Die vorliegende Studie konnte im Längsschnitt nur punktuelle Effekte der Binnendifferenzierung im Deutschunterricht in der Sekundarstufe auf die Lesekompetenzentwicklung feststellen. Die Gründe hierfür können mannigfaltig sein. So könnte eine der Stärken einer Large-Scale-Studie wie dem NEPS, welche Daten über eine sehr große Stichprobe hinweg erhebt, auch gleichermaßen eine Limitation bedeuten: Lehrkräfte berichten in qualitativen Interviews von einer breiteren Methodenvielfalt und Handhabung von differenzierenden Maßnahmen im Unterricht, welche die hier genutzte verkürzte Skala zur Binnendifferenzierung nicht umfassend abbilden kann.<sup>2</sup> Mit dem hier abgefragten Konstrukt für den Längsschnitt können keine Aussagen hinsichtlich der konkreten Implementation, Qualität und Validität verschiedener Maßnahmen getroffen werden, insbesondere fehlt die konkrete Handhabung binnendifferenzierter Förderung z.B. im Sinne von Vermittlung unterschiedlicher Lern- und Lesestrategien. Die verbleibenden Items zur Binnendifferenzierung beziehen sich vor allem auf Aufgabenmaterialien mit verschiedenen Schwierigkeitsniveaus, angepasstes Lerntempo oder Wiederholungen. Für die genannten Aspekte ist diese Skala jedoch im Rahmen von Large-Scale-Assessments, wie berichtet, gut einsetzbar.<sup>3</sup>

In den bisherigen NEPS-Schulkohorten wird die diagnostische Kompetenz der Lehrpersonen nicht erhoben. Fehlende Diagnostikkompetenzen können dazu führen, dass Lehrkräfte „Arbeitsblätter unterschiedlicher Schwierigkeit „per Augenmaß“ verteilen (Vock & Gronostaj, 2017, S. 59). Es kann nicht ausgeschlossen werden, dass die ausbleibenden Effekte der binnendifferenzierenden Maßnahmen auch auf mangelhafte Diagnosefähigkeiten zurückführbar sind.

Trotz dieser allgemeinen Einschränkungen erreichen die Analysen dieses Beitrages dennoch eine Annäherung an das effektive binnendifferenzierende Lehrer\*innenhandeln im Unterricht von gegenwärtigen deutschen Sekundarstufenschulen. Es ist gelungen zu zeigen, dass die Deutsch-Lehrpersonen im Mittel durchaus angeben, in ihren Klassen schüler- und gruppenorientierten binnendifferenzierenden Unterricht mit schwierigkeitsangepassten Aufgaben und lerntempo-anpasseten Extramaterialien durchzuführen.

Auf dem Weg zu einer vielfältigen Unterrichtspraxis, welche Schüler\*innen möglichst passend und individualisiert unterstützen und fördern könnte, ist es wesentlich, weitere Befunde bezüglich der Umsetzung von Differenzierung in der Praxis, ihrer Wirkung auf Leistungserfolge, aber ebenso auf die Motivation und die wahrgenommene Wirkung für die Schüler\*innen zu erlangen. Hier könnten weitere Studien, die sich fachspezifisch und längsschnittlich orientieren und die Praxis der Lehrpersonen umfassender und detaillierter erfassen, zu einer Optimierung der täglichen Unterrichtspraxis beitragen. Auch wäre es hinsichtlich eines gelingenden

---

2 Beispielsweise wird nicht nach dem Einsatz spezifisch adaptiver Unterrichtsformen gefragt, wie z.B. dem Gruppenpuzzle oder Stationenlernen (vgl. Warwas, Hertel & Labuhn, 2011).

3 Diese vier Items wurden beispielsweise bei IGLU innerhalb einer eindimensionalen Faktorstruktur ebenfalls bestätigt. Dort heißt das latente Konstrukt „Individuelle Förderung“ und enthält weitere 3 Items (Schulz-Heidorf, 2018, S. 186).

Umsetzens von Binnendifferenzierung angezeigt, wenn Lehrpersonen durch unterstützende Fortbildungen ihr diesbezügliches Instrumentarium erweitern (vgl. Bohl, Batzel & Richey, 2012).

## Danksagung

Diese Arbeit nutzt Daten des Nationalen Bildungspanels (NEPS): Startkohorte Klasse 5, doi:10.5157/NEPS:SC3:8.0.1. Die Daten des NEPS wurden von 2008 bis 2013 als Teil des Rahmenprogramms zur Förderung der empirischen Bildungsforschung erhoben, welches vom Bundesministerium für Bildung und Forschung (BMBF) finanziert wurde. Seit 2014 wird NEPS vom Leibniz-Institut für Bildungsverläufe e.V. (LIfBi) an der Otto-Friedrich-Universität Bamberg in Kooperation mit einem deutschlandweiten Netzwerk weitergeführt.

## Literatur

- 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 et al. (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–137). Opladen: Leske + Budrich.
- Beck, E., Baer, M., Guldimann, T., Bischoff, S., Brühwiler, C., Müller, P. et al. (2008). *Adaptive Lehrkompetenz: Analyse und Struktur, Veränderbarkeit und Wirkung handlungssteuernden Lehrerwissens*. Münster: Waxmann.
- Becker-Mrotzek, M., Lindauer, T., Pfost, M., Weis, M., Strohmaier, A. & Reiss, K. (2019). Lesekompetenz heute – eine Schlüsselqualifikation im Wandel. In K. Reiss, M. Weis, E. Klieme & O. Köller (Hrsg.), *PISA 2018 – Grundbildung im internationalen Vergleich* (S. 21–46). Münster: Waxmann.
- Blömeke, S. & Müller, C. (2009). Zum Zusammenhang von Allgemeiner Didaktik und Lehr-Lernforschung im Unterrichtsgeschehen. In M. A. Meyer, S. Hellekamps & M. Prenzel (Hrsg.), *Perspektiven der Didaktik* (S. 239–258). Wiesbaden: VS.
- Blossfeld, H.-P., Roßbach, H.-G. & Maurice, J. von (Hrsg.). (2011). *Education as a life-long process. The German National Educational Panel Study (NEPS)* (vol. 14). Wiesbaden: VS.
- Bohl, T., Batzel, A. & Richey, P. (2012). Öffnung – Differenzierung – Individualisierung – Adaptivität. Charakteristika, didaktische Implikationen und Forschungsbefunde verwandter Unterrichtskonzepte zum Umgang mit Heterogenität. In T. Bohl, M. Bönsch, M. Trautmann & B. Wischer (Hrsg.), *Binnendifferenzierung* (S. 40–69). Immenhausen bei Kassel: Prolog.
- Bönsch, M. (2011). *Erfolgreicheres Lernen durch Differenzierung im Unterricht*. Braunschweig: Westermann.
- Bönsch, M. (2012). Strategien zur Lernprozessoptimierung – Innere Differenzierung. In T. Bohl, M. Bönsch, M. Trautmann & B. Wischer (Hrsg.), *Binnendifferenzierung* (S. 9–23). Immenhausen bei Kassel: Prolog.
- Bräu, K. (2005). Individualisierung des Lernens – Zum Lehrerhandeln bei der Bewältigung eines Balanceproblems. In K. Bräu & U. Schwerdt (Hrsg.), *Heterogenität als Chance. Vom produktiven Umgang mit Gleichheit und Differenz in der Schule* (S. 129–150). Münster: LIT.
- Bräu, K. & Schwerdt, U. (Hrsg.). (2005). *Heterogenität als Chance. Vom produktiven Umgang mit Gleichheit und Differenz in der Schule*. Münster: LIT.

- 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). Münster: LIT.
- Corno, L. & Snow, R.E. (1986). Adapting teaching to individual differences among learners. In M. C. Wittrock (Hrsg.), *Handbook of research on teaching* (S. 605–629). New York: Macmillan.
- Ditton, H. (2000). Qualitätskontrolle. In A. Helmke (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule* (S. 73–92). Weinheim: Beltz.
- Eberle, T., Kuch, H. & Track, S. (2016). Differenzierung 2.0. In M. Eisenmann & T. Grimm (Hrsg.), *Heterogene Klassen – Differenzierung in Schule und Unterricht* (S. 1–36). Baltmannsweiler: Schneider Hohengehren.
- Eisenmann, M. & Grimm, T. (Hrsg.). (2016). *Heterogene Klassen – Differenzierung in Schule und Unterricht*. Baltmannsweiler: Schneider Hohengehren.
- Fischer, L., Rohm, T., Gnambs, T. & Carstensen, C. H. (2016). *Linking the data of the competence tests* (NEPS Survey Paper No. 1). Leibniz Institute for Educational Trajectories: National Educational Panel Study.
- Frahm, S., Goy, M., Kowalski, K., Sixt, M., Strietholt, R., Blatt, I. et al. (2011). Transition and development from lower secondary to upper secondary school. In H.-P. Blossfeld, H.-G. Roßbach & J. von Maurice (Hrsg.), *Education as a lifelong process. The German National Educational Panel Study (NEPS)* (vol. 14, S. 217–232). Wiesbaden: VS. <https://doi.org/10.1007/s11618-011-0188-1>
- Gehrer, K. & Artelt, C. (2013). Literalität und Bildungslaufbahn: Das Bildungspanel NEPS. In C. Rosebrock & A. Bertschi-Kaufmann (Hrsg.), *Literalität erfassen: bildungspolitisch, kulturell, individuell* (S. 168–187). Weinheim: Beltz Juventa.
- Gehrer, K., Zimmermann, S., Artelt, C. & Weinert, S. (2012). *The assessment of reading competence (Including sample items for grade 5 and 9)*, National Educational Panel Study. Retrieved from [https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/1-o-o/com\\_re\\_2012\\_en.pdf](https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/1-o-o/com_re_2012_en.pdf)
- Gehrer, K., Zimmermann, S., Artelt, C. & Weinert, S. (2013). NEPS framework for assessing reading competence and results from an adult pilot study. *Journal for Educational Research Online*. 5(2), 50–79.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung*. Münster: Waxmann.
- Hanke, P. (2005). Unterschiedlichkeit erkennen und Lernprozesse in gemeinsamen Lernsituationen fördern. Förderdiagnostische Kompetenzen als elementare Kompetenzen im Lehrerberuf. In K. Bräu & U. Schwerdt (Hrsg.), *Heterogenität als Chance. Vom produktiven Umgang mit Gleichheit und Differenz in der Schule* (S. 115–128). Münster: LIT.
- Hattie, J. A. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement. „Reveals teaching’s holy grail“*. Abingdon: Routledge.
- Hattie, J. A. (2018). *Visible learning plus: 250+ influences on student achievement*. Abgerufen am 06.06.2019 unter <https://us.corwin.com/sites/default/files/2507influencesfi10.1.2018.pdf>
- Helmke, A. (2007). *Unterrichtsqualität erfassen, bewerten, verbessern* (Schulisches Qualitätsmanagement). Seelze: Kallmeyer.
- Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Franz Emanuel Weinert gewidmet. Seelze: Klett Kallmeyer.
- Helmke, A. & Weinert, F.E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 71–177). Göttingen: Hogrefe.
- Hertel, S. (2014). Adaptive Lerngelegenheiten in der Grundschule: Merkmale, methodisch-didaktische Schwerpunktsetzungen und erforderliche Lehrerkompetenzen. In

- B. Kopp (Hrsg.), *Individuelle Förderung und Lernen in der Gemeinschaft* (S. 19–34). Wiesbaden: Springer.
- Klafki, W. & Stöcker, H. (1986). Innere Differenzierung des Unterrichts. *Zeitschrift für Pädagogik*, 22(4), 497–546.
- Klieme, E. & Warwas, J. (2011). Konzepte der Individuellen Förderung. *Zeitschrift für Pädagogik*, 57(6), 805–818.
- Köller, O. (2014). What works best in school? In E. Terhart (Hrsg.), *Die Hattie-Studie in der Diskussion* (S. 24–37). Seelze: Kallmeyer Klett.
- Krannich, M., Jost, O., Rohm, T., Koller, I., Carstensen, C. H., Fischer, L. et al. (2017). *NEPS Technical report for reading: Scaling results of Starting Cohort 3 for grade 7 (NEPS Survey Paper No. 14)*. Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Lankes, E.-M. (2004). Leseunterricht in der Grundschule: Unterschiede zwischen Lehrkräften im internationalen Vergleich. *Zeitschrift für Erziehungswissenschaft*, 7(4), 551–568.
- Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B. & d'Apollonia, S. (1996). Within-class grouping. A meta-analysis. *Review of Educational Research*, 66(4), 423–458.
- McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, 60, 577–605. <https://doi.org/10.1146/annurev.psych.60.110707.163612>
- Messner, H. (2009). Eine integrationsfähige Schule für alle. In D. Bosse & Posch (Hrsg.), *Schule 2020 aus Expertensicht: Zur Zukunft von Schule, Unterricht und Lehrerbildung*. (S. 111–116). Wiesbaden: VS.
- Muthén, L. K. & Muthén, B. O. (2012). Mplus User's Guide. Seventh Edition. Los Angeles.
- Nusser, L., Carstensen, C. H. & Artelt, C. (2015). Befragung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf Lernen: Ergebnisse zur Messinvarianz. *Empirische Sonderpädagogik* (2), 99–116.
- Paradies, L. & Linser, H. J. (2017). *Differenzieren im Unterricht. Sekundarstufe I + II*. Berlin: Cornelsen.
- Pelz, S., & Zielonka, M. (2017). *NEPS technical report: Implementation of the ISCED – 97, CASMIN and years of education classification schemes in SUF Starting Cohort 4* (NEPS Research Data Paper). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Pohl, S. & Carstensen, C. H. (2013). Scaling of the competence tests in the National Educational Panel Study – Many questions, some answers, and further challenges. *Journal for Educational Research Online*, 5(2), 189–216.
- Pohl, S., Haberkorn, K., Hardt, K. & Wiegand, E. (2012). *NEPS technical report for reading: Scaling results of Starting Cohort 3 in fifth grade* (NEPS Working Paper No. 15), University of Bamberg, National Educational Panel Study.
- Praetorius, A.-K., Klieme, E., Herbert, B. & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of three basic dimensions. *ZDM – The International Journal on Mathematics Education*, 50(3), 407–426.
- Prenzel, M., Seidel, T., Lehrke, M., Rimmele, R., Duit, R., Euler, M., Geiser, H., Hoffmann, L., Müller, C. & Widodo, A. (2002). Lehr-Lernprozesse im Physikunterricht – eine Videostudie. *Zeitschrift für Pädagogik*, 45. Beiheft, 139–156.
- Roeder, P. M. (1997). Binnendifferenzierung im Urteil von Gesamtschullehrern. *Zeitschrift für Pädagogik*, 43(2), 241–259.
- Rosebrock, C. & Nix, D. (2011). *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung*. Baltmannsweiler: Schneider.
- Saalfrank, W.-T. (2012). Differenzierung. In E. Kiel (Hrsg.), *Unterricht sehen, analysieren, gestalten* (S. 65–96). Bad Heilbrunn: Klinkhardt.

