

Empirische Bildungsforschung –
eine Standortbestimmung

Journal for Educational Research Online
Journal für Bildungsforschung Online

Volume 11 (2019), No. 1

Nele McElvany, Miriam Marleen Gebauer
& Cornelia Gräsel (Hrsg.)

Empirische Bildungsforschung – eine Standortbestimmung



Waxmann 2019
Münster • New York

JERO
Journal for Educational Research Online
Journal für Bildungsforschung Online
www.j-e-r-o.com

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:

Andreas Sander, Institut für Schulentwicklungsforschung, Technische Universität Dortmund, Phone: 0049 (0)2 31/75 58 292,

Michael Rochnia, Institut für Bildungsforschung in der School of Education, Bergische Universität Wuppertal, Phone: 0049 (0)2 02/43 95 129

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 register as a reader under www.j-e-r-o.com 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.

Please register as an author on the website www.j-e-r-o.com if you want to submit your article considering the submission guidelines.

ISBN 978-3-8309-3994-8

ISSN 1866-6671

© Waxmann Verlag GmbH 2019

Steinfurter Straße 555, 48159 Münster, Germany

Internet: www.waxmann.com, E-mail: info@waxmann.com

Print: Hubert & Co., Göttingen

Cover: Anne Breitenbach, Münster

Setting: Stoddart Satz- und Layoutservice, Münster

All rights reserved

Inhalt

<i>Nele McElvany, Miriam Marleen Gebauer & Cornelia Gräsel</i> Empirische Bildungsforschung – eine Standortbestimmung.....	5
<i>Dieter Lenzen</i> Empirische Bildungsforschung in Deutschland: Rückblick, Zukunft und Bedeutung für die Erziehungswissenschaft – Fragen an Professor Lenzen.....	9
<i>Hans-Peter Blossfeld, Gwendolin Josephine Blossfeld & Pia Nicoletta Blossfeld</i> Soziale Ungleichheiten und Bildungsentscheidungen im Lebensverlauf: Die Perspektive der Bildungssoziologie.....	16
<i>Ludger Woessmann</i> Die Abwägung berufsspezifischer und allgemeiner Bildung über den Lebenszyklus: Eine bildungsökonomische Perspektive.....	31
<i>Nils Berkemeyer, Björn Hermstein, Sebastian Meißner & Ina Semper</i> Kritische Schulsystementwicklungsforschung: Ein normativ-analytischer Forschungsansatz der schulischen Ungleichheitsforschung	47
<i>Ingrid Gogolin</i> Lernende mit Migrationshintergrund im deutschen Schulsystem und ihre Förderung – Forschungstraditionen und aktuelle Entwicklungen.....	74
<i>Knut Schwippert</i> Was wird aus den Büchern? Sozialer Hintergrund von Lernenden und Bildungsungleichheit aus Sicht der international vergleichenden Erziehungswissenschaft	92

Birgit Eickelmann, Julia Gerick & Mario Vennemann
Unerwartet erfolgreiche Schulen im digitalen Zeitalter –
Eine Analyse von Schulmerkmalen resilienter Schultypen
auf Grundlage der IEA-Studie ICILS 2013118

Nele McElvany & Franziska Schwabe
Geschlechtsunterschiede im Lesen an digitalen Geräten?
Analysen zur Rolle von Motivation und Selbstkonzept..... 145

Olaf Köller, Jennifer Meyer, Steffani Saß & Jürgen Baumert
Neue Analysen zu einem alten Thema: Effekte von
Intelligenz und Motivation auf Schulleistungen..... 166

Contents

<i>Nele McElvany, Miriam Marleen Gebauer & Cornelia Gräsel</i> Empirical Educational Research A look at where we are.....	5
<i>Dieter Lenzen</i> Empirical educational research in Germany: Retrospection, future and importance for educational science – Questions to Professor Lenzen.....	9
<i>Hans-Peter Blossfeld, Gwendolin Josephine Blossfeld & Pia Nicoletta Blossfeld</i> Social inequality and educational decisions over the life course: An educational sociology perspective	16
<i>Ludger Woessmann</i> Facing the life-cycle trade-off between vocational and general education in apprenticeship systems: An economics-of-education perspective	31
<i>Nils Berkemeyer, Björn Hermstein, Sebastian Meißner & Ina Semper</i> Critical educational system development research: A normative-analytical research approach to school-based inequality research	47
<i>Ingrid Gogolin</i> Learners with a migration background in the German school system and their support – Tradition and current trends.....	74
<i>Knut Schwippert</i> What’s about the books? Social background of students and <i>educational opportunities</i> from the perspective of international large-scale surveys	92

Birgit Eickelmann, Julia Gerick & Mario Vennemann
Unexpectedly successful schools in a digital age –
An analysis of school characteristics of resilient school
types based on the IEA study ICILS 2013.....118

Nele McElvany & Franziska Schwabe
Gender gap in reading digitally?
Examining the role of motivation and self-concept..... 145

Olaf Köller, Jennifer Meyer, Steffani Saß & Jürgen Baumert
New analyses of an old topic: Effects of intelligence and
motivation on academic achievement..... 166

Nele McElvany, Miriam Marleen Gebauer & Cornelia Gräsel

Empirische Bildungsforschung – eine Standortbestimmung

Special Issue Editorial

Empirische Bildungsforschung – ein Blick zurück

In den letzten zwei Dekaden hat die Empirische Bildungsforschung in Deutschland einen beispiellosen Aufschwung erlebt und durch zahlreiche Studien und deren Forschungsbefunde richtungsweisende Erkenntnisse gewonnen und der Bildungslandschaft von der primären über die sekundäre bis hin zu der tertiären Bildungsetappe bereitgestellt. Ausgangspunkt dieser intensiven Forschungsarbeit waren die sogenannte *empirische Wende* seit Anfang der 1990er Jahre und die große Aufmerksamkeit, die insbesondere die Befunde der Studien TIMSS 1995 und PISA 2000 bei allen Beteiligten im Bildungskontext und in der Öffentlichkeit erfuhr. Es folgten zahlreiche strukturelle Entwicklungen auf unterschiedlichen Ebenen. Zu diesen gehörten im Bereich der Forschung vor allem die Etablierung großer Förderprogramme durch die Deutsche Forschungsgemeinschaft und das Bundesministerium für Bildung und Forschung, der Ausbau der wissenschaftlichen Infrastruktur sowie die Gründung einer eigenen Fachgesellschaft. Entwicklungen im Bildungswesen umfassten substantielle Reformen wie beispielsweise die Einführung von Bildungsstandards oder die Veränderungen von Schulformen in den Ländern, verbunden auch mit der Etablierung und Weiterentwicklung einer systematischen und einheitlichen Bildungsmonitoringstrategie. Diese Entwicklungen wurden durch einen fruchtbaren Austausch zwischen Bildungsforschung, Bildungspraxis, Bildungsadministration und Bildungspolitik begleitet. Die gesellschaftliche Relevanz vieler in der Empirischen Bildungsforschung bearbeiteter Themen bedingte zusätzlich ein aufmerksames Auge der interessierten Öffentlichkeit.

Prof. Dr. Nele McElvany · PD Dr. Miriam Marleen Gebauer, Technische Universität Dortmund, Institut für Schulentwicklungsforschung (IFS), Vogelpothsweg 78, 44227 Dortmund, Deutschland
E-Mail: nele.mcelvany@tu-dortmund.de
miriam.gebauer@tu-dortmund.de

Prof. Dr. Cornelia Gräsel, Universität Wuppertal, Institut für Bildungsforschung in der School of Education, Gaußstr. 20, 42119 Wuppertal, Deutschland
E-Mail: graesel@uni-wuppertal.de

Empirische Bildungsforschung – zentrale aktuelle Stränge

Die internationalen Schulleistungsvergleichsstudien waren wichtige Ausgangspunkte der Empirischen Bildungsforschung in Deutschland und bilden immer noch ein kontinuierliches Gerüst der Aktivitäten, doch ist die Bildungsforschung längst geprägt durch eine Vielzahl unterschiedlicher Studiendesigns, Ansätze und Methoden. Dabei stehen je nach inhaltlichem Schwerpunkt das Bildungssystem, die Schule als Organisation, Klassen bzw. Lerngruppen oder individuelle Schülerinnen und Schüler und ihre Familien im Fokus. Institutionelle, soziale und individuellen Bedingungen von Bildung wie auch Bildungsprozesse und Bildungserträge werden dabei unter anderem aus den Blickwinkeln der Erziehungswissenschaft, der Pädagogischen Psychologie, der Fachdidaktiken, der Rehabilitationswissenschaft, der Bildungssoziologie und der Bildungsökonomie betrachtet. Als große thematische Stränge lassen sich unter anderem die Schulentwicklungs-, Lehrkraft- und Unterrichtsforschung, die Ungleichheitsforschung in Bezug auf sozioökonomischen Hintergrund, Zuwanderungshintergrund oder Geschlecht sowie die Analyse individueller Schülermerkmale in den Bereichen Kognition, Motivation und Emotion identifizieren, die ihrerseits mit einer Weiterentwicklung und Prüfung der jeweiligen Methoden und Testverfahren einhergehen. Gerade die Synergien, die bei der theoretischen und empirischen Arbeit an übergeordneten Forschungsthemen unter Einbezug unterschiedlicher wissenschaftlicher Perspektiven entstehen, machen das Potential der Empirischen Bildungsforschung für umfassende und relevante Erkenntnisse aus.

In dem vorliegenden Themenheft werden zentrale Fragestellungen der Empirischen Bildungsforschung aus unterschiedlicher disziplinärer Sichtweise und Forschungstradition betrachtet. Der Band führt wichtige theoretische Entwicklungslinien, den aktuellen empirischen Erkenntnisstand sowie zukunftsweisende Perspektiven zusammen und beruht dabei auf deutsch- und englischsprachigen Beiträgen. Damit wird ein Überblick bereitgestellt, der in der Forschungstradition der Empirischen Bildungsforschung einerseits die disziplinären Perspektiven würdigt und andererseits den Erkenntnisertrag disziplinübergreifender Forschung herausstellt.

Einleitend gibt *Dieter Lenzen* aus erziehungswissenschaftlicher Perspektive Antworten auf grundsätzliche Fragen zur Geschichte, Begrifflichkeit und aktuellen sowie zukünftigen Rolle der Empirischen Bildungsforschung. Das Interview *Empirische Bildungsforschung in Deutschland: Rückblick, Zukunft und Bedeutung für die Erziehungswissenschaft* zeichnet dabei wichtige Etappen der Entwicklung nach und setzt sich in diesem Zusammenhang auch mit dem Verhältnis von Empirischer Bildungsforschung und Erziehungswissenschaft auseinander.

Hans-Peter Blossfeld, *Gwendolin Josephine Blossfeld* und *Pia Nicoletta Blossfeld* nehmen die bildungssoziologische Perspektive ein und rücken die Lebensverlaufsperspektive in den Fokus. Ihr Beitrag *Soziale Ungleichheiten und*

Bildungsentscheidungen im Lebensverlauf: Die Perspektive der Bildungssoziologie befasst sich mit dem Einfluss der Familie auf die Bildungschancen der Kinder und stellt dabei die primären, sekundären und tertiären Effekte der sozialen Herkunft in den Vordergrund.

Die Lebenslaufperspektive wird aus bildungsökonomischer Perspektive auch in dem Beitrag *Facing the life-cycle trade-off of vocational education in apprenticeship systems: An economics-of-education perspective* von Ludger Woessmann aufgegriffen. In dem Text steht das berufliche Ausbildungssystem mit seinen langfristigen Effekten in einer sich verändernden Wirtschaftswelt im Fokus der Analysen, aus deren Befunden konkrete Empfehlungen für die Gestaltung eines zukunftsorientierten Ausbildungssystems abgeleitet werden.

Nach Familie und Ausbildungssystem richten Nils Berkemeyer, Björn Hermstein, Sebastian Meißner und Ina Semper in ihrem Beitrag *Kritische Schulsystementwicklungsforschung: Ein normativ-analytischer Forschungsansatz der schulischen Ungleichheitsforschung* die Aufmerksamkeit auf die Schule. Die Autorengruppe plädiert für eine gesellschafts- und gerechtigkeits-theoretisch fundierte Weiterentwicklung der gegenwärtigen Schulentwicklungsforschung und zeigt am Beispiel des Nationalen Bildungsberichts analytische Potentiale auf.

Aus der interkulturell vergleichenden Perspektive kommend betrachtet Ingrid Gogolin in ihrem Beitrag *Lernende mit Migrationshintergrund im deutschen Schulsystem und ihre Förderung – Forschungstraditionen und aktuelle Entwicklungen*. Dabei werden die aktuellen Perspektiven zunächst historisch eingeordnet, bevor die aktuelle Befundlage zu Geklärtem und Ungeklärtem dargestellt und durch Befunde aus zwei Studien ergänzt wird.

Knut Schwippert lenkt den Fokus auf die methodische Herausforderung adäquate Messinstrumente zur Verfügung zu haben und stellt die Frage: *Was wird aus den Büchern? Sozialer Hintergrund von Lernenden und Bildungsungleichheit aus Sicht der international vergleichenden Erziehungswissenschaft*. Der Indikator des familiären Buchbesitzes wird trotz der zunehmenden Verbreitung digitaler Medien auf der Basis empirischer Analysen als geeignetes Instrument für die Erfassung des ökonomisch-kulturellen Hintergrunds von Familien eingestuft.

Das zentrale Thema der Digitalisierung in Schulen wird im Kontext der international vergleichenden Studien in dem Beitrag *Unerwartet erfolgreiche Schulen im digitalen Zeitalter – eine Analyse von Schulmerkmalen resilienter Schultypen auf Grundlage der IEA-Studie ICILS 2013* von Birgit Eickelmann, Julia Gerick und Mario Vennemann thematisiert. Die Autorengruppe richtet den Blick auf technologische (Schul-)Entwicklung im Bildungssystem und geht der Frage nach, welche institutionellen Merkmale und Ressourcen die *computer literacy* bei Schülerinnen und Schülern an Schulen in herausfordernder Lage begünstigen können.

Ebenfalls das Thema Digitalisierung aufgreifend widmen sich Nele McElvany und Franziska Schwabe mit dem Beitrag *Gender gap in reading digitally? Examining the role of motivation and self-concept* den offenen Fragen, die im Zuge der zunehmend digital- anstellen von papierbasierten Kompetenzerfassung im Kontext von Bildungsstudien bestehen. Motivationale Merkmale im Bereich

des Lesens, nicht aber im Bereich des Bearbeitens digitaler Aufgaben erweisen sich als prädiktiv für Lesekompetenzen und bieten neben der Stabilität des Geschlechtereffekts erste Validitätshinweise.

Die pädagogisch-psychologische Perspektive ist auch Ausgangspunkt des abschließenden Beitrags. In diesem betrachten *Olaf Köller, Jennifer Meyer, Steffani Saß* und *Jürgen Baumert* kognitive und motivationale Merkmale von Schülerinnen und Schülern unter der Überschrift *New analyses of an old topic: Effects of intelligence and motivation on academic achievement* und analysieren deren Bedeutung für Leistungsmerkmale.

Empirische Bildungsforschung – Wilfried Bos

Das vorliegende Themenheft widmen wir – anlässlich seines Abschieds aus dem aktiven Universitätsdienst – Wilfried Bos, der diese Zeitschrift initiierte und über viele Jahre gemeinsam mit Cornelia Gräsel die Schriftleitung inne hatte. Wilfried Bos hat als Empirischer Bildungsforscher, langjähriger Direktor des Instituts für Schulentwicklungsforschung (IFS) an der TU Dortmund, GEBF-Gründungsmitglied und EERA Network Convenor durch eine Vielzahl an interdisziplinären Studien und Projekten sowie durch sein bildungspolitisches Engagement die Bildungsforschung nicht nur innerhalb Deutschlands, sondern auch auf europäischer Ebene maßgeblich geprägt und mitgestaltet. Neben dem richtungsweisenden Einfluss, den er über die jahrelange Leitung von großen internationalen Schulleistungsvergleichsstudien wie IGLU, TIMSS und ICILS sowie die Etappenleitung im NEPS im Bereich des Bildungsmonitorings und der Qualitätssicherung im Bildungswesen hatte, waren und sind ihm in seiner Arbeit insbesondere die Thematik der sozialen Ungleichheiten im deutschen Bildungssystem und die Förderung der Schulentwicklung durch eine enge Zusammenarbeit von Wissenschaft und Bildungspraxis wichtig.

Die Autorinnen und Autoren dieses Themenhefts sind Weggefährtinnen und Weggefährten, mit denen Wilfried Bos an den unterschiedlichen Wirkungsstätten seiner wissenschaftlichen Laufbahn zusammengearbeitet hat. Ihnen danken wir ebenso herzlich wie auch den Gutachterinnen und Gutachtern, die ebenfalls zum Kreise der Weggefährtinnen und Weggefährten oder des von Wilfried Bos seinerzeit geförderten wissenschaftlichen Nachwuchses gehören: Kathrin Dederling, Sabine Hornberg, Anke Hußmann, Ramona Lorenz, Wolfram Rollett, Katja Scharenberg, Renate Schulz-Zander, Tobias Stubbe und Ariane Willems.

Dortmund, Wuppertal im Januar 2019 –
Nele McElvany, Miriam M. Gebauer und Cornelia Gräsel

Dieter Lenzen

Empirische Bildungsforschung in Deutschland:

Rückblick, Zukunft und Bedeutung für die
Erziehungswissenschaft – Fragen an Professor Lenzen

Zusammenfassung

Professor Dieter Lenzen nimmt in dem Interview Stellung zu verschiedenen Themenfeldern, die die Empirische Bildungsforschung im Speziellen, aber auch die Erziehungswissenschaft im Allgemeinen betreffen. Unter anderem werden die Entwicklung der Empirischen Bildungsforschung der letzten Jahre und der Stand heute fokussiert. Ebenfalls wird die „Eltern-Kind-Beziehung“ zwischen Erziehungswissenschaft und Empirischer Bildungsforschung thematisiert. Abschließend wird ein Blick in die Zukunft gewagt und die zukünftige Position der Empirischen Bildungsforschung im Wissenschaftssystem diskutiert sowie mögliche Zukunftsthemen angesprochen, welche die Bildungsforschung vermutlich noch beschäftigen werden.

Empirical educational research in Germany:

Retrospection, future and importance for educational science –
Questions to Professor Lenzen

Abstract

In this interview, Professor Dieter Lenzen comments on diverse issues focusing on empirical educational research, but also on educational science in general. He discusses, among other topics, the development and current situation of empirical educational research. Additionally, he addresses the “parent-child relationship” in educational science and empirical educational research. Finally, the interview turns to the future position of empirical educational research in the scientific system and possible prospective issues.

Prof. Dr. Dr. h.c. Dieter Lenzen, Universität Hamburg, Mittelweg 177, 20148 Hamburg, Deutschland
E-Mail: praesident@uni-hamburg.de

JERO:

Welches sind die zentralen Verdienste der Empirischen Bildungsforschung, wenn Sie auf die Entwicklung der letzten Jahre zurückblicken? Welches sind die blinden Flecken?

Dieter Lenzen:

Die Empirische Bildungsforschung, wenn der bestimmte Artikel *die* überhaupt gerechtfertigt ist angesichts der großen Spannweite zwischen quantitativen und qualitativen Methoden, besteht gewiss in der Durchsetzung oder Wiederdurchsetzung des Realitätsprinzips und der Beendigung von pädagogischen Lebenslügen. Was bedeutet das?

Historisch gesehen bot die Erziehungswissenschaft sensu Pädagogik im Sinne ihres geisteswissenschaftlichen Konzepts noch vor der Wende zum 20. Jahrhundert etwas, was für die damalige Gesellschaft von besonderer Bedeutung war: normative Orientierung. Vergessen wir nicht: die Katastrophe der Napoleonischen Kriege, die Restauration in der Mitte des 19. Jahrhunderts, der flüchtige Sieg im Krieg von 1870/71 und dann die Katastrophe des 1. Weltkrieges und das Ende des Kaiserreiches – mehr als 100 Jahre Deutungs- und Orientierungsbedarfe für eine Gesellschaft. Und: die Neigung des politischen Systems, alle dort nicht lös-baren Probleme auf das Erziehungssystem zu verschieben, seien es verlorene Kriege oder gewonnene, die angeblich der deutsche Volksschullehrer gewonnen habe. Höhenflüge und Verbrechen, und insbesondere dann, wenn es im politischen System nicht weiterging – alles das landete im Erziehungssystem als der säkularisierten Deutungs- und Orientierungsmaschinerie, an deren Hebeln zuvor die Repräsentanten des religiösen Systems gesessen hatten.