- Scharenberg, K. (2012). *Leistungsheterogenität und Kompetenzentwicklung. Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie.* Münster: Waxmann.
- Scharl, A., Fischer, L., Gnambs, T. & Rohm, T. (2017). *NEPS technical report for reading: Scaling results of Starting Cohort 3 for grade 9 (NEPS Survey Paper No. 20)*, Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Scheerens, J. (2000). *Improving school effectiveness*. Paris: Unesco.
- Schneider, C. & Ludwig, P. H. (2012). Auswirkungen von Maßnahmen der inneren Differenzierung auf Schulleistung und Fähigkeitsselbstkonzept im Vergleich zu Auswirkungen der äußeren Differenzierung In T. Bohl, M. Bönsch, M. Trautmann & B. Wischer (Hrsg.), *Binnendifferenzierung* (S. 72–106). Immenhausen bei Kassel: Prolog.
- Schrader, F.-W. & Helmke, A. (2008). Unter der Lupe: Wie in deutschen Klassenzimmern differenziert wird. *Magazin für die Grundschule*, 20, 3–5.
- Schroeders, U. & Gnambs, T. (2018). Degrees of freedom in multigroup confirmatory factor analyses. Are models of measurement invariance testing correctly specified? *European Journal of Psychological Assessment*, 1–9. <https://doi.org/10.1027/1015-5759/a000500>
- Schulz-Heidorf, K. (2016). Adapted teaching: A chance to reduce the effect of social origin? A comparison between Germany and Norway, using PIRLS 2011. *Tertium Comparationis*, 22(2), 230–259.
- Schulz-Heidorf, K. (2018). Individuelle Förderung im Unterricht: Eine Chance, Schulerfolg von der sozialen Herkunft zu entkoppeln? Eine Diskussion erwartungswidriger Befunde. In A. Krüger, F. Radisch, A. S. Willems, T. Häcker & M. Walm (Hrsg.). *Empirische Bildungsforschung im Kontext von Schule und Lehrer\*innenbildung* (S. 181–193). Bad Heilbrunn: Klinkhardt.
- Schwab, S. & Helm, C. (2015). Überprüfung von Messinvarianz mittels CFA und DIF-Analysen. *Empirische Sonderpädagogik* (3), 175–193.
- Slavin, R. E., Cheung, A., Groff, C. & Lake, C. (2008). Effective reading programs for middle and high schools. A best-evidence synthesis. *Reading Research Quarterly*, 43(3), 290–322.
- Steyer, R. (2009). Modeling true intraindividual change in SEM. In T. D. Little, K. U. Schnabel & J. Baumert (Hrsg.), *Modeling longitudinal and multilevel data. Practical issues, applied approaches, and specific examples* (S. 109–126). Mahwah, N.J.: Lawrence Erlbaum.
- Trautmann, M. & Wischer, B. (2009). Das Konzept der Inneren Differenzierung – eine vergleichende Analyse der Diskussion der 1970er Jahre mit dem aktuellen Heterogenitätsdiskurs. In M. A. Meyer, S. Hellekamps & M. Prenzel (Hrsg.), *Perspektiven der Didaktik* (S. 159–172). Wiesbaden: VS.
- Trautmann, M. & Wischer, B. (2011). *Heterogenität in Schule. Eine kritische Einführung*. Wiesbaden: VS.
- Vock, M. & Gronostaj, A. (2017). *Umgang mit Heterogenität in Schule und Unterricht*. Berlin: Friedrich-Ebert-Stiftung.
- Warm, T. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450.
- Warwas, J., Hertel, S. & Labuhn, A. S. (2011). Bedingungsfaktoren des Einsatzes von adaptiven Unterrichtsformen im Grundschulunterricht. *Zeitschrift für Pädagogik*, 57(6), 854–867.
- Weinert, F. E. (1997). Notwendige Methodenvielfalt. Unterschiedliche Lernfähigkeiten erfordern variable Unterrichtsmethoden. *Friedrich Jahresheft*, 15, 50–52.
- Weinert, F. E. & Helmke, A. (1996). Der gute Lehrer: Person, Funktion oder Fiktion? *Zeitschrift für Pädagogik* (34), 223–233.

- Wischer, B. (2007). Wie sollen LehrerInnen mit Heterogenität umgehen? Über „programmatische Fallen“ im aktuellen Reformdiskurs. *Die Deutsche Schule*, 99(4), 422–433.
- Wischer, B. (2008). Binnendifferenzierung ist ein Wort für das schlechte Gewissen des Lehrers. *Erziehung und Unterricht* (9–10), 714–722.
- Wischer, B. & Trautmann, M. (2010). Innere Differenzierung als (unterschätzte) pädagogische Herausforderung. Zu den Grenzen einer Reformstrategie. In S. Boller (Hrsg.), *Innere Differenzierung in der Sekundarstufe II. Ein Praxishandbuch für Lehrer/innen* (S. 158–166). Weinheim: Beltz.

Divan Mouton, Hui Zhang & Bernhard Ertl

## **German university student's reasons for dropout: Identifying latent classes**

### **Abstract**

*The reasons for students discontinuing studies in university courses are manifold and often represent a combination of different factors. Theoretical work and empirical findings have already conceptualized various factors relevant to dropouts, but little is known about the interplay of these factors. The present paper deals with this topic and analyses how different classes of dropouts can be defined on the basis of the reasons for their termination. The study draws on data from the National Education Panel Study (NEPS), Starting Cohort 5 – first semester students, which asked drop-outs about the reasons for their drop-out with a 24-item Reasons for Dropout Questionnaire (RDQ). Longitudinally, the important reasons for dropout at the beginning of studies related to the student's relationship to the study program or institution, but they do not correspond to important reasons at the end, which mainly relate to socio-economic reasons. In addition, most drop-outs occur at the beginning of studies and then steeply decline over time. A latent class analysis revealed 12 specific profiles of students that are mainly characterized by different combinations of reasons for dropout. These classes correspond to other contextually relevant constructs beyond the RDQ, such as performance and socio-economic-factors, academic self-concept and the intention to dropout. Such profiles should be considered for targeted interventions.*

### **Keywords**

*Dropout; University; Latent class; NEPS*

---

Divan Mouton, M.Sc. (corresponding author) · Hui Zhang, M.Sc. · Prof. Dr. Bernhard Ertl,  
Institute of Education, Universität der Bundeswehr München, Werner-Heisenberg-Weg 39,  
85579 Neubiberg, Germany  
e-mail: [divan.mouton@unibw.de](mailto:divan.mouton@unibw.de)  
[hui.zhang@unibw.de](mailto:hui.zhang@unibw.de)  
[bernhard.ertl@unibw.de](mailto:bernhard.ertl@unibw.de)

# Gründe für den Studienabbruch deutscher Universitätsstudierender: Analyse latenter Klassen von Studienabrechern

## Zusammenfassung

*Die Gründe für Studienabbrüche in Universitätsstudiengängen sind mannigfaltig und stellen dabei oft ein Zusammenspiel verschiedener Faktoren dar. Theoretische Arbeiten sowie empirische Befunde konzeptualisieren bereits verschiedene Faktoren, die für Studienabbrüche relevant sind; allerdings ist bisher wenig über das Zusammenspiel dieser Faktoren bekannt. Die vorliegende Arbeit behandelt dieses Thema und analysiert, wie sich verschiedene Klassen von Studienabrechern an Hand der Gründe für ihren Abbruch definieren lassen. Die Studie greift auf Daten des Nationalen Bildungspanels (NEPS), Startkohorte 5 – Erstsemester zurück, das Studienabrecher mit einem 24-Item Fragebogen zu Gründen für ihren Studienabbruch befragt hat. Im Zeitverlauf bezogen sich die primären Gründe für den Studienabbruch zu Beginn des Studiums auf die Beziehung des Studierenden zum Studiengang oder zur Institution und veränderten sich hin zum Ende des Studiums auf hauptsächlich sozioökonomische Gründe. Darüber hinaus erfolgen die meisten Studienabbrüche zu Beginn des Studiums und nehmen im Laufe der Zeit stark ab. Eine latente Klassenanalyse konnte 12 spezifische Profile von Studienabrechern identifizieren, die überwiegend aus der Kombination von verschiedenen Abbruchgründen bestehen. Diese Klassen korrespondieren mit anderen kontextuell relevanten Konstrukten jenseits des Fragebogens, wie z.B. Leistung, sozioökonomischen Faktoren, akademischem Selbstkonzept und der Intention zum Studienabbruch. Solche Profile sollten für zielgerichtete Interventionen berücksichtigt werden.*

## Schlagworte

*Studienabbruch; Universität; Latente Klassenanalyse; NEPS*

## 1. Introduction

The rate of student dropout in formal higher education for the past 50 years has been a concern both in Germany and internationally. One aspect of this concern is the lack of qualified professionals to occupy important positions in industry (Kruppe & Baumann, 2019). Initiatives in Germany have targeted reducing the number of student dropouts as a strategy to increase the number of professionals who can enter the labor market (see e.g. “Wege zu mehr MINT-Absolventen”, Brossardt, 2012).

Among OECD countries in 2005 an average of 31 % of students completed their studies without their first degree. German students fared better with only 23 % unable to obtain their first degree before completion (OECD, 2008). In 2006

the German educational system undertook a major reform through the Bologna process to introduce the degree qualification framework and phase out the diploma system (Keeling, 2006). While the effects of the reform process on drop-out rates were unclear to some (Horstschräer & Spietsma, 2015), others suggested that there was a strong rise in dropouts following the changes (Heublein, 2014). Heublein (2014) found that on average 28 % of German bachelor students in 2010 and 2012 completed their studies without their first degree. These dropout rates differ between subject areas, with it being notably higher for STEM (Science, Technology, Engineering, & Mathematics) subjects (see Heublein & Schmelzer, 2018; Ispphording & Qendrai, 2017; Xianglei & Mathew, 2013). Differences in drop-out rates between STEM and Non-STEM subjects have been criticized by German governmental authorities (e.g. Bayerischer Oberster Rechnungshof, 2019, p. 216).

The reasons why students choose to discontinue their studies are important guidelines to dictate how university policies are made to address the issue of student dropout. Ispphording and Wozny (2017) point out two opposing perspectives for institutional accountability. On the one hand, should the reasons for student dropout lie within the sphere of influence of the university institutions, such as poorly organized study programs or a lack of academic support for students, then the institutions would be responsible to ameliorate the causes. This stance was also suggested by the Bayerischer Oberster Rechnungshof (2019). On the other hand, if reasons for dropout are outside of the institution's control, such as a student's educational background or socio-economic status of the parent's home, then broadening access to education will disadvantage the university (Ispphording & Wozny, 2017). This paper will first theoretically conceptualize the phenomenon of dropout before an analysis of how reasons for dropout develop longitudinally, an investigation into the need for a multi-causal model of dropout and the identification of distinct latent classes of dropout students: The paper will conclude by contextualizing the latent dropout classes using relevant indicators for dropouts.

## 2. Conceptualizing dropout

The phenomenon of dropout can be seen from either the perspective of the university institution who loses a student or from the perspective of the student who abandons pursuing a qualification. On the one hand, the institution's main aim is to retain their students (Tinto, 2006) or avoid the conditions which cause student attrition (Bean, 1980; Ishitani, 2006). On the other hand, students are interested in their persistence within the higher education system until their degree completion (Berger & Braxton, 1998; Cabrera, Nora, & Castaneda, 1993). The definition of dropout to an institution can be at odds to that of their students who leave to complete their degree in another university institution. This study will define a drop-out as students who had an initial educational goal to complete at least a bachelor's degree, but who did not achieve it (Berger, Ramirez, & Lyons, 2012). Furthermore,

the dropout status may be temporary as a student could re-enter the system and then dropout again several times on their path to graduation. Naturally, students may make corrections to their choice of study subject, their institution or both as their experience within their studies changes their perspectives (Seidman, 2012). In order to avoid these conceptual issues our study limited the definition of dropout to study first-time dropouts only, without considering students who return to their studies.

## 2.1 Theories and models of dropout

Spady (1971), Tinto (1975), and later Bean (1980) started investigating dropout as the retention of students became a priority within the 1970's. Spady's (1971) *Undergraduate Dropout Process Model* became the foundational literature for understanding student dropout and retention. According to the model, students need to integrate into the institution's social and academic systems in order to foster strong relationships with peers and faculty, thus motivating the student to persist was an essential task of the university. Failure to integrate was then theoretically analogous to the withdrawal process seen within the suicide phenomenon espoused by Durkheim (1997), which is a view shared by Tinto (1975). According to Spady (1971), social integration is influenced by the student's normative congruence to the institution and support from friendships while academic integration is influenced by grades and intellectual development.

Tinto (1975) further elaborated on Spady's model (1971) with his *Student Integration Model*, which is the most widely cited model of student dropout. Tinto expanded on the description of social and academic integration, and how student's attributes, such as their family background, skills and abilities, influence their goals and commitment to their institution. Tinto's model has been criticized for placing most of the responsibility on the student's inability to integrate, and less regard on external factors (Cabrera et al., 1993).

Bean's *Student Attrition Model* (1980) departed from Spady (1971) and Tinto's model (1975) through the criticism of Durkheim's (1997) suicide theory, claiming that it did not present enough evidence to support this connection. Bean's Model is based primarily on Price's law of employee turnover in work organizations (Price, 1977). The claim states that students would dropout for analogous reasons as employees leaving a company. Bean also argued that student satisfaction closely links to organizational determinants. Analogous concepts such as employees' "pay" or "wages" correspond to student's grade point average (GPA). In Bean's model the dropout outcome is dependent on performance indicators, socio-economic status variables and organizational determinants, while institutional commitment and satisfaction are intervening variables.