Erfolgreich konnte ein Erziehungssystem dann sein, wenn es versprach, der nachwachsenden Generation die im politischen System formulierten und gewünschten Orientierungen zu vermitteln. Dass das politische System dabei einem anderen Kommunikationscode (Macht/Nicht-Macht) verpflichtet war als das Erziehungssystem (Wissen/Nicht-Wissen, Können/Nicht-Können, etc.) wurde nicht gewusst. Parsons und Luhmann waren noch nicht geboren.

Wenn man von einer kleinen Episode der Forderung nach empirischer Erziehungswissenschaft am Beginn des 20. Jahrhunderts absieht, hat sich an dieser Realitätsverweigerung oder -verdrehung bis zum Ende des 2. Weltkrieges wenig geändert, gleichgültig ob reformpädagogisch gehofft oder nationalsozialistisch gefordert wurde. Die für sie günstigste Deutung, die der geisteswissenschaftlichen Pädagogik mit ihrem Versagen während der Nazizeit nachträglich prädi-ziert werden konnte, war die, dass die Reduktion der Methodologie auf „Deutung“ systematisch darauf angewiesen ist, dass etwas geschehen sein muss, bevor man es deutet. Ausschwitz musste also geschehen sein, bevor man es deuten konnte. Geisteswissenschaftliche Pädagogik konnte sich insoweit nicht als emanzipatorische Avantgarde verstehen. Dieses hat sich erst langsam entwickelt, verstärkt sicher durch die Studentenbewegung in den 60er und 70er Jahren, aber weiterhin unter der Missachtung von Kausalitäten. Ob große Hypothesen eines

Wilhelm Reich oder eines Theodor W. Adorno empirisch valide sein würden, ob es also strukturelle „Ursachen“ für den faschistischen Charakter gibt, sei es sexuelle Beschränkung oder autoritäre Erziehung, um diese Frage hat sich geisteswissenschaftliche Pädagogik – ebenso wenig wie emanzipatorische – ernsthaft gekümmert.

Die in den 70er Jahren im Gefolge des Positivismusstreits bedeutsamer werdende Empirische Bildungsforschung allerdings auch nicht, da sie eigenartigerweise immer stärker mit konservativen politischen Positionen verbunden war. Dieses verdankt sich dem (unglücklichen) Umstand, dass das „Gute“, Emanzipatorische in der Geisteswissenschaft gesehen wurde und nicht in einem Wissenschaftsverständnis, das zu Unrecht mit Erbsenzählerei konnotiert wurde. Emanzipation als Erziehungsziel wäre möglicherweise erfolgreicher gewesen, wenn man sich von Anfang an Gedanken über „Gütekriterien“ gemacht und den Schulterchluss mit der Empirischen Bildungsforschung gesucht hätte.

Da dieses nicht geschah, hat Empirische Bildungsforschung eines Anlasses bedurft, der wiederum außerhalb des Erziehungssystems entstand, eher im System der Wirtschaft, um diesem methodologischen Ansatz zu seinem Recht zu verhelfen. Die ökonomische Krise nach der globalen Wende hat nach Erklärungen, wenn nicht nach Schuldigen gesucht und die internationalen Vergleichsstudien boten sich als Erklärungen an. Tatsächlich sind sie aber Beschreibungen, deren besondere Leistung darin bestand, zum richtigen Zeitpunkt auf ein zweifellos massives Problem des deutschen Erziehungssystems (sensu Ausbildungssystems) zu verweisen. Dieses ist die historische Leistung von Teilen der Empirischen Bildungsforschung, die daran beteiligt waren. Das politische System hat darauf reagiert und nicht jeder Empirische Bildungsforscher hat sich am Anfang der Neigung enthalten können, handfeste politische Empfehlungen zu geben, obwohl aus Beschreibungen zunächst ja einmal nichts folgt.

Wenn man also nach „blinden Flecken“ sucht, dann gibt es vielleicht auf dem Hemd der Erziehungswissenschaft eher einen bleichen als einen blinden Fleck: Ich meine das Theoriedefizit etlicher Ansätze der empirischen Erziehungswissenschaft. Während geisteswissenschaftlich orientierte Pädagogen hauptsächlich Theorien und Deutungen lieferten und kein valides Wissen, so könnte man extrem zugespitzt sagen, war es bei Teilen der empirischen Erziehungswissenschaft umgekehrt: eine Akkumulation von Fakten und Zusammenhängen, deren Befunde allerdings nicht aus der Verfolgung von theoriegestützten Hypothesen folgten, sondern zunächst einmal aus schierer Beobachtung.

JERO:

Sind die (inter)nationalen Schulleistungsvergleichsuntersuchungen „Segen oder Fluch“ für die Empirische Bildungsforschung?

Dieter Lenzen:

Ich würde nicht fragen, ob die (inter)nationalen Schulleistungsvergleichsuntersuchungen „Segen oder Fluch“ sind, sondern ob sie Anfang, Höhepunkt oder

Ende der Empirischen Bildungsforschung sein könnten. Von hinten her beantwortet: Das Ende sind sie sicher nicht, da die Empirische Bildungsforschung sicher nicht auf PISA und Co. beschränkt ist, sondern einen im liberalistischen, politischen und wirtschaftlichen System selbstverständlichen Gedanken des Wettbewerbs bedient. Politiker glauben zu wissen, wo sie stehen, auch wenn manche Forschungsergebnisse, landes-, bezirks- oder gar schulscharf geschaltet, unter Kuratel gestellt wurden. Daraus könnte man schließen, dass eine wirkliche Problembearbeitung gar nicht gewünscht wird, sondern vorzugsweise eine Bestätigung der eigenen Politik. Der Jubel über die statistisch einer gerade signifikanten „Verbesserung“ im Mathematikbereich zeigt ja, worum es eigentlich geht: Demonstration von Aktivismus und Erfolg.

Höhepunkt? – wahrscheinlich am ehesten. Der „PISA-Hype“ ist vorbei. Das wirtschaftliche System hat keinen Bedarf daran, zurzeit jedenfalls nicht, ein Versagen auf ein anderes System zu verschieben, ebenso wenig das politische System. Zurzeit finden Probleme nicht im Medium von internationalen Schulleistungen, sondern im Medium von Gefährdung/Nichtgefährdung durch islamistische Terroristen statt. Hier wäre selbst die Erwartung an das Erziehungssystem zu kühn, nämlich Terroranschläge durch Pädagogik zu verhindern. Gleichwohl: Die Übernahme einer Integrationsaufgabe wird erwartet. Das hat aber weniger mit Schulleistungen zu tun als mit der Anpassung an gesellschaftliche Normen. Damit beginnt sofort die Frage, nach welchen Normen denn angepasst oder integriert werden soll. Leitkultur ja oder nein, und in gewisser Weise sind wir damit wieder in der Mitte der 70er Jahre. Wie operationalisiert man denn Leitkultur, so dass messbar wäre, ob ein Integrant einer Leitkultur folgt?

Mit dieser Antwort ist die Alternative „Anfang der Empirischen Bildungsforschung“ im Prinzip mitbeantwortet. Über den Anfang ist sie längst hinweg. Solange nicht erneut massive Fragen gesellschaftlicher Orientierung in den Mittelpunkt rücken, ist an einen Bedeutungsverlust Empirischer Bildungsforschung im nennenswerten Maße nicht zu denken. Sie hat ihren Platz erworben und muss keine großen Anstrengungen unternehmen, um sich zu behaupten.

JERO:

Welche Bedeutung hat die Empirische Bildungsforschung für die Erziehungswissenschaft in Deutschland – und welche Bedeutung hat die Erziehungswissenschaft für die Empirische Bildungsforschung?

Dieter Lenzen:

Die Empirische Bildungsforschung hat das erziehungswissenschaftliche Genre in gewisser Weise rehabilitiert. Man denke nur an die Ketten von gehässigen Journalistenkommentaren, die die Erziehungswissenschaft immer noch als Vehikel der Gesellschaftsveränderung fürchteten, sich mit dem Abdruck von Dissertationstiteln über das Fach lustig machten oder schlicht seine Einstellung verlangten. Das dieses nicht mehr so ist, dass also die Erziehungswissenschaft als breites, nach der Medizin wohl größtes Fach, stabilisiert werden konnte, ist im

Wesentlichen das Verdienst der Empirischen Bildungsforschung. Niemand bezweifelt mehr, dass man mit „der“ Erziehungswissenschaft nichts anfangen könne, sondern im Gegenteil, dass sie gefährlich sei, ist kein Thema mehr.

Welche Bedeutung die Erziehungswissenschaft für die Empirische Bildungsforschung hat, das ist schwerer zu sagen. Wer ist „die Erziehungswissenschaft“? Empirische Bildungsforschung ist ja ein Bestandteil von Erziehungswissenschaft. Die Frage heißt also so gestellt, welche Bedeutung hat die Mutter für das Kind oder die Polizeiwache für den Polizisten? – Ohne Mütter und Wachstube wäre der Beruf des Kindes oder des Polizeibeamten eine frostige Angelegenheit. Empirische Bildungsforschung ist kein Fach, sondern eher eine Fachrichtung oder die Bezeichnung für eine Methodologie. Empirische Forschung kommt ohne Theoriebildung nicht aus. „Die“ Erziehungswissenschaft wird sich allerdings anstrengen müssen, einen fruchtbaren Dialog mit ihren „Kindern“ in theoretischer Hinsicht so zu führen, dass Theorien und operationalisierbare Hypothesen entwickelt werden. Wenn sie das nicht tut, besteht die Gefahr einer endgültigen Abkopplung und das Risiko, dass die DFG tatsächlich glaubt, Empirische Bildungsforschung sei ein Fach. Das ist genauso unsinnig, als wenn eine Forschungsförderungsinstitution ein Fach „Hermeneutik“ von der Germanistik abkoppeln würde, oder Homiletik von der Theologie oder theoretische Physik von der Physik. Insofern haben Mutter und Kind gefälligst zueinanderzufinden, und um die Metapher totzureiten: Die Mutter Erziehungswissenschaft muss Theorien für das Leben des Kindes entwickeln und umgekehrt ist das Kind vielleicht Garant dafür, dass die Mutter sich modernisiert.

JERO:

Die Erziehungswissenschaft hat die Erziehung in ihrem Namen, die Bildungsforschung die Bildung. Welchen Bildungsbegriff sehen Sie aus erziehungswissenschaftlicher Sicht in der Empirischen Bildungsforschung verfolgt?

Dieter Lenzen:

Das ist eine interessante Frage. Ich glaube, dass der Gebrauch des Wortes „Bildung“ im Kompositum „Bildungsforschung“ eher unbedacht und zufällig gewählt wurde. Empirische „Erziehungsforschung“ hätte nicht gepasst, weil damit das Schulsystem nicht erfasst worden wäre. Gleiches gilt für „Empirische Ausbildungsforschung“, da es hier nicht nur um Berufsschulen geht. Präziser war aus meiner Sicht der ältere Begriff der „Lehr-/Lernforschung“, der aber sehr unhandlich und kaum nachvollziehbar war. Im Rückblick auf Klaus Prange muss man aber sagen, dass die Frage vorweg zu klären ist, ob das Spezifikum dessen, womit Erziehungswissenschaft sich beschäftigt oder auch durch was das Erziehungssystem gekennzeichnet ist, Erziehung, Bildung oder doch eher Lernen ist. Mit dieser Frage ist allerdings auch ein Grundsatzproblem berührt: Ist die Konzentration auf Lernen eine Reduktion des gesamten Prozesses des Aufwachsens auf die Dimension von Wissens- und Kompetenzerwerb? Oder ließe sich eine Operation finden, bei der auch das mit Sozialisation Gemeinte oder das mit Erziehung Verfolgte umgriffen wäre? Diese

Debatte ist im Grunde unabgeschlossen, und eine für alle überzeugende Lösung hat sich nicht gefunden. Insofern hat der Begriff „Bildungsforschung“ immerhin für sich, dass Alltagsmenschen damit etwas anfangen können. Sie assoziieren Bildungssystem und meinen damit meistens ohne dies nur die Schule. Das dem Bildungsbegriff innewohnende empathische und sogar humanistische Moment hat bei der Begriffsbildung von „Empirischer Bildungsforschung“ meines Wissens nie eine Rolle gespielt. Andererseits entwickelt ein solcher Begriff eine Eigendynamik, ähnlich der der Theologie, die ja eigentlich in ihrem Namen ein Wissen über Gott insinuiert. Da über ein lediglich aus Offenbarung abgeleitetes Etwas aber nicht gewusst sondern an es nur geglaubt werden kann, wäre der Begriff heute aufzugeben. Auf diesen Gedanken kommt indessen keiner. Also wird es wohl bei der Empirischen Bildungsforschung bleiben, bis jemandem etwas Neues einfällt.

Ein bisschen bedauerlich könnte es sein, dass das mit Bildung einmal gemeinte, nämlich die Selbstbildung der Person als der Humanität verpflichtete Persönlichkeit im Blick auf eine „Höherbildung der Menschheit“ banalisiert sein könnte. Im Augenblick bleibt uns vermutlich nichts anders übrig als von „Bildung I“ und „Bildung II“ zu sprechen. Das ist natürlich nicht ganz ernst gemeint.

JERO:

Welche zukünftige Entwicklung antizipieren Sie für die Empirische Bildungsforschung? Welche (Forschungs-)Schwerpunkte empfehlen Sie ihr für die nächste(n) Dekade(n)?

Dieter Lenzen:

Mit Entwicklungsprognosen für eine Fachrichtung wäre ich sehr vorsichtig. Das Wissenschaftssystem tendiert, schon aufgrund der Projektionen aus dem politischen System, gern dazu, Fragestellungen und Konjunkturen aufzunehmen, um sie zu einem Forschungsfeld, einer Fachrichtung oder sogar zu einem ganzen Fach zu verbauen. Die Wirtschaftswissenschaft im Sinne von Betriebswirtschaft ist ein solches Beispiel, die Linguistik war es in den 70er Jahren des vorangehenden Jahrhunderts und auch die Erziehungswissenschaft gehört dazu. Ich erinnere mich noch gut an die Debatten am Ende der 60er Jahre bei der Einführung des „Diplompädagogen“, dessen Berufs- bzw. Titelbezeichnung zu dem Entwicklungsstand der „Erziehungswissenschaft“ (und nicht „Pädagogik“) eigentlich unpassend war. „Dipl. Erz. Wiss.“ wäre aber zu absurd gewesen, „rer. ed.“ ist es immer noch. Die Selbstteilungsfreudigkeit der Wissenschaften ist ungebrochen und in staatlichen Hochschulsystemen auch erfolgreich. Denn Politiker können durch die Einrichtung entsprechender Lehrstühle Aktivität demonstrieren, wenn ein eigentlich politisch zu lösendes Problem in das Wissenschaftssystem (ähnlich der Praxis in Bezug auf das Bildungssystem) stattfindet.

Aber zu der eigentlichen Frage: Die Empirische Bildungsforschung wird einen gesunden, nicht hysterischen Platz im Rahmen der Sozialwissenschaften haben, hoffentlich als Bestandteil einer Erziehungswissenschaft, zu der sie nach der

„Ablösung vom Elternhaus“ wieder ein normalisiertes Verhältnis auf höherem Niveau entwickeln kann.

Was künftige Forschungsschwerpunkte angeht, so hängen diese leider nicht von Empfehlungen ab. Durch die Umstellung des Wissenschaftssystems auf ein Wettbewerbssystem, innerhalb dessen Politik durch themengebundene Ausschreibungen Schwerpunkte festlegt, sind erhebliche Freiheitsmaße auch für die Erziehungswissenschaft verloren gegangen. Wer teure Forschung realisieren möchte, ist, so gesehen, auf die Deutsche Forschungsgemeinschaft angewiesen, die noch als letzte Bastion die Freiheit der Wissenschaft auch im Sinne der freien Findung von Themen, unabhängig von gesellschaftlichen Konjunkturen verteidigt. Die Frage müsste also eher heißen, welche Schwerpunkte wohl künftig durch Politik – und ein bisschen auch durch Wirtschaft (deren Bedeutung allerdings in einem umgekehrt proportionalen Verhältnis zu den Fördersummen steht) – gesetzt werden. Und dann werden es wieder die im politischen System ungelösten, nicht lösbaren oder zur Lösung nicht gewünschten Probleme sein: allem voran Migration, Radikalisierung, religiöse Erziehung, sodann Re-Nationalisierung bis Provinzialisierung, Identitätssicherung (nicht -findung!), schlimmstenfalls sogar Verteidigungsbereitschaft, Normkonformität und manches von dem, was in der Folge der 70er Jahre überwunden schien. Dem werden sich auch weiterhin kompetenzorientierte Thematiken anschließen, wenn es um die Optimierung des Verhältnisses von Lerneffekten und deren Einsatz im öffentlichen und vielleicht auch privaten Leben geht. Möglicherweise werden auch Themen eine Rolle spielen, die mit der vielleicht eingeführten ökonomischen Grundsicherung zusammenhängen, nämlich Modi der Sinnfindung und Motivationssicherung – es wäre schön, wenn wir es schaffen könnten, bei allen Schwierigkeiten der Operationalisierung auch der Empirischen Bildungsforschung Aufgaben zu übertragen, die mit den Entstehungsbedingungen humanen Handelns und Existierens durch Erziehung, Bildung und Sozialisation der Subjekte hindurch verbunden sind: mehrdimensionale Bildung.

Hans-Peter Blossfeld, Gwendolin Josephine Blossfeld &
Pia Nicoletta Blossfeld

Soziale Ungleichheiten und Bildungsentscheidungen im Lebensverlauf: Die Perspektive der Bildungssoziologie

Zusammenfassung

Der vorliegende Beitrag setzt sich das Ziel, auf der Grundlage soziologischer Theorien und ausgewählter Forschungsergebnisse den Einfluss der Familie auf die Bildungschancen der Kinder aus einer Lebensverlaufsperspektive aufzuzeigen. Im Vordergrund stehen dabei die sogenannten primären, sekundären und tertiären Effekte der sozialen Herkunft. Die Ergebnisse zeigen, dass in Deutschland die soziale Herkunft noch immer zu den wichtigsten Faktoren gehört, die die Bildungschancen von Schülern beeinflussen. Dabei beeinflusst die Herkunftsfamilie die Bildungskarriere kumulativ. Auch über die Kohorten hinweg ist die Wirkung der sozialen Herkunft auf die Bildungschancen relativ stabil.

Schlagworte

Soziale Ungleichheit; Bildungsentscheidungen; Lebensverlauf; Bildungssoziologie

Social inequality and educational decisions over the life course: An educational sociology perspective

Abstract

Based on sociological theories and selected empirical research findings, this article aims to outline how families influence their children's educational opportunities from a life-course perspective. The focus of this contribution is on the so-

Prof. Dr. Hans-Peter Blossfeld (corresponding author), Fakultät Sozial- und Wirtschaftswissenschaften, Otto-Friedrich-Universität Bamberg, Feldkirchenstr. 21, 96052 Bamberg, Deutschland

E-Mail: hans-peter.blossfeld@uni-bamberg.de

Dr. Gwendolin Josephine Blossfeld, Fakultät Humanwissenschaften, Otto-Friedrich-Universität Bamberg, Augustenstr. 6, 96047 Bamberg, Deutschland

E-Mail: gwendolin.blossfeld@uni-bamberg.de

Dr. Pia Nicoletta Blossfeld, Institut für Soziologie, Universität Leipzig, Beethovenstraße 15, 04107 Leipzig, Deutschland

E-Mail: pia.blossfeld@uni-leipzig.de

called primary, secondary and tertiary effects of social origin. The results show that social background is still one of the most important factors that influence educational opportunities of students in Germany. The influence of social origin on educational careers is cumulative. Across cohorts, the impact of social background on educational opportunities seems to be relatively stable.