Another contribution provided by Bean (1985) was the concept of dropout intention which serves as an early-warning sign for dropout students before they actually decide to dropout (Cabrera et al., 1993). Despite the predictive strength of

dropout intention, students do still leave university despite having no intentions to dropout. Such students “can be accurately specified after the fact, but not predicted” due to the nature of their reasons for dropout, such as illness or family crisis (Bean, 1985, p. 36)

## **2.2 Dropout within the German university context**

Theoretical concepts espoused by Tinto (1975) and Bean (1980) were developed for students on North American campuses which do not directly relate to the German university context (Heublein & Wolter, 2011). The most comprehensive study of dropout within Germany was conducted by Heublein and colleagues in the winter through to the summer semester in 2007/2008 (Heublein et al., 2010) and again in the summer semester 2014 (Heublein et al., 2017). Heublein, Hutzsch, Schreiber, Sommer and Besuch (2010) implemented a study on students reasons for dropout based on problem-centered interviews from a representative sample. A variety of reasons for dropout were identified and analyzed, which yielded eight factors including (1) performance requirements being too high, (2) financial issues, (3) failing intermediate and final exams, (4) a lack of motivation to study, (5) inadequate study conditions (6) professional reorientation, (7) family or personal problem situations and (8) dropping out due to illness. Performance problems, a lack of motivation to study and financial problems were found to be the dominant reasons for dropout.

The *Model for Student Dropout Processes*, developed by Heublein and colleagues (Heublein et al., 2010; Heublein et al., 2017), describes three phases: the pre-study phase, the study phase and the decision to drop out. The pre-study phase involves the relatively stable traits and background-characteristics that the student brings to the study condition. In contrast, the study phase is a dynamic process explained by the complex interplay between internal and external factors during the study situation. While internal factors are the concrete expression of students within their study situation, such as study behaviors, study motivation and performance, external factors are the environmental conditions outside of their control, such as living conditions and alternatives to the current studies, which affect student's individual study situations. A student's dropout decision is therefore rarely due to only a single reason but rather an idiosyncratic combination of factors that accumulate to reinforce their decision to drop out.

Klein and Stocké (2016) agree that multi-causal models, such as that espoused by Heublein et al. (2010), are theoretically necessary to understand the complex nature of dropout but argue that most German language dropout studies frequently use inappropriate methods of analysis to understand the significance and explanatory power of individual factors. Klein and Stocké (2016) also identified evidence that German university dropout factors are time-variant, specifically contrasting the early and late stages of studies.

Furthermore, many comparisons have been made with US-American studies, which should not be directly related to the German university context without certain reservations (Heublein & Wolter, 2011)

### **2.3 Factors affecting dropout**

Understanding the impact of individual dropout factors is of great importance when deciding which factors to include in a study of dropout. A traditional type of predictor for dropout are *performance indicators*, such as high school GPA and standardized test scores (Robbins et al., 2004). High school GPA has been found to better predict college graduation compared to standardized colleague admission tests (e.g. Scholastic Aptitude Test, SAT, Galla et al., 2019) because the latter heavily focus on students' cognitive abilities over a short time-frame, while high school GPA illustrate a students' competencies over a longer time period and also captures non-achievement factors related to success, such as students' attitude, behavior, effort and their ability to self-regulate.

Individuals' subjective thoughts about their abilities could also influence university students' dropout (Robbins et al., 2004). For example, a student's poor academic self-concept, which relates to their thoughts and feelings in reference to themselves as an academic object, can hinder the expression of their abilities. A student's opinion about their subjective likelihood to complete a degree (Esser & Stocké, n.d.) or, conversely, their intention to dropout (Bean, 1985) are also strong predictors of student's goal commitment.

Another important traditional dropout predictor is *socio-economic status* which can indicate barriers for students to successfully earn a degree (Robbins et al., 2004; Caldas & Bankston, 1997). Robbins and colleague's meta-analysis shows a small effect of financial supports on university students' retention. Socio-economic barriers do not only include financial indicators but is also related to parents' social status (Caldas & Bankston, 1997). A lower socio-economic status has shown evidence to reduce the likelihood to reach degree completion among first generation university students (Robbins et al., 2004). Ishitani (2006) found that several variables to affect dropout across years in a 6-year study. Ma and Cragg (2013) also indicate differences in the impact of dropout factors over time.

Robbins et al. (2004) concluded that a multidimensional model is necessary to understand the complex interplay of all the factors that influence dropout. However, existing literature on the dropout phenomenon (Bean, 1980; Berger & Braxton, 1998; Cabrera et al., 1993; Spady, 1971; Tinto, 1975) has predominantly used "variable-centered" approaches (see Muthén & Muthén, 2000) that focus on relationships amongst university dropout factors. These variable-centered methods, such as path analysis, factor analysis or structural equation models, require arbitrary cutoffs to discover factor structures common to the entire sample and do not consider the interplay between factors amongst classes of individuals. Comprehensive typologies of student's reasons for dropping out have been sugge-

sted (Griesbach, Lewin, Heublein, & Sommer, 1998) and attempts to create typologies using person-centered approaches have been made, albeit with data that lacks additional predictors to contextualize the classes (Voelkle & Sander, 2008).

### 3. Research question

This paper will follow such a person-centered approach and delve into the drop-out phenomenon by conducting a latent class analysis (LCA). First the development of dropout factors will be discussed longitudinally before focus turns to the multi-causal aspect of dropout. Finally, the interplay between the reasons for drop-out that form distinct latent classes of student dropouts are described using both the item means and other contextually similar variables. The research questions are therefore:

*RQ1: How far can students' reasons for dropout be distinguished longitudinally?*

This first research question is rather exploratory. However, literature presents the evidence points to the unique effects of several variables that predict dropout vary over time (Ishitani, 2006, Ma & Cragg, 2013, Klein & Stocké, 2016). Some drop-out factors such as false expectations would be expected to drop out early. Other factors such as the broader socio-economic issues, which include family and financial problems, are less clear when they will dropout out according to the literature.

*RQ2: How far can single reasons explain students' dropout?*

The broader literature suggests that single reason for dropout, such as performance problems cannot entirely explain the variance amongst dropouts (Voelkle & Sander, 2008). The presence of a combination of factors is expected rather than single, stand-alone reasons for dropout (Robbins et al., 2004).

*RQ3: How far can latent dropout classes be distinguished by their reasons for dropout?*

Based on our theoretical background presented by Heublein et al. (2010) and Heublein et al. (2017) we expect that several classes can be identified with different reasons for dropout, and that these may reflect several combinations of drop-out factors.

*RQ4: How far can latent dropout classes be distinguished by contextual variables?*

We expect differences between the classes to be further enhanced using variables that are contextually similar to salient dropout factors such as performance and socio-economic indicators (Robbins et al., 2004). An advantage of using contextual variables within NEPS are that they were collected early on in the course of

studies, which mitigate issues regarding causality from retrospectively investigating reasons for dropout in RQ3.

## 4. Methods and procedure

### 4.1 Sample

The sample analyzed is a cohort of first year university students from the National Education Panel Study (NEPS) studied from 2010 to 2016 (SC5: 10.0.0, Blossfeld et al., 2011). Students completed the Reasons for Dropout Questionnaire (RDQ) if they indicated that they had either temporarily discontinued or abandoned their studies during one of the regular interview episodes. The data set comprises of a total of 1461 responses. 74 responses came from students who had already answered the RDQ once before. Only the first response was retained from every unique dropout. After correcting for purposely unanswered responses by males regarding pregnancy, we removed all missing values and maintained 1353 dropouts. Finally, we identified nine outliers as scoring three standard deviations above the median of all items. The outliers were subsequently removed and remaining 1344 (43 % male; 57 % female) student's responses for subsequent analyses.

### 4.2 Measures

#### 4.2.1 Reasons for dropout

The Reason for Dropout Questionnaire (RDQ, Cronbach's  $\alpha = .77$ , see FDZ-LIfBi, 2018, pp. 1272–1295) consists of 24 items which were adapted from Heublein et al. (2010) questionnaire version with 31 items. All items are rated on a six-point Likert scale, ranging from 1 (*plays no role at all*) to 6 (*plays a very important role*), based on the extent to which students think it is an important reason for their dropout.

#### 4.2.2 Intentions for dropout

Intentions to dropout was measured by five items (Cronbach's  $\alpha = .85$ ) from Trautwein, Kölner and Watermann (2004). Three items focused on dropping out from university and two items focused on changing the particular study field. All items are rated on a four-point Likert scale, ranging from 1 (*does not apply at all*) to 4 (*applied completely*), based on how strongly students have the intention to dropout such as "I will complete these studies no matter what". Items were reverse coded as appropriate.

#### **4.2.3 Academic self-concept**

Academic self-concept is based on four items (Cronbach's  $\alpha = .83$ ) from Dickhäuser, Schöne, Spinath and Stiensmeier-Pelster (2002). All items are rated on seven-point Likert scale, ranging from 1 (*low*) to 7 (*high*), based on how students would describe themselves in terms of their studies.

#### **4.2.4 High school Grade Point Average (GPA)**

High school GPA was collected using the German 1.0 to 5.0 scale where a 1.0 is a higher score.

#### **4.2.5 Subjective likelihood of graduation**

Subjective likelihood of graduation is one item used to ascertain how likely students think that they will complete the degree program and graduate (Esser & Stocké, n.d.).

#### **4.2.6 Socio-economic status**

The “HISEI Score” was calculated following an approach used in a PISA study (OECD’s Programme for International Student Assessment) by choosing the higher value of the two parents International Socio-Economic Index of Occupational Status (ISEI-o8; Ganzeboom, 2010). The ISEI-o8 has a range from 10, being the less prestigious occupations, up to 90 (Ganzeboom & Treiman, 1992).

In addition, students’ financial background was measured by one item, asking how well students do get by with the money they have in an average month rated on a five-point Likert scale, ranging from 1 (*very poor*) to 5 (*very good*).

#### **4.2.7 Study satisfaction**

Study satisfaction was measured by one item and rated on a ten-point Likert scale, ranging from 0 (*completely dissatisfied*) to 10 (*completely satisfied*), asking how far the students was satisfied with higher education.

#### **4.2.8 Field of study**

Eight study subject domains were classified according to the German Statistics Office (Destatis, 2018).

### 4.3 Statistical analysis

This paper will apply an LCA, which is a statistical method that identifies individuals with distinct typologies or systematic diverging patterns of responses to a set of input variables, to classify subgroups of student dropout. LCA is a type of finite mixture model of distributions which constructs classes by maximizing the homogeneity within, while also maximizing the heterogeneity between each subgroup (Baudry, Raftery, Celeux, Lo, & Gottardo, 2010; Collins & Lanza, 2009). LCA has an advantage over other traditional “variable-centered” approaches by not relying on cutoffs to identify classes. For this study we are interested in grouping students with similar reasons for dropout together. For the LCA, we applied the *MCLUST* package (Fraley, Raftery, Scrucca, Murphy, & Fob, 2018) using the R Studio statistical software (R Core Team, 2019).<sup>1</sup>

Significant differences amongst classes on contextual variables are measured using means and confidence intervals (Field, Miles, & Field, 2019).

## 5. Results

### 5.1 Sample characteristics

The sample consisted of 1344 students (57 % female) with a mean age of 26.8 years at first dropout ( $SD = 5.7$ ;  $min = 21$ ;  $max = 67$ ). Considering that the NEPS panel had an oversampling of women (60 % women in the full sample; see Ertl & Hartmann, 2019), the proportion of female students is less than expected. Dropouts mainly distribute across four main Destatis (2018) subject groups: Linguistic and cultural studies (25 %), Law, economics and social science (23 %), Mathematics and natural science (25 %) and Engineering (17 %; see Supplement D).

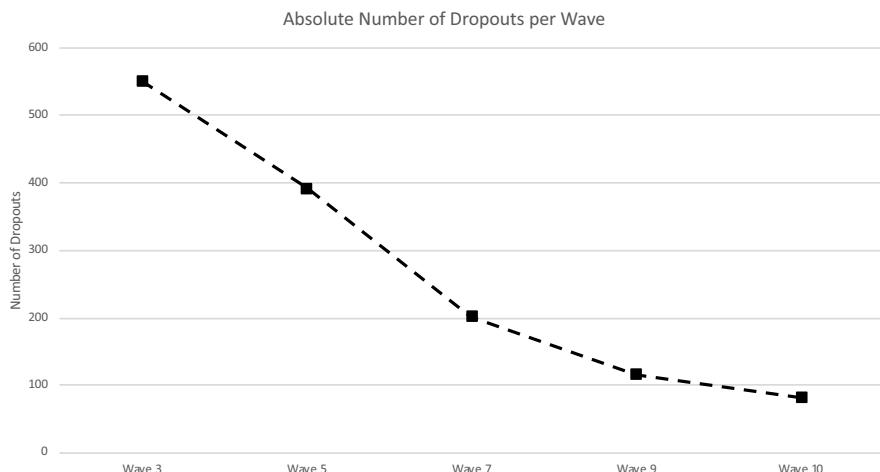
### 5.2 Research question 1: Longitudinal effects amongst of reasons for dropout

Figure 1 shows a continuous decline in dropout numbers longitudinally from 551 in wave three down to 82 in wave ten.

---

<sup>1</sup> According to Haughton, Legrand, and Woolford (2009), *MCLUST* “performs latent class cluster analysis on continuous data” (p. 88). Although other authors like Collins and Lanza (2010, p. 6) would label that latent profile analysis, we will further keep the term of latent class analysis.

Figure 1: Absolute numbers of dropout students per wave. Exact numbers can be found in Supplement A1.



We furthermore could observe a slight decline of students' intention to drop out from 1.96 in wave two to 1.61 in wave eight in the dropout students' cohort which was far less developed in the total cohort (see Supplement A2). Looking now at the impact of the different reasons for dropout, we find several reasons have significantly higher impact on early dropouts in wave three, especially the relation between the student and the study program/institution, specifically: *study requirements, overcrowded lectures, anonymity in the university, and false expectations* (see Table 1). In contrast reasons that have a higher impact at the end of the study course relate more to the students' life circumstances, specifically, the need for *earning money quickly, family reasons, incompatibility of degree course and employment, and illness*.