Keywords

Social inequality; Educational decisions; Life course; Educational sociology

1. Einleitung

In Deutschland gehört die soziale Herkunft zu den wichtigsten Faktoren, die die Bildungschancen von Schülern beeinflussen. So haben die großangelegten internationalen Schulleistungstudien wie die *Progress in International Reading Literacy Study* (PIRLS/IGLU), die *Trends in International Mathematics and Science Study* (TIMSS) und das *Programme for International Student Assessment* (PISA) immer wieder gezeigt, dass die Kompetenzen der deutschen Schüler im internationalen Vergleich stark von der sozialen Herkunft abhängen (Bos, 2008; Bos, Wendt, Köller & Selzer, 2012; Bos, Eickelmann & Gerick, 2014; Hußmann et al., 2017; Reiss, Sälzer, Schiepe-Tiska, Klieme & Köller, 2016). Darüber hinaus haben langfristig angelegte soziologische Studien nachgewiesen, dass trotz massiver Bildungsexpansion und vielfältiger Reformen im Bildungssystem in den letzten Jahrzehnten die herkunftsspezifischen Bildungschancen relativ unverändert geblieben sind (Becker, 2003; Blossfeld, 1984; Blossfeld, 1989; Hadjar & Becker, 2017; Müller & Haun, 1994; Shavit & Blossfeld, 1993). Damit wird die Hartnäckigkeit des Einflusses der sozialen Herkunft auf die Bildungsentscheidungen der Kinder, Jugendlichen und jungen Erwachsenen in Deutschland erklärungsbedürftig.

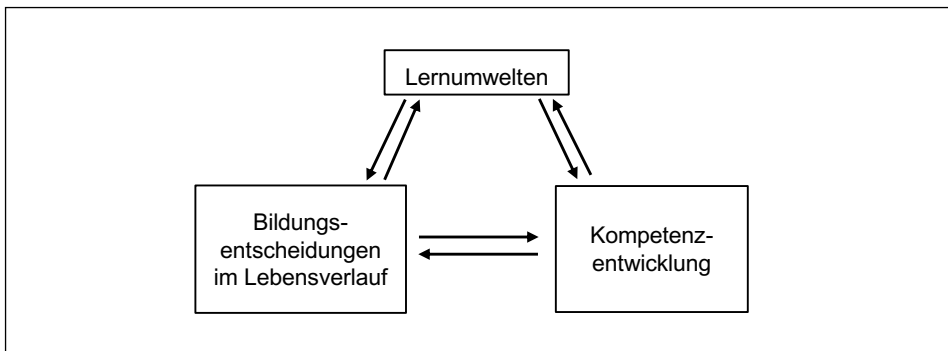
Der vorliegende Beitrag setzt sich das Ziel, auf der Grundlage soziologischer Theorien und ausgewählter Forschungsergebnisse den Einfluss der Familie auf die Bildungschancen der Kinder aus einer Lebensverlaufsperspektive aufzuzeigen. Im Vordergrund stehen dabei die sogenannten primären, sekundären und tertiären Effekte der sozialen Herkunft. Auf dieser Grundlage wird sich zeigen, wie die Familie die Bildungskarriere kumulativ beeinflusst und der Einfluss der sozialen Herkunft über die Kohorten relativ stabil bleibt.

2. Bildungsentscheidungen im Lebensverlauf

Nach der Geburt ist die Herkunftsfamilie in der Regel der wichtigste Ort, an dem sich primäre Sozialisationsprozesse vollziehen und in dem die Grundlagen der kognitiven, emotionalen und sozialen Entwicklung der Kinder gelegt werden (Knoppick, Becker, Neumann, Maaz & Baumert, 2016). Aber auch nach dieser

frühen Lebensphase unterstützen Familien mit und neben den zunehmend stärker werdenden Bildungsinstitutionen (wie den Kindertagesstätten, der Schule, den Einrichtungen der Berufsausbildung und des tertiären Bildungsbereichs) die kognitive Entwicklung des Einzelnen, seine soziale und kulturelle Integration in die Gesellschaft und die Entfaltung seines vorhandenen Entwicklungspotentials im Lebensverlauf. Im Kontext der Familien werden insbesondere zentrale Bildungsentscheidungen für die Kinder und Jugendlichen gefällt. Dabei verhalten sich Familien nicht nur passiv, sondern reagieren durchaus strategisch auf die Strukturen des Bildungssystems und seine Veränderungen durch Bildungsreformen, in dem Versuch, das Beste für ihre Kinder zu tun (Blossfeld, Kulic, Skopek & Triventi, 2017; Blossfeld, 2018). Dabei variieren nicht nur die sozialen Bedeutungen sukzessiver Bildungsübergänge, sondern auch die Mechanismen, nach denen die Bildungsentscheidungen in den Familien getroffen werden nach der sozialen Herkunft.

Abbildung 1: Das Zusammenwirken von Bildungsentscheidungen, Lernumwelten und Kompetenzentwicklungen im Lebenslauf



Anmerkungen. Eigene Darstellung

Die Lebensverlaufsforchung zeigt, dass heute die Bildungsprozesse vom Baby bis zum Greis meist kumulativ verlaufen (Blossfeld, Kilpi-Jakonen & Vono de Vilhena, 2014; Blossfeld, Kulic, Skopek & Triventi, 2017; Blossfeld, Buchholz, Skopek & Triventi, 2016). Das heißt, sie vollziehen sich nach dem Matthäus-Prinzip: Wer in einem bestimmten Lebensalter bereits ein höheres Bildungs- oder Kompetenzniveau erreicht hat, der hat auch im nächsten Schritt die jeweils bessere Chance, diese Vorteile noch weiter auszubauen. Kompetenzunterschiede zwischen verschiedenen sozialen Gruppen verschärfen sich deswegen tendenziell über den Bildungs- und Lebensverlauf. Abbildung 1 stellt diesen dynamischen Zusammenhang schematisch dar. Durch herkunftsspezifische Unterschiede in den Bildungsentscheidungen werden die Lernenden in verschiedene Lernumwelten kanalisiert, wo sie jeweils unterschiedliche Lernerfahrungen machen und Kompetenzfortschritte erzielen, die dann im weiteren Bildungsverlauf die Grundlage für die Entscheidungen der Familien und ihrer Kinder darstellen.

3. Herkunftsressourcen und Bildungsentscheidungen

Die soziologische Bildungsforschung konzentriert sich heute vor allem auf das Ausmaß und die Ursachen von nach sozialstrukturellen Gruppen variierenden Bildungsentscheidungen wie der Wahl einer Kindertagesstätte, einer weiterführenden Schule, der beruflichen Ausbildung, der Entscheidung für oder gegen das Abitur oder ein Hochschulstudium, die Fortführung einer Bildungskarriere oder die Teilnahme an einer beruflichen Weiterbildung. Dabei treten die Einflüsse der Herkunftsfamilie selbst dann auf, wenn bei den Kindern und Jugendlichen vergleichbare Kompetenzniveaus vorliegen. Deshalb ist zu klären, welche Bedeutung schichtspezifische Bildungsaspirationen, Erfolgserwartungen und Bewertungen von Kosten und Nutzen von Bildungsabschlüssen für die Ungleichheit der Bildungsprozesse heute haben.

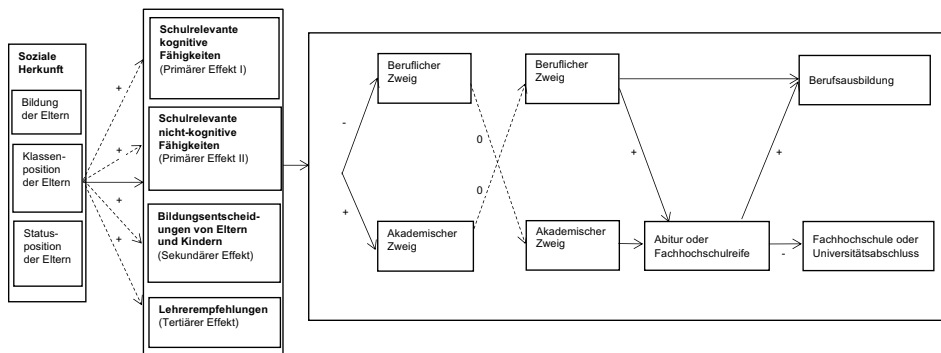
Die Soziologie geht davon aus, dass die Bildungsentscheidungen von Eltern und ihren Kindern von den verfügbaren Ressourcen der Herkunftsfamilie beeinflusst werden. Bukodi und Goldthorpe (2013) unterscheiden hier zwischen der Bildung der Eltern, ihrer Klassenposition und ihrem sozialen Status (siehe Abbildung 2). Jede dieser Ressourcen wirkt über eigene Mechanismen auf die Bildungsentscheidungen von Eltern und ihren Kindern (Blossfeld, 2018). Dabei repräsentiert die *Bildung der Eltern* die *Unterstützung der Kinder bei den Hausaufgaben* und das *Beraten über die Wege im Bildungssystem* (Bukodi & Goldthorpe, 2013). Höher gebildete Eltern können ihren Kindern in der Schule helfen, wenn diese Probleme haben. Außerdem haben sie selber Erfahrung mit dem akademischen Bildungsweg und wissen, dass man vielleicht auch nur mit durchschnittlichen Schulleistungen das Gymnasium und ein Universitätsstudium erfolgreich absolvieren kann (Blossfeld, 2018).

Die *Klassenposition der Eltern* steht für die *ökonomischen Ressourcen*, die Familien zur Verfügung stehen um ihren Kindern in der Schule zu helfen (Bukodi & Goldthorpe, 2013). Familien in höheren Klassenpositionen verfügen meistens über mehr finanzielle Ressourcen, die es Ihnen erlauben bei schulischen Problemen ihrer Kinder eine Nachhilfe zu bezahlen und sie können ihren Kindern ein förderliches Lernumfeld schaffen in dem zum Beispiel jedes Kind ein eigenes Zimmer bekommt (Blossfeld, 2018). Des Weiteren gibt die Klassenposition der Eltern darüber Auskunft *wie weitsichtig Familien die Bildungskarrieren ihrer Kinder planen können* (siehe Zeithorizonttheorie; Blossfeld, 2018; Hillmert & Jacob, 2003). Eltern in höheren Klassenpositionen haben meistens sicherere Arbeitsplätze und bessere Karrieremöglichkeiten, die es ihren Kindern erlauben von Anfang an eine kostspieligere und langwierigere akademische Ausbildung in Betracht zu ziehen (Blossfeld, 2018).