### 5.3 Research question 2: Multi-dimensional reasons for dropout

Descriptively, the number of students who provide a single major reason for dropout with no other moderately important reason are relatively few (53 cases; 4 %). The cases with a single major reason for dropout mostly provide reasons such as illness, studying abroad or family-based reasons. The trend is similar for cases with both two and three major reasons, where both groups each indicated 28 students whose predominant reasons for dropout related to either family reasons, including child care and pregnancy, or illness. Table 2 gives an overview of all the reasons that students indicated for one, two and three major reasons for dropout.

**Table 1:** Means, standard deviations, *t*-test results, and effect sizes comparing early (wave 3) and late (wave 10) dropout students on RDQ. Information for mean per wave can be found in Supplement E.

	Early (Wave 3)		Late (Wave 10)		<i>t</i>	df	<i>p</i>	<i>d</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>					
Perform req.	2.67	1.60	2.55	1.59	—	0.64	106.92	.52	.08
Study req.	2.96	1.62	2.51	1.56	►	2.42	108.88	.02	.28
Exam mat.	2.98	1.66	2.73	1.66	—	1.29	106.71	.20	.15
Fail exam	2.69	1.97	3.15	2.04	◀	-1.88	104.58	.06	.23
Lecture supp.	2.72	1.58	2.51	1.64	—	1.05	104.47	.30	.13
Study org.	2.72	1.54	2.62	1.73	—	0.48	101.01	.63	.06
Overcrowded	2.26	1.53	1.70	1.33	►	3.54	115.64	.00	.38
Anonymity	1.91	1.31	1.56	1.03	►	2.75	123.38	.01	.27
Practical work	3.25	1.78	2.95	1.83	►	1.37	105.24	.17	.17
Practical relev.	2.81	1.70	2.70	1.90	—	0.52	101.19	.60	.07
Interest in sub.	3.41	1.86	3.11	1.78	►	1.44	109.00	.15	.16
Interest in prof.	2.43	1.69	2.20	1.49	—	1.31	114.26	.20	.14
False expec.	3.50	1.73	2.91	1.71	►	2.90	107.20	.01	.34
Fin problems	2.27	1.67	2.67	1.81	◀	-1.87	102.39	.06	.24
Earn money	1.79	1.41	2.28	1.79	◀	-2.35	96.63	.02	.33
Pregnancy	1.23	1.01	1.35	1.27	—	-0.87	96.74	.38	.12
Child care	1.39	1.22	1.46	1.31	—	-0.47	103.01	.63	.06
Fam reasons	1.87	1.64	2.39	1.87	◀	-2.36	100.30	.02	.31
Personal suit.	2.64	1.52	2.53	1.55	—	-0.59	105.71	.56	.07
Incompatibility	2.14	1.61	2.65	1.88	◀	-2.35	99.43	.02	.31
Job offer	2.21	1.73	1.99	1.58	—	1.18	112.27	.24	.13
Lack opportun.	2.02	1.53	1.93	1.39	—	0.61	113.72	.54	.07
Illness	1.46	1.27	2.06	1.87	◀	-2.82	92.46	.01	.44
Abroad	1.17	0.77	1.17	0.81	—	0.01	103.51	.99	.01

Note. *d* = effect size measured in Cohen's *d*; Perform req. = Could not meet performance requirements for the studies; Study req. = Study requirements too high; Exam mat. = Extensive contents of study and examination; Fail exam = Did not pass examinations; Lecture supp. = Lack of support from lecturers; Study org. = Lack of organization of the studies; Overcrowded = Overcrowded lectures; Anonymity = Anonymity in the university; Practical work = Wanted to do practical work; Practical relev. = Lack of professional and practical relevance of the studies; Interest in sub. = Decrease interest in subject; Interest in prof. = Lack of interest in the professions possible with the degree obtained; False expec. = Wrong expectations in relation to the studies; Fin problems = Financial shortage; Earn money = Wanted/had to earn money as quickly as possible; Pregnancy = Pregnancy; Child care = Incompatibility of study and child care; Family reasons = Family reasons; Personal suit. = Doubt about personal suitability for the degree course; Incompatibility = Incompatibility of study and gainful activity; Job offer = Interesting job offer; Lack opportun. = Poor job opportunities in my field of study; Illness = Illness. Abroad = Study abroad or internship abroad.

**Table 2:** Outstanding items for cases with singular reasons for dropout. For only one outstanding item as reason for dropout please see Supplement I.

Items	Frequency
1 Family	40
2 Pregnancy	31
3 Child care	27
4 Illness	23
5 Abroad	13
6 Incompatibility	10
7 Practical work	10
8 Job offer	9
9 Fail Exam	8
10 Fin Problems	5
11 Earn money	3
12 Interest in sub.	3
13 Perform req.	3
14 Practical relev.	2
15 Study org.	1
16 Study req.	1
17 Lack Opportun.	1
18 False expec.	1
19 Anonymity	1
Total	192

*Note.* Frequency = Frequency of outstanding reasons for dropout if a person named either one, two or three major reasons in the decision to dropout (of either 5 or 6 on the 1-6 RDQ scale), while all other items within the 24-item questionnaire have scores of less than 3; Family = family reasons, Pregnancy = pregnancy, Child care = incompatibility of degree course and child care, Illness = illness, Abroad = degree course abroad or internship abroad, Incompatibility = incompatibility of degree course and employment, Practical work = wanted to do practical work, Job offer = interesting job offer, Fail exam = did not pass examinations, Fin prob. = financial problems, Earn money = wanted/had to earn money as quickly as possible, Interest in sub. = loss of interest in subject, Perform req. = unable to meet performance requirements for the degree course, Practical relev. = Lack of professional and practical relevance of the studies, Study req. = course requirements too high, Lack opportun. = poor job opportunities in my field of study, False expec. = Wrong expectations in relation to the studies, Anonymity = Anonymity in the university.

#### **5.4 Research question 3: Description of the classes by RDQ means scores**

Analysis of the data revealed that the best solution related to a 12-class model (Fraley & Raftery, 1998; see Supplements B, C1, C2). Comparisons amongst the 12 classes yielded three sub-groups based on class size. Three large classes with between 200-350 cases, four medium classes with between 90-110 cases and five

small classes with between 20-40 cases were classified. Class descriptions are ordered according to class sizes and based on items with class mean scores of above 3.5 (see Supplement F). In the following, we will characterize these classes:

### (1) Large classes (200-350 cases)

Class 1 (low interest/performance problems;  $n = 328$ ; 24 %) is the largest class and presents a pattern of both performance problems and a low interest in studies, specifically *failing the exam* ( $M = 4.09$ ; for all means see Supplement F), *false expectations* ( $M = 3.84$ ), *low interest in subject* ( $M = 3.75$ ) and *study requirements are high* ( $M = 3.61$ ). In addition, Class 1 contains the largest number of males in the sample ( $n = 197$ , 34 % of the male sample) despite only comprising of 61 % males within the class. This class also has one of the youngest means for age at dropout of 25 years. Notably, Class 1 had the highest proportion of mathematics and natural sciences students, and engineering students (see Supplement D).

Class 4 (job needed;  $n = 249$ ; 19 %) is the second largest class and indicates higher scores on the *financial problems* ( $M = 3.29$ ), *incompatibility of degree course and employment* ( $M = 3.09$ ) and *interesting job offer* ( $M = 3.18$ ). Class 4 has 51 % females and an average age at dropout of 27 years.

Class 7 (low interest;  $n = 221$ ; 16 % of the sample) is the third largest class. The main reasons include *low interest in subject* ( $M = 3.78$ ; for all means please see Supplement F), *desire for practical work* ( $M = 3.47$ ) and *false expectations* ( $M = 3.35$ ). Furthermore, this class comprises of 64 % females and an average age at dropout of 25 years.

### (2) Medium-sized classes (90-110 cases)

Class 2 (unsure;  $n = 103$ ; 8 %) provides no clear reason for dropout according to RDQ mean scores. Overall, class 2 has the lowest scoring main reason for dropout and the lowest mean score for all reasons for dropout combined. The most prominent items are *low interest in subject* ( $M = 3.05$ ), *desire for practical work* ( $M = 2.93$ ) and *false expectations* ( $M = 2.80$ ). Other reasons for dropout include *degree course abroad or internship abroad* ( $M = 2.72$ ) and reasons related to performance problems (mean ranging from 2.46 to 2.54). This class consists of 66 % females and a mean age at dropout of 25 years.

Class 3 (ambiguous;  $n = 96$ ; 7 %) provides a broad range of possible reasons such as *desire for practical work* ( $M = 3.94$ ), *false expectations* ( $M = 3.58$ ), *low interest in subject* ( $M = 3.52$ ), but also indicate that *financial problems* ( $M = 3.55$ ) and *illness* ( $M = 3.50$ ) played a considerable role in the decision. Class 3 consists of 65 % females and an average age at dropout of 27 years.

Class 11 (illness;  $n = 102$ ; 8 %) reported that *illness* ( $M = 3.32$ ) was the primary problem and, to a moderate degree, also performance problems such as *failing the exam* ( $M = 2.84$ ) are the main reasons for dropout. All other reasons for dropout play almost no role in this class' decision to drop out ( $M < 2.00$ ). Class 11 consists of 66 % females with a mean age at dropout of 25 years.

Class 12 (family only;  $n = 93$ ; 7 %), similarly to class six, reported *pregnancy* ( $M = 4.09$ ), *family reasons* ( $M = 4.33$ ) and *child care* ( $M = 3.77$ ), while all remaining reasons play almost no role in this class' decision to drop out ( $M < 1.80$ ). In addition, 15 % of this class were recorded as dropping out more than once, which is the highest percentage of any class. This class comprises of 81 % females and an average age at dropout of 30 years.

### *(3) Small classes (<40 cases)*

Class 5 (challenged;  $n = 27$ ; 2 %) reported several major reasons for dropout primarily related to *family issues*, specifically *child care* ( $M = 4.96$ ) and *family reasons* ( $M = 4.44$ ). In addition, they reported *incompatibility of degree course and employment* ( $M = 4.33$ ), *financial problems* ( $M = 4.00$ ) and *lack of support from lecturers* ( $M = 3.96$ ). Class 5 has 66 % females and an average age at dropout of 30 years.

Class 6 (family plus;  $n = 24$ ; 1 %) relates specifically to *family reasons* ( $M = 4.33$ ), *pregnancy* ( $M = 4.29$ ) and *child care* ( $M = 3.25$ ) as well as *financial problems* ( $M = 3.04$ ), whilst other reasons for dropout play a lesser role in the decision, i.e. other reasons for dropout have means below 3.00. This is the smallest class with 96 % females and an average age at dropout of 30 years.

Class 8 (going abroad;  $n = 26$ ; 2 %) scored the highest on *degree course abroad or internship abroad* ( $M = 5.08$ ) and *desire for practical work* ( $M = 3.73$ ). This small class comprises of 65 % females and an average age at dropout of 26 years.

Class 9 (study requirements;  $n = 37$ ; 3 %), presents multiple reasons for dropout including *too much exam material* ( $M = 3.62$ ), *desire for practical work* ( $M = 3.46$ ), and *illness* ( $M = 3.57$ ). This class comprises of 70 % females and an average age at dropout of 32 years.

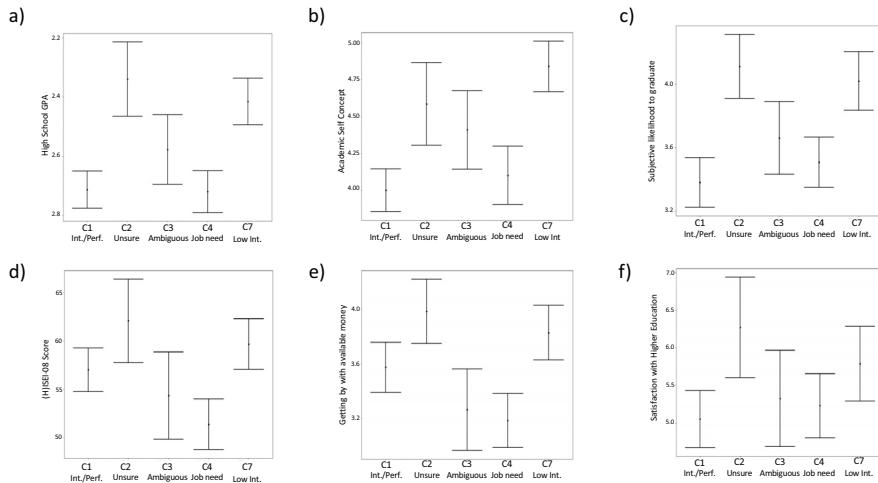
Class 10 (desire for alternative employment;  $n = 38$ ; 3 %) students indicated the highest scores for the *desire for practical work* ( $M = 4.20$ ) and *interesting job offer* ( $M = 3.79$ ).

## **5.5 Research question 4: Class descriptions according to contextual variables**

Research question 4 focuses on characterizing these classes using contextual variables associated with dropout. For these analyses, we only focus on five of the large and medium classes but exclude Class 11 (illness) and Class 12 (family only) because their reasons for dropout are less predictable.

Figure 2 shows the comparison of the five classes' scores on six contextual variables, some of which relate to performance and socio-economic factors.

**Figure 2:** Means and 95 % confidence intervals for students high school GPA (a), academic self-concept at wave 2 (b), their subjective evaluation of their likelihood to graduate (c), their estimation how far they're getting by with their money (e), and their satisfaction with higher education (f) for the Class 1 (low interest/performance problems), Class 2 (unsure), Class 3 (ambiguous), Class 4 (job needed), and Class 7 (low interest). All values can be found in Supplement H1-H6; Means and standard deviations for all classes could be found in Supplement G.



Notably, Class 1 (low interest/performance problems) as well as Class 4 (job needed) show significantly lower values regarding their high school GPA, their academic self-concept, and their subjective likelihood to graduate than Class 2 (unsure) and Class 7 (low interest), while Class 3 (ambiguous) is in between the former classes with partially overlapping confidence intervals. Socio-economic factors regarding parents' status (HISEI) and financial challenges showed comparable results between Class 2 (unsure) and Class 7 (low interest), as both classes' parents had a higher HISEI and less issues with money, while students of Class 3 (ambiguous) and Class 4 (job needed) indicated lower scores on these socio-economic factors. Class 1 (low interest/performance problems) presented moderate scores on the socio-economic factors which situates them between these four classes. Results regarding students' satisfaction with higher education are less clear, as Class 2 (unsure) shows the highest satisfaction, while Class 1 (low interest/performance problems) and Class 4 (job needed) report a noticeably lower satisfaction with higher education.

## 6. Discussion

The aim of the study was to shed light on the dropout phenomenon using a person-centered approach to reveal latent classes of dropout students based on their reasons for dropout. Therefore, the paper first presented information about drop-outs and reasons for dropout longitudinally before looking into the issue of multi-causality of dropouts. The latent class analysis revealed 12 classes and the follow up analysis disclosed that they further distinguish with respect to contextual variables. We will now discuss the respective results in detail:

### 6.1 Longitudinal effects

Regarding longitudinal effects, we could see that most students drop out very early and their numbers constantly decline over time. Of the four reasons for early dropout, two of them relate to study conditions, namely *anonymity in the university* and *over-crowdedness*. These reasons may occur due to lower social integration with their peers which weakens their institutional and goal commitments, and puts them at greater risk to dropout (Tinto, 1975). A further reason, *false expectations*, was highlighted by Klein and Stocké (2016) as a potential indicator of information deficits that students have prior to entering their studies, which to a lesser extent also relates to *study requirements*. Class 1 (low interest/performance problems) could be a prototype class for early dropout, as it embodies more early dropout reasons than any other classes. Positive social participation with faculty or student counselling could be utilized to produce positive effects on students' intention to dropout and prevent withdrawal behavior from the educational system.

The reasons which significantly impacted later dropout included *earn money quickly*, *incompatibility of degree course and employment*, *family reasons*, and *illness*. The two former reasons encompass an increased financial burden both in the immediate sense, that money needed to be earned quickly, as well as the future employment opportunities which the current studies may not afford them. Dropout in this sense is potentially related to socio-economic factors within the family, which ties into another late impact reason, namely *family reasons*. Heublein and Wolter (2011) cited that gainful employment alongside studies may lead students to dropout early. However, students who did not acquire gainful employment parallel to their studies may redirect their focus to finding full-time work to alleviate some of their financial burdens. Family and employment-related dropout as described by Class 6 (family plus), Class 10 (desire for alternative employment), or Class 12 (family only) could represent prototypes of late dropout. Finally, *illness* is a crisis which cannot be predicted (Bean, 1985) but the results suggest that it sets in at the later stages of studies.

## 6.2 Multi-dimensional reasons for dropout

Research questions 2, 3, and 4 applied a person-centered approach that analyzed reasons for dropouts among sub-groups, leading to 12 latent classes of dropout students. RQ2 illustrated that 53 cases (4 %) represent a mono-causal dropout which implies a marginal significance. This trend also applies for the two- and three-reason combinations which represents a further 8 %.

The LCA provided three classes that are possibly distinguishable by only a single dropout factor, specifically Class 8 (going abroad), Class 11 (illness) and Class 12 (family only), who together represent 17 % of the sample. However, as mentioned previously, both Class 11 and Class 12 had a low intention to dropout which theoretically made their dropout far less predictable (Bean, 1985). These results support the need for a multi-dimensional model to understand the majority of the dropout phenomenon (Robbins et al., 2004) and provides evidence for following a multi-causal approach such as an LCA (Muthén & Muthén, 2000).

Focusing on the majority of the sample, the classes reveal that the three largest classes each emphasize a loss of interest in their subject and false study expectations as their major reasons for dropout. This corresponds to previous findings related to information deficits that students experienced prior to the start of studies, which then leads to false expectations (Klein & Stocké, 2016; Brandstätter, Grillich, & Farthofer, 2006). While Class 7 (low interest), Class 1 (low interest/performance problems) and Class 4 (job needed) each indicated a loss of interest in their subject and false study expectations as major reasons for dropout, Class 1 also indicated performance problems while Class 4 indicated finance problems. These three classes together account for 60 % of the total sample and raise questions about how these patterns develop. The issue may stem from a lack of quality career counselling or inappropriate choice of study courses, which are causes that develop in the pre-study phase according to Heublein et al. (2010). Alternatively, these students may have chosen to study their subject for reasons other than being interested in the contents of their studies, which corresponds to previous findings by Heublein et al. (2017). Although these three classes appear to have lost interest in their study subject, they did not indicate a loss of interest in later profession as playing as great a role in their decision to drop out. This suggests a disconnect between their experiences within their studies and their future career roles. These students may decide to vocationally re-orientate to more practical work, which is both pulled by the attractiveness of practical and gainful employment and pushed by the failures experienced within the higher educational system (Griesbach et al., 1998). Heublein et al. (2017) noted how students desired to move away from the primarily theory-based higher education system to do more practical and relevant work. Hence, these students may likely switch their study field, or join the labor market without completing their studies.

Performance indicators did not manifest a stand-alone dropout class. Other than a very small sample of single-reason dropouts in RQ1 (see Table 2), who only indicated failing an exam, performance indicators alone do not explain why drop-

out occurs (Voelkle & Sander, 2008). The two largest classes indicated how performance problems combined with either low interest in the study subject expressed by Class 1 (low interest/performance problems), or socio-economic burdens expressed by Class 4 (job needed). Such differences could be verified by RQ4 which uncovered that both Class 1 and Class 4 reported a less beneficial performance background and, additionally, Class 4 had a less beneficial socio-economic background when compared to that of Class 7 (no interest), who had both a superior performance and socio-economic background.

A different pattern can be seen by Class 11 (illness;  $n = 102$ ; 8 %) and Class 12 (family only;  $n = 93$ ; 7 %) that present a remarkably similar profile according to contextual variables investigated. Both indicate a low intention to dropout reflected by a high subjective likelihood to complete a degree. They both reported high academic self-concepts and relatively average high school GPA compared to other classes. Most importantly, they represent classes whose reasons for dropout are outside the influence of university institutions (Isphording & Wozny, 2017; Klein & Stocké, 2016) and cannot be predicted as the nature of their reasons for dropouts are in response to a personal crisis (Bean, 1985). Their decision to drop out are not a failure on either the part of the institution nor the student. These students had no intention to dropout but rather may have less perceived control over their academic success (Respondek, Seufert, Stupnisky, & Nett, 2017). Reducing obstacles for students with family duties and providing support at the appropriate times may be a promising way to alleviate the challenges facing Class 12.

### 6.3 Conclusion

The results of this study broadly support Heublein et al. (2017) synopsis that the available dropout models were limited in their ability to only explain certain groups of students, and that various reasons for dropout are needed to encompass the full scope of the dropout phenomenon. The introduction of latent class analysis as a method to identify various subgroups of dropouts can overcome this limitation in a practical and generalizable manner. Using LCA, as opposed to factor-based methods, has solved some issues but also illuminated further complexity within the dropout phenomenon. Previous dropout models needed to be separated by demographics such as gender (Spady, 1971), or special groups such as non-traditional students (Bean & Metzner, 1985) to grasp at least some homogeneity within dropout subgroups. Within our study the LCA approach has the advantage of being able to classify subgroups based on their primary reasons for dropout as well as demographic differences (Collins & Lanza, 2009). Using LCA to allow multiple predictors to interplay amongst classes appears to provide more easily identifiable and generalizable groups with clearer interpretations within other educational purposes (see Ertl, Luttenberger, & Paechter, 2014; Luttenberger, Aptarashvili, Ertl, Ederer, & Paechter, 2014; Quirk, Grimm, Furlong, Nylund-Gibson, & Swami, 2016).

The results of this study go further than the well-established models espoused by Spady (1971) and Tinto (1975). In our analysis, only one class (Class 5: challenged), representing 2 % of the sample, distinguished a lack of lecturer support as a reason for dropout. In addition, other items related to social integration within the institution, such as *overcrowded lectures* and *anonymity in the university* played only a minor role in distinguishing any of the classes, although they did present a significant impact on early dropout amongst the sample as a whole (see Supplement E and Table 1). This may be due to the structure of German universities that do not combine academic and social life on campus, which is often typical of North American campus universities (Heublein & Wolter, 2011). This situation also provides an alternative perspective to the universal role of institutional commitment as espoused by Tinto (1975).

A peculiarity within our study, which corresponded to previous literature, was the proportion of classes who indicated a *lack of interest in the study subject* and *false study expectations* (Heublein et al., 2017). This finding was cited as evidence for an information deficit at the start of studies (Klein & Stocké, 2016). Whether this is an artifact of the relatively recent reform within the German education system or of poorly counselled students is unclear.

## 6.4 Limitations

The sample of dropouts who answered the reason for dropout questionnaire does not exhaust the total number of dropouts from this panel study. As is customary within dropout research, many of the students left the panel study without any information about why they discontinued their studies. For this reason, there may be individuals who could have produced further latent classes. The possibility of further other classes is supported by existence of the “unsure” class, which may be an indication that the reason for dropout which specifically affected them was not mentioned within the list of 24 items.

Lastly, a dropout can become a non-dropout once they re-enter the educational system at any time, either continuing studying the original subject or switching to a new one. Tracking and understanding why students depart from an institution therefore remains a difficult yet important task within dropout research.

## 7. Significance

Identifying students who are at risk of prematurely ending their studies and providing them with the appropriate intervention has been the goal of research into the dropout phenomenon. A clearer understanding of not only the factors which lead to dropouts but the interplay between these dropout factors can provide val-

able information to keep institutions accountable to their students while pointing out to areas for effective management of at-risk students.

This study both built on previous models of student dropout, specifically Spady (1971), Tinto (1975), Bean (1980) and for the German context (Heublein et al., 2010; Heublein et al., 2017), while also advancing on this research by identifying latent classes of students dropping, and thereby revealing two essential aspects. Firstly, dropout research from the North American context has only limited power to explain dropout phenomena in Germany. Constructs like social integration that were seen as essential by Spady (1971) and Tinto (1975) may only play a marginal role in distinguishing between dropout classes. The second aspect relates to the increased value in contrasting between variable-centered and person-centered approaches. The patterns shown by the latent class analysis allow variables to independently impact each subgroup in a different way, which can reveal classes with opposing impacts from the same variable. For example, classes of students characterized by low interests in their study subject and false study expectation also displayed an inclination for practical work despite the fact that these three items belong to two separate factors in previous analyses (Heublein et al., 2010; Heublein et al., 2017). In addition, classes with a low interest in their study subject did not strongly indicate a low interest in the profession, despite these two items representing one factor in these studies.

While variable-centered approaches give indications of the overarching effects of dropout factors (Bean, 1980; Heublein et al. 2010; Heublein et al., 2017; Spady. 1971; Tinto, 1975), this study was able to quantitatively identify specific student profiles that provide a starting point for designing targeted and practical interventions. Confirming the existence of these classes using models with specific and directed combinations of dropout variables, such as performance and socio-economic background factors, could further disclose the development of specific subgroups of dropout students.

## Acknowledgments

This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort First-Year Students, doi:10.5157/NEPS:SC5:10.0.0. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

## References

- Baudry, J. P., Raftery, A. E., Celeux, G., Lo, K., & Gottardo, R. (2010). Combining mixture components for clustering. *Journal of Computational and Graphical Statistics*, 19(2), 332–353. doi:10.1198/jcgs.2010.08111

- Bayerischer Oberster Rechnungshof. (2019). *Jahresbericht 2019* [2019 Annual report]. <https://www.orh.bayern.de/berichte/jahresberichte/aktuell/jahresbericht-2019.html>
- Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education*, 12(2), 155–187. doi:10.1007/bf00976194
- Bean, J. P. (1985). Interaction effects based on class level in an explanatory model of college student dropout syndrome. *American Educational Research Journal*, 22(1), 35–64. doi: 10.3102/00028312022001035
- Bean, J. P., & Metzner, B. S. (1985). A conceptual model of non-traditional undergraduate student attrition. *Review of Educational Research*, 55(4), 485–540. doi:10.2307/1170245
- Berger, J. B., & Braxton, J. M. (1998). Revising Tinto's interactionalist theory of student departure through theory elaboration: Examining the role of organizational attributes in the persistence process. *Research in Higher Education*, 39(2), 103–118. doi: 10.1023/A:1018760513769
- Berger, J. B., Ramirez, G. B., & Lyons, S. (2012). Past to present: A historical look at retention. In A. Seidman (Ed.), *College student retention: Formula for student success* (2nd ed., pp. 7–34). New York, NY: Rowman & Littlefield. doi: 10.1353/rhe.2006.0017
- Blossfeld, H.-P., Roßbach, H.-G. & Maurice, J. von (Eds.). (2011). *Education as a life-long process. The German National Educational Panel Study (NEPS)*. Wiesbaden: VS.
- Brandstätter, H., Grillich, L., & Farthofer, A. (2006). Prognose des Studienabbruchs [Prognosis of dropout]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 38(3), 121–131. doi: 10.1026/0049-8637.38.3.121
- Brossardt, B. (2012). *Studdienabbruch und Fachkräftemangel vermeiden: Das Engagement von bayme vbm vbw* [Avoiding a student drop-out and a lack of skilled workers: Engagement from bayme vbm vbw's commitment]. [Press release]. Staatsministerium für Wissenschaft, Forschung und Kunst. <https://docplayer.org/38432882-Studienabbruch-und-fachkraeftemangel-vermeiden-das-engagement-von-bayme-vbm-vbw.html>
- Cabrera, A. F., Nora, A., & Castaneda, M. B. (1993). College persistence: Structural equations modelling test of an integrated model of student retention. *The Journal of Higher Education*, 64(2), 123–139. doi:10.2307/2960026
- Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *The Journal of Educational Research*, 90(5), 269–277. doi:10.1080/00220671.1997.10544583
- Collins, L. M., & Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Wiley-Blackwell. doi:10.1111/j.1751-5823.2010.00122\_5.x
- Destatis. (Ed.). (2018). *Studierende an Hochschulen – Fächersystematik* [Students at universities – Classification of subjects]. [https://www.destatis.de/DE/Methoden/Klassifikationen/BildungKultur/StudentenPruefungsstatistik.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Methoden/Klassifikationen/BildungKultur/StudentenPruefungsstatistik.pdf?__blob=publicationFile)
- Dickhäuser, O., Schöne, C., Spinath, B., & Stiensmeier-Pelster, J. (2002). Die Skalen zum akademischen Selbstkonzept [The scales for the academic self-concept]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 23(4), 393–405. doi: 10.1024/0170-1789.23.4.393
- Durkheim, E. (1997). *Suicide: A study in sociology* (J. Spaulding, Trans.; Ed.). Free Press. (Original work published 1897).
- Ertl, B., & Hartmann, F. G. (2019). The interest profiles and interest congruence of male and female students in STEM and non-STEM fields. *Frontiers in Psychology*, 10(897). doi:10.3389/fpsyg.2019.00897