Die *Statusposition der Eltern* (im Sinne von Weber, 1976) repräsentiert das *soziale Netzwerk* das einer Familie zur Verfügung steht um ihren Kindern in der Schule zu helfen und deren Bildungsaspirationen beeinflusst (Bukodi & Goldthorpe, 2013; Erikson, 2016). So sollten Freunde, Verwandte und Bekannte

von Familien mit einem hohen Status eher hohe Bildungsaspirationen haben. Diese hohen Erwartungen werden auf die Familien und ihre Kinder projiziert, so dass diese Kinder eine nichtakademische Bildung nicht mehr als Option wahrnehmen (Blossfeld, 2018). Außerdem steht der Status der Eltern nach Bukodi und Goldthorpe (2013) auch für die *kulturellen Ressourcen* die Kinder in Familien erlernen und die mit den Mittelschichtstandards, die im akademischen Bildungsweg von Lehrern und Schulen erwartet werden, besser harmonieren (Blossfeld, 2018).

Abbildung 2: Der Einfluss der drei Herkunftsressourcen über primäre, sekundäre und tertiäre Effekte auf die Bildungskarriere



Anmerkungen. Eigene Darstellung in Anlehnung an Blossfeld, Blossfeld und Blossfeld (2015) und Blossfeld (2016)

4. Primäre, sekundäre und tertiäre Effekte im Bildungsverlauf

Die soziologische Bildungsforschung unterscheidet vier Arten der sozialen Herkunftseffekte an sukzessiven Bildungsübergängen (siehe Abbildung 2): den primären Effekt I (schulrelevante kognitive Fähigkeiten), den primären Effekt II (schulrelevante nicht-kognitive Fähigkeiten), den sekundären Effekt (familiäre Bildungsentscheidungen) und den tertiären Effekt (Lehrerempfehlungen).

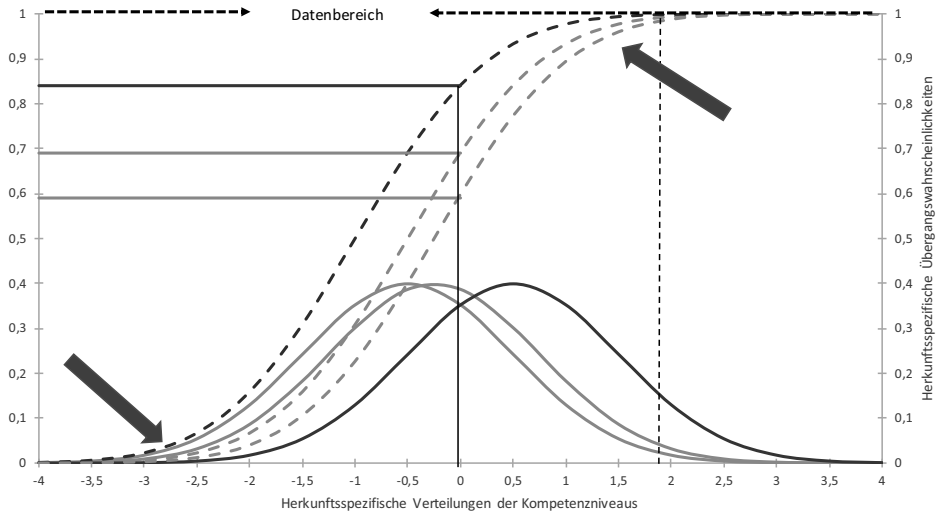
Ungleichheiten in den kognitiven Fähigkeiten, welche in der frühen Kindheit durch Unterschiede im Familienhintergrund geprägt werden, werden häufig als primäre Effekte der sozialen Herkunft auf den Bildungserfolg bezeichnet (vgl. Boudon, 1974). Die Bedeutung der kognitiven Fähigkeiten in der Vorhersage von Bildungserfolg in der Schule und Erfolg im späteren Erwachsenenalter ist in der Literatur gut dokumentiert (Almlund, Duckworth, Heckmann & Kautz, 2011). Dieser Effekt ist stark durch die Bildungsressourcen der Eltern bestimmt, da die Eltern die bedeutendsten Agenten sind, die die primären und weitestgehend auch sekundären Sozialisationsprozesse gestalten. Es wird dabei angenommen, dass

ihre Bildung einen starken Einfluss auf die alltägliche Interaktion mit den Kindern und das Niveau der kognitiven Stimulation in der häuslichen (Lern-)Umwelt hat (Blossfeld, Blossfeld & Blossfeld, 2015).

Der Erfolg an den verschiedenen Bildungsübergängen wird auch durch *nicht-kognitive Fähigkeiten* bestimmt (vgl. primärer Effekt II in Abbildung 2). In der Psychologie versteht man unter nicht-kognitiven Fähigkeiten Persönlichkeitsvariablen wie zum Beispiel Metakompetenzen, Selbstkonzept, Selbstregulation und soziale Kompetenzen (Weinert et al., 2011). Elterliche Bildung spielt dabei eine absolut zentrale Rolle für die Entwicklung der nicht-kognitiven Kompetenzen von Kindern (Almlund, Duckworth, Heckman & Kantz, 2011; Carneiro & Heckman, 2005; Kohn & Slomczynski, 1990). Eltern können dabei die Ziele und die Wahrnehmung sowie das Selbstkonzept und Wahlmöglichkeiten der Kinder durch das Informations- und Erfahrungsangebot beeinflussen. Kinder aus höhergebildeten Familien profitieren dabei auch von einem stärkeren akademischen Klima zu Hause. Über den Lebensverlauf hinweg sind Schüler immer häufiger dazu gezwungen eigene Bildungsentscheidungen zu treffen, sodass nicht-kognitive Fähigkeiten mit steigendem Alter immer mehr an Bedeutung gewinnen.

Im Zentrum der soziologischen Analyse steht der sekundäre Effekt der sozialen Herkunft (Boudon, 1974; vgl. Abbildung 2). Der sekundäre Effekt beschreibt die unterschiedlichen Bildungsentscheidungen verschiedener Herkunftsfamilien bei vergleichbaren kognitiven Fähigkeiten ihrer Kinder. Dabei spielen Kosten-Nutzen-Überlegungen von Familien und ihren Kindern die entscheidende Rolle. Die (subjektiv) erwarteten Kosten höherer Bildungsabschlüsse sind dabei für Familien mit geringen Herkunftsressourcen in der Regel höher. Gleichzeitig wird der subjektive Nutzen und die Erfolgswahrscheinlichkeit von diesen Familien als geringer eingeschätzt (Boudon, 1974). Deswegen schätzen Familien mit geringen Herkunftsressourcen typischerweise den Wert eines zukünftigen höheren Bildungsniveaus niedriger ein (höhere Zeitpräferenzen) und entscheiden sich daher seltener für akademisch anspruchsvollere und wirtschaftlich lohnenswertere Laufbahnen (hohe Risikoaversion) (Breen, van de Werfhorst & Jaeger, 2014; Erikson & Jonsson, 1996a). Aus einer soziologischen Perspektive hat der Mechanismus des Statuserhalts eine besondere Bedeutung (Breen & Goldthorpe, 1997, S. 283). Danach streben Eltern für ihre Kinder mindestens einen Bildungsabschluss an der mit ihrem vergleichbar ist. Das heißt, dass unterschiedliche Herkunftsfamilien verschiedene Bildungsziele für ihre Kinder haben. Zum Beispiel bedeutet ein mittlerer Bildungsabschluss für die eine Familie einen Aufstieg, während er für eine akademische Familie ein Abstieg ist.

Abbildung 3: Zusammenhänge zwischen primären und sekundären Effekten auf die Übergangentscheidung in Gymnasium nach Erikson et al. (2005)



Anmerkungen. Eigene Darstellung

Lange Zeit stellte sich die Frage nach einer adäquaten Methode, die Größenordnung der primären und sekundären Effekte gemeinsam zu schätzen. Erikson, Goldthorpe, Jackson, Yaish & Cox (2005) entwickelten eine Methode, die das Zusammenwirken von primären und sekundären Effekten bei verschiedenen sozialen Herkunftsguppen ermöglicht. Abbildung 3 verdeutlicht diesen Ansatz. In Abbildung 3 werden die primären Effekte durch unterschiedliche glockenförmige Kompetenzverteilungen (durchgezogene Linien) für drei Herkunftsguppen dargestellt. Aufgrund von herkunftsspezifischen Sozialisationseffekten ist die Kompetenzverteilung von Kindern aus Familien mit niedrigen Herkunftsressourcen nach links verschoben. Die Kompetenzverteilung von Kindern aus Familien mit mittleren Herkunftsressourcen liegt in der Mitte und die Kompetenzverteilung von Kindern aus Familien mit hohen Herkunftsressourcen ist nach rechts verschoben. Gleichzeitig werden in Abbildung 3 mit gestrichelten Linien die unterschiedlichen Übergangswahrscheinlichkeiten an das Gymnasium für die Kinder aus den drei Herkunftsguppen dargestellt. Abbildung 3 zeigt, dass die Übergangswahrscheinlichkeit an das Gymnasium für Kinder aus Familien mit besseren Herkunftsressourcen (linke S-Kurve) bei gleichem Kompetenzniveau höher ist als für Kinder aus Familien mit mittleren Herkunftsressourcen (mittlere S-Kurve) und diese wiederum höher ist als für Kinder aus Familien mit geringen Herkunftsressourcen (rechte S-Kurve). Abbildung 3 demonstriert, dass auch Kinder aus privilegierten Familien eine niedrige Übergangsrage an das Gymnasium haben, wenn deren Kompetenzen niedrig sind. Außerdem zeigt Abbildung 3, dass Kinder aus benachteiligten Familien eine hohe Übergangsrage auf das Gymnasium

haben, wenn ihre Kompetenzen hoch sind. Interessant ist in Abbildung 3 vor allem der Bereich der mittleren Kompetenzen. Bei ein und derselben mittleren Kompetenz, entscheiden sich privilegierte Familien häufiger dafür ihr Kind auf ein Gymnasium zu schicken als Familien mit geringeren Herkunftsressourcen. Mit anderen Worten, der herkunftsspezifische sekundäre Effekt ist im mittleren Kompetenz- und Notenbereich am größten.

Schließlich zeigt unser theoretisches Modell in Abbildung 2, dass die soziale Herkunft durch die Lehrerempfehlungen und Bildungsinstitutionen einen Einfluss auf die Bildungsentscheidungen hat (tertiärer Effekt, Esser, 2016). Es gibt zahlreiche empirische Belege dafür, dass Kinder höherer sozialer Herkunft, bei gleichen schulischen Leistungen und Schulnoten wie Kinder niedrigerer sozialer Herkunft, einfacher eine Lehrerempfehlung für den akademischen Bildungszweig erhalten und häufiger an der nächsthöheren Bildungsinstitution aufgenommen werden. Für diese Bevorzugung von Kindern aus Familien mit höheren Herkunftsressourcen gibt es mindestens drei Gründe (Ditton, 2010): (a) Generell bescheinigen Lehrer Kindern aus Familien mit besseren Herkunftsressourcen höhere schuladäquate nicht-kognitive Fähigkeiten (Erikson & Jonsson, 1996b). Es scheint also so zu sein (wie von Bourdieu (1973) beschrieben), dass höhergebildete Eltern ihre Kinder mit einem kulturellen Kapital ausstatten welches sie im Schulkontext erfolgreicher sein lässt. (b) Grundschullehrer unterstellen bessergebildeten Eltern, dass sie eher in der Lage sind ihre Kinder bei Problemen auf der höheren Schule zu unterstützen. Folglich schreiben sie diesen Kindern höhere Erfolgswahrscheinlichkeiten im Hinblick auf deren Bildungskarriere zu. Und (c) üben besser gebildete Eltern generell mehr Druck auf Lehrer und Bildungsinstitutionen aus, falls dies aus ihrer Sicht nötig sein sollte.

Auch bundeslandspezifische institutionelle Rahmenbedingungen haben einen erheblichen Einfluss auf die Bildungsentscheidungen und somit auf die Reproduktion sozialer Ungleichheit. Größere elterliche Entscheidungsspielräume beim Übergang zu den weiterführenden Schulen (wie zum Beispiel in Hessen), bieten insbesondere für Kinder aus Familien mit besseren Ressourcen höhere Chancen höhere Bildung zu erlangen. In Schulsystemen, in denen die Übergangsempfehlungen stärker auf Noten basieren (wie beispielsweise in Bayern), ist in letzter Instanz die Lehrerempfehlung ausschlaggebend und die Bildungsaspirationen privilegierter Familien werden dadurch etwas gebremst. Zum Beispiel zeigt Dollmann (2016) in einer Untersuchung, dass Kinder aus benachteiligten Familien dann eine größere Chance haben, auf das Gymnasium zu gehen, wenn die Übergangsentscheidungen stärker auf die tatsächlichen Schulnoten bezogen sind. Dadurch werden die überschießenden Bildungsaspirationen von Mittelschichteltern mit Kindern mit mittlerer Kompetenz beschränkt.

5. Bildungsentscheidungen und Mobilität zwischen den akademischen und nichtakademischen Bildungswegen im Lebensverlauf

In den letzten Jahrzehnten hat sich das Bildungssystem in Deutschland immer mehr von einem dreigliedrigen zu einem zweigliedrigen Bildungssystem entwickelt (siehe Abbildung 2; Helbig & Nikolai, 2015; Tillmann, 2010). Die Haupt- und Realschulen werden dabei zunehmend zusammengelegt (Schulen mit mehreren Bildungsgängen) und das Gymnasium bleibt als herausgehobene Schulform bestehen. Diese zweigliedrige Betrachtungsweise von akademischen und nichtakademischen Bildungswegen ist von Vorteil, da man so die zentralen Bildungsentscheidungen und -übergänge über einen langen historischen Zeitraum vergleichen kann (Blossfeld, 2018).

In Abbildung 2 werden die zentralen Übergänge in den akademischen und nichtakademischen Bildungswegen sowie die Übergänge zwischen ihnen im Karriereverlauf vereinfacht dargestellt. Blossfeld (2018) zeigt auf der Grundlage von Längsschnittdaten des Nationalen Bildungspanels (NEPS), dass über die Kohorten der Anteil der Kinder die nach der Grundschule den Übergang zum akademischen Bildungszweig machen, angestiegen ist. Entsprechend haben die Übergangsraten der Kinder nach der Grundschule zum nichtakademischen Bildungszweig über die Kohorten abgenommen. Obwohl damit alle Kinder verschiedener Herkunftsgruppen von dieser Öffnung profitiert haben, bleiben die relativen herkunftsspezifischen Bildungsungleichheiten bestehen.

Trotz der Bildungsreformen, die das Ziel hatten die Durchlässigkeit zwischen den Schulen in der Sekundarstufe I zu erhöhen, ist die Mobilität zwischen den verschiedenen Bildungszweigen im Kohortenvergleich erstaunlich gering. Falls es Mobilität von Schülern gibt, ist Abwärtsmobilität häufiger als Aufwärtsmobilität. Kinder von niedrig gebildeten Eltern haben dabei eine höhere Abwärtsmobilität und Kinder von hoch gebildeten Eltern sowie einem hohen Status sind eher aufwärts mobil.

Im Vergleich zur Sekundarstufe I, hat die Aufwärtsmobilität nach einem ersten Bildungsabschluss auf dem nichtakademischen Bildungsweg über die Kohorten hinweg deutlich zugenommen. Jugendliche mit einem Haupt- oder Realschulabschluss entscheiden sich von Kohorte zu Kohorte immer häufiger dafür, eine Hochschulzugangsberechtigung nachzuholen. Vor allem Kinder aus privilegierten Familien nutzen diesen zweiten Bildungsweg mit einer höheren Wahrscheinlichkeit (Blossfeld, 2018).

Gleichzeitig hat der Anteil der Schüler mit einer Hochschulzugangsberechtigung, die sich dann für eine berufliche Ausbildung entscheiden, über die Kohorten deutlich zugenommen. Hierbei handelt es sich vor allem um junge Erwachsene aus Familien mit geringen Herkunftsressourcen. Kinder mit einer Hochschulzugangsberechtigung, die aus Familien mit hohen Herkunftsressourcen stammen, entscheiden sich hingegen häufiger für einen tertiären Bildungsabschluss.

6. Die Stabilität der Herkunftseffekte im Zuge der Bildungsexpansion

Es existieren verschiedene, konkurrierende Theorien über die herkunftsspezifische Veränderung der Bildungschancen im Zuge der Bildungsexpansion. Erstens, die Modernisierungstheorie, die besagt, dass die Bildungsexpansion zu einem allgemeinen Rückgang der Ungleichheit der herkunftsspezifischen Bildungschancen führt (vgl. z.B. Lenski, 1966; Treiman, 1970). Nach dieser Theorie expandiert das Bildungssystem als Reaktion auf die funktionalen Anforderungen in modernen Gesellschaften. In diesem Modernisierungsprozess werden alle Eltern zunehmend besser durch die Massenmedien über die Bildungschancen und -möglichkeiten ihrer Kinder informiert, sodass der sekundäre Effekt der sozialen Herkunft abnehmen sollte. Darüber hinaus *sollten nach dieser Theorie* Auswahlverfahren im Bildungssystem rationaler und meritokratischer werden, was dann auch zu einem Sinken des tertiären Effekts der sozialen Herkunft führen sollte. Bildungschancen sollten somit zunehmend nur noch von den Schülerleistungen abhängen (die primären Effekte I und II der sozialen Herkunft). Die große Mehrzahl der empirischen Ergebnisse zeigt, dass die Vorhersagen der Modernisierungstheorie im Grunde nicht eingetroffen sind. Weiterhin sind die Bildungsentscheidungen verschiedener Herkunftsgruppen von großer Bedeutung.

Ein angemesseneres Modell scheint vielmehr die Theorie der kulturellen Reproduktion zu sein (Bourdieu, 1973; Bowles & Gintis, 1976; Collins, 1979). Sie behauptet, dass Bildungszertifikate die soziale Ungleichheit auf dem Arbeitsmarkt legitimieren. Vertreter dieser Theorie erkennen jedoch einen inhärenten Konflikt zwischen der Auswahl- und Sozialisationsfunktion von Bildung. Eine wesentliche Rolle von Bildungsinstitutionen ist die Integration von Kindern mit niedriger sozialer Herkunft in das vorherrschende Wertesystem der Gesellschaft (Meyer, Ramirez, Rubinson & Boli-Bennett, 1977). Die Bildungsexpansion der Sekundarstufe II steht im Einklang mit der steigenden Nachfrage benachteiligter Gruppen nach mehr Bildung. Mit anderen Worten, im Zuge der Bildungsexpansion von unten wird nicht nur das Erreichen der Grundschulbildung, sondern auch die Sekundarstufe I und II universal und zunehmend unabhängig von der sozialen Herkunft (Shavit & Blossfeld, 1993). In diesem Prozess sind über die Kohorten hinweg Familien mit besseren Herkunftsressourcen in der Regel die Vorreiter in den Bildungsentscheidungen, während Familien mit niedrigen Herkunftsressourcen häufig die Nachzügler sind (Blossfeld, Blossfeld & Blossfeld, 2015). Gleichzeitig wollen sich die privilegierten Gruppen ihre Vorteile im oberen Teil des Ungleichheitssystems erhalten. Sie bewahren ihre Privilegien durch den Erwerb eines Hochschulabschlusses. Dies ist aufgrund der Persistenz von primären, sekundären und tertiären Effekten der sozialen Herkunft möglich (siehe Abbildung 2). Das bedeutet, dass im Zuge des Ausbaus der tertiären Bildung der Anstieg der Bildungsnachfrage von Kindern höherer sozialer Herkunft immer größer sein sollte als der Anstieg der Bildungsnachfrage von Kindern niedriger sozialer Herkunft.