- Ertl, B., Luttenberger, S., & Paechter, M. (2014). Stereotype als Einflussfaktoren auf die Motivation und die Einschätzung der eigenen Fähigkeiten bei Studentinnen in MINT-Fächern [Stereotypes as factors influencing motivation and the assessment of one's own abilities among female students in MINT subjects]. *Gruppendynamik und Organisationsberatung*, 45(4), 419–440. doi:10.1007/s11612-014-0261-3
- Esser, H., & Stocké, V. (n.d.). Forschungsprojekt “Bildungsaspirationen, Bezugsgruppen und Bildungsentscheidungen” (Mannheimer Bildungspanel). Kinderbefragung, 6. Welle.
- FDZ – LIffBi. (2018). Codebook. NEPS Starting Cohort 5 – First-Year Students: From Higher Education to the Labor Market [Scientific Use File Version 10.0.0]. Research Data Center LIffBi.
- Field, A., Miles, J., & Field, Z. (2019). *Discovering statistics using SPSS* (5th ed.). Los Angeles: Sage.
- Fraley, C., & Raftery, A. E. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis. *The Computer Journal*, 41(8), 579–588. doi:10.1093/comjnl/41.8.578
- Fraley, C., Raftery, A. E., Scrucca, L., Murphy, T. B., & Fob, M. (2018). Package mclust (Version 5.4.2). *The R Journal*.
- Galla, B. M., Shulman, E. P., Plummer, B. D., Gardner, M., Hutt, S. J., Goyer, J. P., & Duckworth, A. L. (2019). Why high school grades are better predictors of on-time college graduation than are admissions test scores: The roles of self-regulation and cognitive ability. *American Educational Research Journal*, 56(6), 2077–2115. doi:10.3102/0002831219843292
- Ganzeboom, H. B. G. (2010). *A new International Socio-Economic Index (ISEI) of occupational status for the International Standard Classification of Occupation 2008 (ISCO-08) constructed with data from the ISSP 2002–2007; With an analysis of quality of occupational measurement in ISSP* [Paper presentation]. Annual Conference of International Social Survey Programme, Lisbon, Portugal.
- Ganzeboom, H. B. G., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56. doi:10.1016/0049-089X(92)90017-B
- Griesbach, H., Lewin, K., Heublein, U., & Sommer, D. (1998). Studienabbruch: Typologie und Möglichkeiten der Abbruchquotenbestimmung [Termination of studies: Typology and possibilities of determining dropout rates]. *HIS: Kurzinformation*.
- Haughton, D., Legrand, P., & Woolford, S. (2009). Review of three latent class cluster analysis packages: Latent Gold, poLCA, and MCLUST. *The American Statistician*, 63(1), 81–91. doi:10.1198/tast.2009.0016
- Heublein, U. (2014). Student drop-out from German higher education institutions. *European Journal of Education*, 49(4), 497–513. doi:10.1111/ejed.12097
- Heublein, U., Ebert, J., Hutzsch, C., Isleib, S., König, R., Richter, J., & Woisch, A. (2017). Zwischen Studienerwartungen und Studienwirklichkeit: der Studienabbrcherinnen und Studienabbrecher und Entwicklung der Studienabbruchquote an deutschen Hochschulen [Between study expectations and reality: dropouts and development of dropout rates in German universities]. DZHW: Forum Hochschule.
- Heublein, U., Hutzsch, C., Schreiber, J., Sommer, D., & Besuch, G. (2010). Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen: Ergebnisse einer bundesweiten Befragung von Exmatrikulierten des Studienjahres 2007/08 [Causes of dropout in Bachelor and conventional courses: Results of a nationwide survey of exmatriculated students in the academic year 2007/08]. HIS: Forum Hochschule.
- Heublein, U., & Schmelzer, R. (2018). Die Entwicklung der Studienabbruchquoten an den deutschen Hochschulen: Berechnungen auf Basis des Absolventenjahrgangs

- 2016 [The development of dropout rates in German universities: calculations based on the 2016 graduate class]. DZHW-Projektbericht.
- Heublein, U., & Wolter, A. (2011). Studienabbruch in Deutschland. Definition, Häufigkeit, Ursachen, Maßnahmen [Dropped out of university in Germany. Definition, frequency, causes, measures]. *Zeitschrift für Pädagogik*, 57(2), 214–236.
- Horstschräer, J., & Spietsma, M. (2015). The effects of the introduction of Bachelor degrees on college enrollment and dropout rates. *Education Economics*, 23(3), 296–317. doi: 10.1080/09645292.2013.823908
- Ishitani, T. T. (2006). Studying attrition and degree completion behavior among first-generation college students in the united states. *The Journal of Higher Education*, 77(5), 861–885. doi: 10.1080/00221546.2006.11778947
- Isphording, I., & Qendrai, P. (2017). *Gender differences in student dropout in STEM* (No. 87). Institute of Labor Economics. <https://www.iza.org/publications/r/197/gender-differences-in-student-dropout-in-stem>
- Isphording, I., & Wozny, F. (2017). *Ursachen des Studienabbruchs – eine Analyse des Nationalen Bildungspanels* [Causes of dropping out of university – an analysis by the National Education Panel]. (No. 82). Institute of Labor Economics. [http://ftp.iza.org/report\\_pdffs/iza\\_report\\_82.pdf](http://ftp.iza.org/report_pdffs/iza_report_82.pdf)
- Keeling, R. (2006). The bologna process and the Lisbon research agenda: The European commission's expanding role in higher education discourse. *European Journal of Education*, 41(2), 203–233. doi:10.1111/j.1465-3435.2006.00256.x
- Klein, D., & Stocké, V. (2016). Studienabbruchquoten als Evaluationskriterium und Steuerungsinstrument der Qualitätssicherung im Hochschulbereich. In D. Großmann & T. Wolbring (Eds.), *Evaluation von Studium und Lehre: Grundlagen, methodische Herausforderungen und Lösungsansätze* (pp. 323–365). Wiesbaden: Springer Fachmedien.
- Kruppe, T., & Baumann, M. (2019). Weiterbildungsbeteiligung, formale Qualifikation, Kompetenzausstattung und Persönlichkeitsmerkmale: IAB-Forschungsbericht [Participation in continuing training, formal qualifications, skills and personality traits: IAB Research Report]. Institute für Arbeitsmarkt- und Berufsforschung.
- Luttenberger, S., Aptarashvili, I., Ertl, B., Ederer, E., & Paechter, M. (2014). Niedrige Übereinstimmung zwischen Interessen und Berufswunsch – Ein bislang vernachlässigtes Risiko in der Berufsorientierung Jugendlicher [Low correspondence between interests and career aspiration – A previously neglected risk in the professional orientation of young people]. *Gruppendynamik und Organisationsberatung*, 45(4), 359–377. doi:10.1007/s11612-014-0252-4
- Ma, Y., & Cragg, K. M. (2013). So close, yet so far away: Early vs. late dropouts. *Journal of College Student Retention: Research, Theory & Practice*, 14(4), 533–548. doi: 10.2190/CS.14.4.f
- Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, 24(6), 882–891. doi: 10.1111/j.1530-0277.2000.tb02070.x
- OECD – Organization for Economic Cooperation and Development. (2008). *Education at a Glance 2008: OECD Indicators*. <http://www.oecd.org/dataoecd/8/24/41271802.pdf>.
- Price, J. L. (1977). *The Study of Turnover*. University of Iowa Press.
- Quirk, M., Grimm, R., Furlong, M. J., Nylund-Gibson, K., & Swami, S. (2016). The association of Latino children's kindergarten school readiness profiles with Grade 2–5 literacy achievement trajectories. *Journal of Educational Psychology*, 108(6), 814–829. doi:10.1037/edu0000087
- R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>

- Respondek, L., Seufert, T., Stupnisky, R., & Nett, U. E. (2017). Perceived academic control and academic emotions predict undergraduate university student success: Examining effects on dropout intention and achievement. *Frontiers in Psychology*, 8(243). doi:10.3389/fpsyg.2017.00243
- Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130(2), 261–288. doi:10.1037/0033-2909.130.2.261
- Seidman, A. (2012). Taking action: a retention formula and model for a student success. In A. Seidman (Ed.), *College student retention: Formula for student success* (2nd ed., pp. 267–284). New York, NY: Rowman & Littlefield.
- Spady, W. G. (1971). Dropouts from higher education: Toward an empirical model. *Interchange*, 2(3), 38–62. doi:10.1007/BF02282469
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89–125. doi:10.2307/1170024
- Tinto, V. (2006). Research and practice of student retention: What next? *Journal of College Student Retention: Research, Theory and Practice*, 8(1), 1–19. doi:10.2190/CoC4-EFT9-EG7W-PWP4
- Trautwein, U., Köller, O., & Watermann, R. (2004). Transformation des Sekundarschulsystems und akademische Karrieren—Zusammenfassung, Diskussion und ein Ausblick [Transformation of the secondary school system and academic careers— Summary, discussion and outlook]. In *Wege zur Hochschulreife in Baden-Württemberg* (pp. 451–471). VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-322-80906-3\_13
- Voelkle, M. C., & Sander, N. (2008). University dropout: A structural equation approach to discrete-time survival analysis. *Journal of Individual Differences*, 29(3), 134–147. doi:10.1027/1614-0001.29.3.134
- Xianglei, C., & Mathew, S. (2013). *STEM attrition: College students' paths into and out of STEM fields: Statistical Analysis Report*. National Center for Education Statistics.

## Appendix

### Supplement A1: Absolute frequency of dropout students per wave

Wave	Frequency
3	551
5	392
7	202
9	117
10	82

### Supplement A2: Descriptive statistics of dropout students' intention to dropout per wave

Wave	Dropout students		Whole sample	
	M	SD	M	SD
2	1.96	0.83	1.49	0.59
4	1.84	0.81	1.45	0.53
6	1.72	0.77	1.45	0.53
8	1.61	0.70	1.43	0.53

*Note.* Dropout students refer to all students who answered the RDQ; The whole sample refers to all other cases who answered the Intention to Dropout Questionnaire within the NEPS dataset.

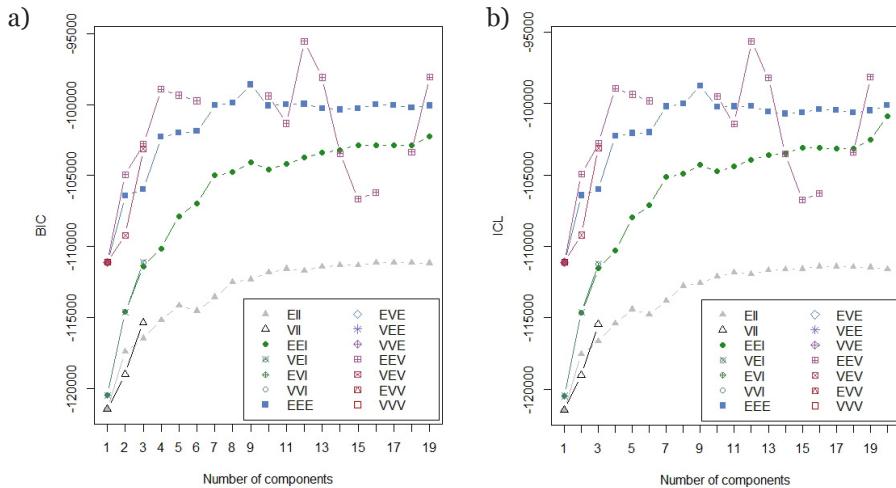
**Supplement B:** Quality indicators for solutions with the different classes: Summary of model fit indices among different classes.

K	df	Model	BIC	ICL	LL
1	324	Multiple Models <sup>a</sup>	-111107.30	-111107.30	-54386.68
2	625	EEV	-104936.30	-104937.80	-50217.11
3	926	EEV	-102772.30	-102786.90	-48050.98
4	1227	EEV	-98937.59	-98947.44	-45049.51
5	1528	EEV	-99340.98	-99361.60	-44167.09
6	1829	EEV	-99741.00	-99811.60	-43282.99
7	NA	NA	NA	NA	NA
8	NA	NA	NA	NA	NA
9	NA	NA	NA	NA	NA
10	3033	EEV	-99405.42	-99485.13	-38778.74
11	3334	EEE	-101334.60	-101445.20	-38659.20
12	3635	EEV	<b>-95543.74</b>	<b>-95638.94</b>	-34679.68
13	3936	EEV	-98088.78	-98193.24	-34868.09
14	4237	EEE	-103463.00	-103539.90	-36471.08
15	4538	EEE	-106656.80	-106742.70	-36983.88
16	4839	EEE	-106184.30	-106270.70	-35663.52
17	NA	NA	NA	NA	NA
18	5441	EEE	-103342.50	-103417.10	-32074.38
19	5742	EEV	-98070.36	-98126.43	<b>-28354.20</b>

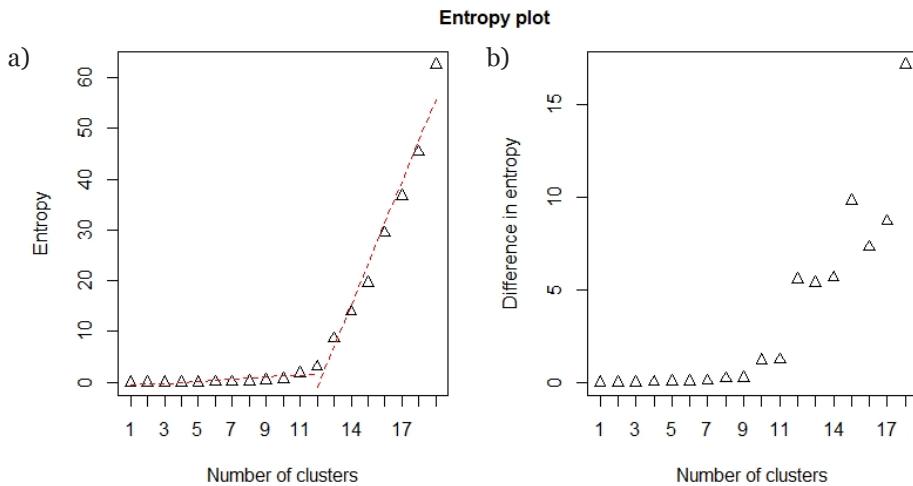
Note. K = number of classes; BIC = Bayesian Information Criterion; ICL = Integrated Complete-data Likelihood; LL = log-likelihood; Bolded values indicated “best” fit for each respective statistic. EEV refers to ellipsoidal, equal volume and equal shape. EEE is ellipsoidal, equal volume, shape, and orientation. NA represents that model does not provide solution according to the classes numbers we requested.

<sup>a</sup>Multiple models match with both model EEV and EEE.

**Supplement C1:** Chart based on BIC (a) and ICL (b) criterion for the best models among different number of classes.



**Supplement C2:** Entropy charts for solutions with a different number of classes



*Note.* (a) Entropy values for the K-cluster combined solutions. (b) Differenced between successive entropy values. An elbow in the plot (a) suggests 12 clusters.

**Supplement D:** Characteristics about classes size, gender and age and study subject at dropout in profiles. Absolute numbers and line percentages (in brackets) unless noted otherwise

	Classes										Total [% of total]		
	1	2	3	4	5	6	7	8	9	10			
Classes size (%)	328 (24)	103 (8)	96 (7)	249 (19)	27 (2)	24 (1)	221 (16)	26 (2)	37 (3)	38 (3)	102 (8)	93 (7)	1344 [100]
Gender (%)													
Male	197 (34)	34 (6)	34 (6)	122 (21)	9 (1)	80 (14)	9 (1)	11 (2)	21 (4)	42 (7)	17 (3)	577 [43]	
Female	131 (17)	69 (9)	62 (8)	127 (17)	18 (2)	23 (3)	141 (18)	17 (2)	26 (4)	17 (2)	60 (8)	76 (10)	767 [57]
Age at Dropout													
Mean (SD)	25 (4)	25 (6)	27 (5)	27 (5)	30 (7)	30 (7)	25 (4)	26 (5)	32 (9)	30 (8)	30 (9)	30 (7)	1339 [99]
Median	24.50	24	25	26	28	28	28.50	24	24	28	26	26	28
Range	21-59	21-62	22-47	21-53	23-48	22-45	21-48	22-49	23-52	23-52	22-67	22-56	
Mean Wave for Dropout	4.76	5.10	5.51	5.31	5.15	6.13	4.75	5.35	5.46	5.66	5.45	5.53	
Study Domain (%)													
Linguistic and cultural studies	44 (13)	38 (11)	25 (7)	59 (17)	11 (3)	6 (1)	69 (20)	11 (3)	13 (4)	13 (4)	25 (7)	33 (10)	347 [26]
Law, economics and social science	54 (18)	16 (5)	20 (7)	56 (18)	7 (2)	8 (3)	61 (20)	7 (2)	14 (5)	12 (4)	25 (8)	26 (8)	306 [23]
Mathematics, natural sciences	111 (33)	25 (7)	25 (7)	65 (19)	6 (2)	3 (1)	46 (14)	4 (1)	3 (1)	7 (2)	29 (9)	14 (4)	338 [25]
Engineering	101 (44)	13 (6)	15 (7)	52 (24)	.	2 (1)	23 (10)	2 (1)	7 (2)	3 (1)	6 (2)	6 (2)	230 [17]
Human medicine/health sciences	4 (8)	4 (8)	3 (5)	4 (8)	.	4 (8)	10 (19)	.	.	3 (5)	12 (22)	9 (17)	53 [4]
Veterinary medicine	.	1 (50)	.	.	.	.	.	.	.	.	1 (50)	2 [1]	
Agricultural-, forest- and nutrition sciences	10 (32)	2 (6)	3 (10)	6 (20)	1 (3)	.	3 (10)	1 (3)	.	3 (10)	2 (6)	31 [2]	
Sports	2 (15)	1 (8)	2 (15)	1 (8)	2 (15)	.	4 (31)	1 (8)	.	.	.	13 [1]	
Arts, aesthetics	.	3 (16)	3 (16)	5 (26)	.	1 (5)	4 (21)	.	.	1 (5)	2 (11)	19 [1]	

## Supplement E: Means of RDQ items per wave

Wave	Perform.req.	Study req.	Exam mat.	Fail exam	Lecture supp.	Study org.	Overcrowded	Anonymity	Practical work	Practical relev.	Interest in sub.	Interest in prof.	False expec.	Earn money	Pregnancy	Child care	Fam reasons	Personal suit.	Incompatibility	Job offer	Lack opportun.	Illness	Abroad	
3	2.67	2.96	2.69	2.72	2.72	2.26	1.91	3.25	2.81	3.41	2.43	3.50	2.27	1.79	1.23	1.39	1.87	2.64	2.14	2.21	2.21	2.03	1.46	1.17
5	2.56	2.77	2.71	2.77	2.44	2.53	2.11	1.67	3.25	2.70	3.05	2.09	2.91	2.21	1.78	1.33	1.46	2.03	2.24	2.11	2.29	1.82	1.60	1.41
7	2.36	2.59	2.62	2.78	2.44	2.68	1.93	1.70	2.96	2.63	2.79	2.03	2.76	2.19	1.99	1.23	1.38	1.97	2.16	2.32	2.17	1.79	1.84	1.43
9	2.28	2.40	2.49	2.70	2.34	2.38	1.83	1.56	3.01	2.49	2.50	1.91	2.49	2.14	1.88	1.68	1.74	2.39	1.99	2.32	2.33	1.91	1.56	1.31
10	2.55	2.51	2.73	3.15	2.51	2.62	1.70	1.56	2.95	2.70	3.11	2.20	2.91	2.67	2.28	1.35	1.46	2.39	2.54	2.66	1.99	1.93	2.06	1.17

Note. Perform req. = performance requirements are high; Study req. = study requirements are high; Exam mat. = too much exam material; Fail exam = failing the exam; Lecture supp. = lack of support from lecturers; Study org. = lack of organization of the studies; Overcrowded = overcrowded lectures; Anonymity = anonymity at the institution; Practical work = desire for practical work; Practical relev. = lack of professional and practical relevance of the studies; Interest in sub-ject; Interest in prof. = lack of interest in the professions possible with the degree obtained; False expec. = false expectations; Fin problems = financial problems; Earn money = earning money quickly; Pregnancy = pregnancy; Child care = childcare; Fam reasons = family reasons; Personal suit. = doubt about personal suitability for the degree course; Incompatibility = incompatibility of degree course and employment; Job offer = interesting job offer; Lack opportun. = poor job opportunities in my field of study; Illness = illness. Abroad = degree course abroad or internship abroad.

**Supplement F:** Means and standard deviations (in brackets) of the responses across the 24-item Reason for Dropout Questionnaire among the twelve dropout classes

Reasons for Dropout Items	C1: low interest/ performance problems (n = 328)	C2: unsure (n = 103)	C3: ambiguous job needed (n = 96)	C4: C5: challenged family plus (n = 249)	C6: C7: low interest (n = 24)	C8: going abroad (n = 221)	C9: study re- quirements alternative employment (n = 37)	C10: desire for illness (n = 38)	C11: C12: family only (n = 102)	C12: (n = 93)
Perform. req.	3.47(1.5)	2.46(1.6)	2.76(1.6)	2.93(1.4)	2.29(1.4)	1.44(0.8)	2.46(1.3)	3.16(1.6)	1.73(1.3)	2.29(1.6)
Study req.	3.61(1.4)	2.49(1.6)	3.30(1.6)	3.07(1.6)	2.79(1.7)	1.72(1.4)	3.08(1.7)	3.30(1.4)	2.17(1.4)	2.16(1.6)
Fail exam	4.09(1.8)	2.53(1.9)	3.08(1.9)	3.04(1.9)	3.04(1.9)	2.04(1.7)	1.45(3.9)	2.62(1.7)	2.68(1.8)	1.67(1.3)
Exam mat.	3.35(1.5)	2.54(1.7)	3.39(1.5)	3.09(1.6)	3.85(1.6)	2.79(1.6)	1.87(1.3)	3.27(1.8)	3.62(1.4)	2.27(1.5)
Lecture supp.	2.99(1.6)	2.29(1.4)	3.22(1.5)	2.83(1.5)	3.96(1.5)	1.87(1.1)	2.27(1.6)	3.12(1.8)	2.51(1.4)	2.40(1.6)
Study org.	3.00(1.5)	2.35(1.4)	3.37(1.4)	2.78(1.5)	3.07(1.6)	2.33(1.6)	2.55(1.7)	3.19(1.7)	2.54(1.5)	2.57(1.7)
Overcrowded	2.22(1.5)	1.57(1.1)	2.94(1.8)	2.46(1.6)	3.30(1.8)	2.12(1.6)	2.15(1.6)	2.62(1.8)	1.54(1.2)	2.00(1.4)
Anonymous	1.93(1.4)	1.58(1.2)	2.09(1.5)	1.79(1.3)	2.52(1.7)	1.58(0.9)	1.82(1.3)	2.00(1.1)	1.78(1.1)	1.87(1.5)
Practical work	3.46(1.7)	2.93(1.5)	3.94(1.7)	3.74(1.7)	3.52(1.6)	2.79(1.6)	3.47(1.9)	3.73(1.8)	3.46(1.7)	4.20(1.6)
Practical relev.	3.06(1.7)	2.54(1.5)	3.07(1.8)	3.01(1.7)	3.26(1.4)	2.46(1.6)	2.93(1.7)	3.31(1.9)	2.89(1.7)	3.12(1.6)
Interest in sub.	3.75(1.7)	3.05(1.8)	3.52(1.9)	2.70(1.5)	2.75(1.5)	2.70(1.5)	2.75(1.9)	2.81(1.5)	3.08(1.8)	3.06(1.5)
Interest in prof.	2.30(1.6)	2.10(1.6)	2.45(1.7)	2.18(1.5)	2.26(1.3)	2.92(1.8)	3.13(1.9)	1.96(1.5)	2.05(1.5)	1.34(0.8)
False expect.	3.84(1.5)	2.80(1.8)	3.58(1.6)	3.25(1.6)	3.22(1.4)	2.79(1.7)	3.35(1.9)	2.96(1.9)	3.08(1.6)	2.21(1.4)
Fin problems	1.87(1.3)	1.12(0.4)	3.55(1.9)	3.61(1.8)	4.00(1.8)	3.04(2.0)	1.38(0.8)	2.50(1.6)	3.00(2.0)	2.52(1.8)
Earn money	1.40(0.8)	1.00(0.0)	2.87(1.9)	3.29(1.7)	3.74(1.9)	2.04(1.6)	1.06(0.3)	2.31(1.7)	2.35(1.9)	2.33(1.6)
Pregnancy	1.00(0.0)	1.01(0.1)	1.00(0.0)	1.00(0.0)	2.44(2.1)	4.29(2.1)	1.00(0.0)	1.00(0.0)	1.00(0.0)	1.00(0.0)
Child care	1.02(0.1)	1.00(0.0)	1.03(0.2)	1.00(0.0)	4.96(1.5)	3.25(2.1)	1.01(0.1)	1.08(0.3)	2.86(2.0)	2.82(2.2)
Fam reasons	1.03(0.2)	1.77(1.3)	2.98(1.9)	2.03(1.5)	4.44(1.5)	4.33(1.8)	1.17(0.6)	2.08(1.7)	3.03(2.0)	2.50(1.8)
Personal suit.	2.83(1.5)	2.55(1.6)	2.84(1.6)	2.49(1.5)	2.81(1.4)	2.12(1.2)	2.33(1.5)	2.12(1.1)	3.08(1.4)	3.36(0.7)
Incompatibility	1.92(1.4)	1.02(0.2)	2.86(1.7)	3.09(1.9)	4.33(1.8)	2.75(1.8)	1.65(1.3)	2.66(1.8)	3.00(1.9)	2.97(1.8)
Job offer	2.06(1.6)	1.10(0.4)	2.65(1.7)	3.18(1.9)	3.07(2.0)	2.04(1.6)	2.50(1.9)	2.96(1.9)	1.00(0.0)	3.79(2.0)
Lack opportun.	1.31(0.6)	1.85(1.4)	2.30(1.7)	2.48(1.7)	3.04(1.9)	1.87(1.2)	2.68(1.8)	2.42(1.4)	2.22(1.7)	1.00(0.0)
Illness	1.00(0.0)	2.12(1.8)	3.50(1.6)	1.00(0.0)	3.07(2.1)	2.88(2.0)	1.01(0.2)	1.65(1.3)	3.57(2.2)	3.32(2.4)
A broad	1.00(0.0)	2.72(2.3)	1.00(0.0)	1.03(0.2)	2.33(2.0)	1.08(0.3)	1.01(0.1)	5.08(1.4)	1.14(0.4)	2.37(2.2)
										1.02(0.1)

Note. Mean scores calculated from six-point Likert scale anchored as 1 – plays no role to 6 – plays major role in decision to drop out. Perform. req. = performance requirements are high; Study req. = study requirements are high; Fail exam = failing the exam; Exam mat. = too much exam material; Lecture supp. = lack of support from lecturers; Study org. = lack of organization of the studies; Overcrowded = overcrowded lectures; Anonymity = anonymity at the institution; Practical work = desire for practical work; Practical relev. = lack of professional and practical relevance of the studies; Interest in sub. = low interest in subject; Interest in prof. = lack of interest in the professions possible with the degree obtained; False expect. = false expectations; Fin problems = financial problems; Personal suit. = doubt about personal suitability for the degree course; Fam reasons = family reasons; Fam. care = childcare; Job offer = interesting job offer; Lack opportun. = poor job opportunities in my field of study; Illness = illness; Abroad = degree course abroad or internship abroad.

Supplement G: Means and standard deviations for the contextual variables for all classes

		Large classes										Medium-sized classes					Small classes																
		1			4			7			2			3			11			12			5			6		8		9		10	
		<i>M</i>	2.17	2.16	2.03	<i>M</i>	1.72	2.11	<i>M</i>	1.49	1.56	<i>M</i>	0.67	<i>M</i>	0.63	0.79	<i>M</i>	2.14	2.11	<i>M</i>	2.03	2.03	<i>M</i>	1.63	2.11	<i>M</i>	2.03	2.03					
Dropout Intentions <sup>a</sup>		<i>M</i>	0.85	0.87	0.90	<i>SD</i>	0.85	0.87	<i>M</i>	0.81	0.87	<i>SD</i>	0.67	<i>M</i>	0.67	0.63	<i>M</i>	0.79	0.76	<i>M</i>	0.69	0.55	<i>M</i>	0.55	0.55	<i>M</i>	0.55	0.55	<i>M</i>	0.55	0.55		
High school GPA		<i>M</i>	2.71	2.72	2.41	<i>SD</i>	0.57	0.57	<i>M</i>	0.59	0.64	<i>SD</i>	0.58	<i>M</i>	0.65	0.60	<i>M</i>	0.60	0.52	<i>M</i>	0.73	0.69	<i>M</i>	0.69	0.63	<i>M</i>	0.69	0.69	<i>M</i>	0.69	0.69		
Academic self-concept <sup>b</sup>		<i>M</i>	4.00	4.10	4.84	<i>SD</i>	0.90	1.10	<i>M</i>	0.90	1.10	<i>SD</i>	1.10	<i>M</i>	1.00	0.90	<i>M</i>	0.80	1.10	<i>M</i>	0.90	1.20	<i>M</i>	1.10	1.10	<i>M</i>	0.90	0.90	<i>M</i>	0.90	0.90		
Subjective likelihood that student will graduate		<i>M</i>	3.38	3.51	4.02	<i>SD</i>	0.97	0.90	<i>M</i>	0.93	0.81	<i>SD</i>	0.85	<i>M</i>	0.86	0.77	<i>M</i>	0.77	0.62	<i>M</i>	1.21	0.75	<i>M</i>	0.88	0.75	<i>M</i>	0.75	0.75	<i>M</i>	0.75	0.75		
HISEI score		<i>M</i>	57.01	51.36	59.68	<i>SD</i>	20.30	51.36	<i>M</i>	59.68	20.80	<i>SD</i>	21.85	<i>M</i>	19.15	20.09	<i>M</i>	54.24	54.32	<i>M</i>	52.96	58.54	<i>M</i>	51.92	53.32	<i>M</i>	53.32	53.32	<i>M</i>	53.32	53.32		
Getting by with money		<i>M</i>	3.57	3.19	3.83	<i>SD</i>	1.07	1.09	<i>M</i>	0.94	0.90	<i>SD</i>	1.04	<i>M</i>	1.15	1.08	<i>M</i>	3.43	3.00	<i>M</i>	3.33	3.67	<i>M</i>	3.22	3.45	<i>M</i>	3.45	3.45	<i>M</i>	3.45	3.45		
Higher education satisfaction		<i>M</i>	5.05	5.23	5.79	<i>SD</i>	2.38	2.49	<i>M</i>	2.59	2.71	<i>SD</i>	2.41	<i>M</i>	2.48	2.43	<i>M</i>	6.18	6.00	<i>M</i>	4.94	4.77	<i>M</i>	4.90	6.86	<i>M</i>	6.86	6.86	<i>M</i>	6.86	6.86		

<sup>a</sup>Dropout intention was taken from wave 2. <sup>b</sup>Academic self-concept was taken from wave 2.

**Supplement H1:** Means and standard deviations and confidence intervals for students' high school GPA

Classes	<i>n</i>	Min.	Max.	<i>M</i>	<i>SD</i>	<i>SE</i>	Lower <i>CI</i>	Upper <i>CI</i>
C1: low interest/ performance problems	183	1.3	3.8	2.65	0.58	0.04	2.56	2.73
C2: unsure	51	1.0	3.5	2.31	0.64	0.09	2.14	2.49
C3: ambiguous	48	1.0	3.7	2.58	0.60	0.09	2.41	2.75
C4: job needed	141	1.0	3.9	2.63	0.61	0.05	2.53	2.74
C7: low interest	110	1.0	3.7	2.35	0.60	0.06	2.24	2.46

**Supplement H2:** Means and standard deviations and confidence intervals for students' academic self-concept

Classes	<i>n</i>	Min.	Max.	<i>M</i>	<i>SD</i>	<i>SE</i>	Lower <i>CI</i>	Upper <i>CI</i>
C1: low interest/ performance problems	147	1.00	6.00	4.00	0.89	0.07	3.86	4.14
C2: unsure	62	1.00	7.00	4.58	1.13	0.14	4.30	4.87
C3: ambiguous	52	2.50	7.00	4.41	0.98	0.14	4.14	4.67
C4: job needed	124	1.00	6.75	4.10	1.13	0.10	3.90	4.30
C7: low interest	99	2.75	7.00	4.84	0.87	0.09	4.67	5.01

**Supplement H3:** Means and standard deviations and confidence intervals for students' subjective likelihood to graduate

Classes	<i>n</i>	Min.	Max.	<i>M</i>	<i>SD</i>	<i>SE</i>	Lower <i>CI</i>	Upper <i>CI</i>
C1: low interest/ performance problems	147	1	5	3.38	0.97	0.08	3.22	3.54
C2: unsure	62	2	5	4.11	0.81	0.10	3.91	4.32
C3: ambiguous	53	1	5	3.66	0.85	0.12	3.43	3.89
C4: job needed	126	1	5	3.51	0.90	0.08	3.35	3.67
C7: low interest	98	1	5	4.02	0.93	0.09	3.84	4.20

**Supplement H4:** Means and standard deviations and confidence intervals for students' parents HISEI

Classes	<i>n</i>	Min.	Max.	<i>M</i>	<i>SD</i>	<i>SE</i>	Lower <i>CI</i>	Upper <i>CI</i>
C1: low interest/ performance problems	307	14.21	88.70	57.01	20.30	1.16	54.74	59.29
C2: unsure	89	14.21	88.70	62.07	20.80	2.21	57.75	66.40
C3: ambiguous	90	14.21	88.70	54.36	21.85	2.30	49.81	58.84
C4: job needed	228	11.74	88.70	51.36	20.12	1.33	48.75	53.97
C7: low interest	201	16.50	88.96	59.68	19.02	1.34	57.04	62.31

**Supplement H5:** Means and standard deviations and confidence intervals for students' estimation of getting by with available money

Classes	<i>n</i>	Min.	Max.	<i>M</i>	<i>SD</i>	<i>SE</i>	Lower <i>CI</i>	Upper <i>CI</i>
C1: low interest/ performance problems	136	1	5	3.57	1.07	0.09	3.39	3.75
C2: unsure	59	2	5	3.98	0.90	0.12	3.75	4.21
C3: ambiguous	49	1	5	3.27	1.04	0.15	2.98	3.56
C4: job needed	118	1	5	3.19	1.09	0.10	2.99	3.38
C7: low interest	87	1	5	3.83	0.94	0.10	3.63	4.03

**Supplement H6:** Means and standard deviations and confidence intervals for students' satisfaction with higher education

Classes	<i>n</i>	Min.	Max.	<i>M</i>	<i>SD</i>	<i>SE</i>	Lower <i>CI</i>	Upper <i>CI</i>
C1: low interest/ performance problems	151	0	10	5.05	2.38	0.19	4.67	5.43
C2: unsure	63	0	10	6.27	2.71	0.34	5.60	6.94
C3: ambiguous	55	0	10	5.33	2.41	0.33	4.69	5.96
C4: job needed	131	0	10	5.23	2.49	0.22	4.80	5.65
C7: low interest	104	0	10	5.79	2.59	0.25	5.29	6.29

## Supplement I: Outstanding single reasons for dropout

Items	Frequency
1 Illness	14
2 Abroad	11
3 Practical work	5
4 Pregnancy	5
5 Incompatibilty	4
6 Job offer	4
7 Fam reasons	4
8 Fail exam	3
9 Lack opportun.	1
10 Interest in sub.	1
Total	52

*Note.* Frequency = Cases present a single item with a score of either 5 or 6\*. All other item scores are either 1 or 2 of the 1-6 RDQ scale. Illness = illness, Abroad = degree course abroad or internship abroad, Practical work = wanted to do practical work, Pregnancy = pregnancy, Incompatibility = incompatibility of degree course and employment, Job offer = interesting job offer, Fam reasons = Family reasons, Fail exam = failing the exam, Lack opportun. = Poor job opportunities in my field of study; Interest in sub. = low interest in subject.

Markus Janssen  
Thomas Wiedenhorn  
(Eds.)

## School adoption in teacher education

Increasing pre-service teachers' responsibility during practice

2020, ca. 72 pages, br., € 19,90,  
ISBN 978-3-8309-4263-4  
[doi.org/10.31244/9783830992639](https://doi.org/10.31244/9783830992639)

Markus Janssen & Thomas Wiedenhorn (Eds.)

## School adoption in teacher education

Increasing pre-service teachers' responsibility during practice



### With contributions by

Sonja Bandorski, Johanna Gosch  
Kirsten Großmann, Markus Janssen  
Andreas Lund, Monika Merket  
Elin Bø Morud, Natalia Pavlova  
Olga Surina, Thomas Wiedenhorn

School adoption is an ambitious and innovative partnership model in teacher education which offers unique opportunities for in-service and pre-service teachers. At its core, teachers leave their school to be adopted by teacher students for one week. While the teachers engage in a professional development course outside the school, they are fully substituted by teacher students, who thus have an increased responsibility for the pupils' learning, for the organizational matters of the school and for their own professional development. In this volume, we present different international concepts of school adoption, lessons learned, and first theoretical considerations. With it, we invite teacher educators in schools, universities, and other institutions to engage into a dialogue about the perspectives school adoption offers for teacher education and teacher education research.

**WAXMANN**

[www.waxmann.com](http://www.waxmann.com)  
[info@waxmann.com](mailto:info@waxmann.com)

Nele McElvany  
 Heinz Günter Holtappels  
 Fani Lauermann  
 Aileen Edele  
 Annika Ohle-Peters  
 (Eds.)

## Against the Odds – (In)Equity in Education and Educational Systems

Dortmunder Symposium der  
 Empirischen Bildungsforschung,  
 vol. 5, 2020, 186 pages, br., € 29,90,  
 ISBN 978-3-8309-4151-4  
 E-Book: € 26,99,  
 ISBN 978-3-8309-9151-9



### With contributions by

Susanne Backes, Jürgen Baumert,  
 Rolf Becker, Hans-Peter Blossfeld,  
 David Glauser, Aiso Heinze,  
 Thorsten Henke, Leonidas Kyriakides,  
 Alexander Lorz, Sara Möser,  
 Emer Smyth, Birgit Spinath,  
 Lysann Zander

(In)Equity in education and educational systems is a topic of outstanding relevance. Factors, like, for example, students' socio-economic status or migration background have been found to be strongly related to academic achievement in various studies and this is a constant finding in large-scale assessment studies – especially for Germany. Does educational inequity accumulate over a child's life span or what compensating factors can be identified? Not least because of the topics' timeless meaning the 5<sup>th</sup> Dortmund Symposium on Empirical Educational Research was focused on questions pertaining to (in)equity. Factors of influence and approaches to overcome (in)equity were discussed by an international and interdisciplinary expert panel, with contributions from Cyprus, Ireland, Switzerland, Luxembourg, and Germany.

**WAXMANN**

[www.waxmann.com](http://www.waxmann.com)  
[info@waxmann.com](mailto:info@waxmann.com)

Marc J. de Vries, Stefan Fletcher,  
Stefan Kruse, Peter Labudde,  
Martin Lang, Ingelore Mammes,  
Charles Max, Dieter Münk,  
Bill Nicholl, Johannes Strobel,  
Mark Winterbottom  
(Eds.)

# The Impact of Technology Education

## International Insights

Center of Excellence for Technology  
Education (CETE), vol. 3, 2020,  
178 pages, br., € 34,90,  
ISBN 978-3-8309-4141-5  
E-Book: € 30,99,  
ISBN 978-3-8309-9141-0

M. J. de Vries, S. Fletcher, S. Kruse,  
P. Labudde, M. Lang, I. Mammes,  
C. Max, D. Münk, B. Nicholl, J. Strobel &  
M. Winterbottom (Eds.)

## The Impact of Technology Education

International Insights



### With contributions by

Victoria Adenstedt, Stefan Fletcher,  
Annika Gooß, Karin Güdel,  
Anni Heitzmann, Christian K. Karl,  
Alexander F. Koch, Stefan Kruse,  
Heide Lukosch, Ingelore Mammes,  
Elizabeth McGregor Jacobides,  
Andreas Müller, Johannes Strobel,  
Yan Sun, Marc J. de Vries, Mark  
Winterbottom

The increasing use of technology in our lives requires not only the qualification of young professionals through vocational training in order to maintain innovation and technical and societal progress, but also a technical education "for everyone", to become a society with technology literacy. A lack of technology activities may not only result in a "technology illiteracy", thus making a responsible participation in social life more difficult, but also has an impact on identity development. Against this background, technology education is getting important and has an impact on various aspects of the personality, e.g. skills, knowledge and interest in technology, which initiate lifelong learning.

With the combination of articles, the editors of Technology Education Vol. III want to give an insight into international approaches of technology education and its impact. Nine authors, respectively teams of authors from various countries present their educational settings and the impact it has for the personality development in technology.

**WAXMANN**

[www.waxmann.com](http://www.waxmann.com)  
[info@waxmann.com](mailto:info@waxmann.com)

Silke Trumpa, Emma Kostiainen,  
Isolde Rehm, Matti Rautiainen  
(Eds.)

## Innovative schools and learning environments in Germany and Finland

Research and findings of  
comparative approaches

Ideas of good and  
next practice

2020, 208 pages, pb, € 34,90,  
with coloured illustrations,  
ISBN 978-3-8309-4169-9  
E-Book: € 30,99,  
ISBN 978-3-8309-9169-4

Silke Trumpa,  
Emma Kostiainen,  
Isolde Rehm,  
Matti Rautiainen  
(Eds.)

# INNOVATIVE SCHOOLS and learning environments in Germany and Finland

Research and findings of comparative approaches

WAXMANN

### With contributions by

Lucia Hefti, Marc Kirschbaum,  
Emma Kostiainen, Laura Kuusela,  
Suvi Nenonen, Katja Ninnemann,  
Lachlan Paterson, Iida-Maria  
Peltomaa, Matti Rautiainen,  
Isolde Rehm, Silke Trumpa,  
Marianne Väliaho, Albrecht Wacker,  
Anne Weidemann-Akkermans

In this book, innovations in the education system are discussed and exemplary developments and projects from Finland and Germany are compared. On the one hand, the focus is on current research results, while on the other hand, examples of good practice from both countries describe remarkable ways of implementation and provide creative suggestions for one's own educational work. These are based on the experience of ten years of international cooperation between two teacher-training university locations.

**WAXMANN**

[www.waxmann.com](http://www.waxmann.com)  
[info@waxmann.com](mailto:info@waxmann.com)