Raftery und Hout (1993) schlagen eine spezielle Version der Reproduktionstheorie für Bildungsübergänge vor. Ihre Maximally Maintained Inequality (MMI) Hypothese besagt, dass der Effekt der Bildungsherkunft auf den Bildungsübergang über die Kohorten hinweg nur dann abnimmt, wenn die privilegierten Gruppen am jeweiligen Bildungsübergang mehr oder minder bereits saturiert sind. Somit führt die Bildungsexpansion vor allem dann zu einem Anstieg der Bildungschancen von Kindern aus Familien mit niedrigen Herkunftsressourcen, wenn im Grunde alle Kinder aus Familien mit besseren Herkunftsressourcen diesen Bildungsübergang bereits erfolgreich absolvieren (Lucas, 2001). Dies bedeutet, dass im Zuge der Bildungsexpansion von unten die Bildungsübergänge der Sekundarstufen I und II über die Geburtskohorten hinweg universaler werden. Wenn jedoch alle Herkunftsgruppen an einem bestimmten Bildungsübergang saturiert sind, sind die Bildungszuwächse für Kinder mit niedriger sozialer Herkunft zum Teil illusorisch. Wenn fast jeder Schüler den Bildungsübergang schafft, dann ist dies keine besondere Leistung mehr und der jeweilige Bildungsübergang wird zu einer Voraussetzung für alle. Mit anderen Worten, höhere Übergangsraten an weitgehend universalisierten Bildungsübergängen verlieren ihre symbolische Bedeutung für die Bildungsungleichheit. Wenn jedoch privilegierte Kinder bei höheren Bildungsübergängen noch nicht saturiert sind, ist es unwahrscheinlich, dass die Bildungsexpansion zu verbesserten Chancen benachteiligter Kinder führt. Vielmehr bleiben ihre Chancen im Zeitverlauf relativ konstant oder reduzieren sich sogar (Blossfeld et al., 2015). Dies ist wiederum auf die primären, sekundären und tertiären Effekte der sozialen Herkunft zurückzuführen (vgl. Abbildung 2). Solange privilegierte Herkunftsfamilien an den verschiedenen Bildungsübergängen nicht saturiert sind, wird die Übergangswahrscheinlichkeit von Kindern aus Familien mit niedrigem Bildungsniveau eingeschränkt sein.

Eine weitere Theorie ist die Effectively Maintained Inequality (EMI) Theorie von Lucas (2001). Auch diese Theorie postuliert an höheren Bildungsübergängen eine starke Konkurrenz zwischen Familien unterschiedlicher sozialer Herkunft, konzentriert sich jedoch auf die qualitativen Bildungsunterschiede bei gleichem Bildungsniveau. Wenn neue Distinktionen auf der tertiären Ebene eingeführt werden, z. B. durch eine zunehmende Differenzierung zwischen Bachelor- und Masterstudiengängen oder zwischen Fachhochschulen und traditionellen Universitäten in Deutschland, prognostiziert diese Theorie, dass soziale Herkunft eine große Rolle dabei spielt, welcher tertiäre Abschluss gewählt wird (Lucas, 2009).

7. Abschließende Bemerkung

Dieser Beitrag weist darauf hin, dass die soziale Herkunft noch immer ein wichtiger Faktor für die Bildungschancen im Lebensverlauf ist. Auch über die Kohorten hinweg ist die Wirkung der sozialen Herkunft auf die Bildungschancen relativ stabil. Das Ziel dieses Beitrags war, die Rolle der Bildungsentscheidungen auf

Grundlage verschiedener Herkunftsressourcen im Zusammenspiel mit primären und tertiären Effekten im Lebensverlauf aufzuzeigen.

Die Lebensverlaufsforschung zeigt, dass heute die Bildungsprozesse vom Baby bis zum älteren Erwachsenen meist kumulativ verlaufen (Blossfeld, Kilpi-Jakonen & Vono de Vilhena, 2014; Blossfeld, Buchholz, Skopek & Triventi, 2016; Blossfeld, Blossfeld & Blossfeld, 2017). Durch herkunftsspezifische Unterschiede in den Bildungsentscheidungen werden die Lernenden in verschiedene Lernumwelten kanalisiert, wo sie jeweils unterschiedliche Lernerfahrungen machen und Kompetenzfortschritte erzielen, die dann im weiteren Bildungsverlauf die Grundlage für die Entscheidungen der Familien und ihrer Kinder darstellen.

Die Organisation des Bildungssystems mit seinen Übergängen bestimmt deswegen, wann im Lebensverlauf Bildungsentscheidungen mit welcher Konsequenz von den Familien (und Lehrern) getroffen werden müssen und in welchem Umfang diese später korrigiert werden können. Auf dieser Grundlage ist zu empfehlen, das Bildungssystem insgesamt möglichst lange offen zu halten und die Anschlussfähigkeit von Bildungsabschlüssen zu gewährleisten (Vermeidung von Bildungssackgassen). Der zweite und der kürzlich eröffnete dritte Bildungsweg (offene Hochschule) stellen hier einen großen Schritt in die richtige Richtung dar, weil hier auch berufliche Qualifikationen zunehmend anerkannt werden.

Aus einer Längsschnittstudie von Fend, Berger und Grob (2009) gibt es allerdings auch empirische Hinweise darauf, dass durch organisatorische Reformen des Bildungssystems (wie etwa durch die Einführung der Förderstufe oder der Gesamtschule) zwar kurzfristig die herkunftsspezifischen Effekte vermindert werden können, dass sich die Herkunftsfamilien dann aber mit ihren Wünschen langfristig doch wieder durchsetzen. Insbesondere bei späteren risikobehafteten Bildungsentscheidungen scheint sich der schulische Einfluss im Lebensverlauf zu verlieren und die familiären Ressourcen treten zunehmend wieder in den Vordergrund. Das heißt, dass sich die Herkunftsfamilien der Mittelschichten, die jeweils versuchen optimal mit ihren Strategien und Ressourcen für ihre Kinder zu sorgen, in unterschiedlichen Schulformen letztendlich immer wieder durchsetzen (Blossfeld, Buchholz, Skopek & Triventi, 2016). Vieles spricht deswegen dafür, dass man nicht nur mehr Frühförderung und Ganztagschulen braucht, sondern einen umfassenderen Ansatz benötigt, der insbesondere die Familien der benachteiligten Herkunftsgruppen stärker in die Bildungsförderung einbezieht (Familienbildung).

Literatur

- Almlund, M., Duckworth, A. L., Heckman, J. J. & Kautz, T. (2011). *Personality psychology and economics*. NBER Working Paper No. 16822. Retrieved from <http://www.nber.org/papers/w16822>
- Becker, R. (2003). Educational expansion and persistent inequalities of education. Utilizing subjective expected utility theory to explain increasing participation rates in upper secondary school in the federal republic of Germany. *European Sociological Review*, 19(1), 1–24.

- Blossfeld, H.-P. (1984). *Bildungsexpansion und Berufschancen*. Frankfurt a. M.: Campus.
- Blossfeld, H.-P. (1989). *Kohortendifferenzierung und Karriereprozeß*. Frankfurt a. M.: Campus.
- Blossfeld, H.-P., Buchholz, S., Skopek, J. & Triventi, M. (Hrsg.) (2016). *Models of secondary education and social inequality* (eduLIFE lifelong learning series). Cheltenham: Edward Elgar Publishing.
- Blossfeld, H.-P., Kilpi-Jakonen, E. & Vono de Vilhena, D. (Hrsg.) (2014). *Adult learning in modern societies. An international comparison from a life-course perspective* (eduLIFE lifelong learning series). Cheltenham: Edward Elgar Publishing.
- Blossfeld, H.-P., Kulic, N., Skopek, J. & Triventi, M. (Hrsg.) (2017). *Childcare, early education and social inequality. An international perspective* (eduLIFE lifelong learning series). Cheltenham: Edward Elgar Publishing.
- Blossfeld, P. N. (2016). *Changes in inequality of educational opportunity. A cross-national comparison and the long-term development in Germany*. Oxford, UK: University of Oxford.
- Blossfeld, P. N. (2018). *Changes in inequality of educational opportunity. The long-term development in Germany*. Wiesbaden: Springer.
- Blossfeld, P. N., Blossfeld, G. J., & Blossfeld, H.-P. (2015). Educational expansion and inequalities in educational opportunity: Long-term changes for East and West Germany. *European Sociological Review*, 31(2), 144–160.
- Blossfeld, P. N., Blossfeld, G. J., & Blossfeld, H.-P. (2016). Changes in educational inequality in cross-national perspective. In M. J. Shanahan, J. T. Mortimer & M. Kirkpatrick Johnson (Hrsg.), *Handbook of the life course* (S. 223–247). Cham: Springer International Publishing.
- Bos, W. (Hrsg.) (2008). *IGLU-E 2006: Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Eickelmann, B. & Gerick, J. (2014). Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern der 8. Jahrgangsstufe in Deutschland im internationalen Vergleich. In W. Bos, B. Eickelmann, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil, R. Schulz-Zander & H. Wendt (Hrsg.), *ICILS 2013 – Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich* (S. 113–145). Münster: Waxmann.
- Bos, W., Wendt, H., Köller, O. & Selter, C. (2012). *TIMSS 2011: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern im internationalen Vergleich*. Münster: Waxmann.
- Boudon, R. (1974). *Educational opportunity, and social inquiry. Changing prospects in Western society*. New York, NY: John Wiley & Sons.
- Bourdieu, P. (1973). Kulturelle Reproduktion und soziale Reproduktion. In P. Bourdieu (Hrsg.), *Grundlagen einer Theorie der symbolischen Gewalt*. Frankfurt a. M.: Suhrkamp.
- Bowles, S. & Gintis, H. (1976). *Schooling in capitalist America*. New York, NY: Basic Books.
- Breen, R. & Goldthorpe, J. H. (1997). Explaining educational differentials. Towards a formal rational action theory. *Rationality and Society*, 9, 275–305.
- Breen, R., van de Werfhorst, H. G. & Jaeger, M. M. (2014). Deciding under doubt. A theory of risk aversion, time discounting preferences, and educational decision-making. *European Sociological Review*, 30(2), 258–270.
- Bukodi, E. & Goldthorpe, J. H. (2013). Decomposing ‘social origins’. The effects of parents’ class, status, and education on the educational attainment of their children. *European Sociological Review*, 29(5), 1024–1039.

- Carneiro, P. & Heckman, J. J. (2005). Human capital policy. In J. J. Heckman & A. B. Krueger (Hrsg.), *Inequality in America. What role for human capital policies?* (S. 77–239). Cambridge, MA: MIT Press.
- Collins, R. (1979). *The credential society. An historical sociology of education and stratification*. New York, NY: Academic Press.
- Ditton, H. (2010). Schullaufbahnen und soziale Herkunft. Eine Frage von Leistung oder Diskriminierung? In S. Aufenanger, F. Hamburger, L. Ludwig & R. Tippelt (Hrsg.), *Bildung in der Demokratie* (S. 79–99). Opladen: Budrich.
- Dollmann, J. (2016). Unwillig oder benachteiligt? Migranten im deutschen Bildungssystem. In R. Becker & W. Lauterbach (Hrsg.), *Bildung als Privileg* (S. 253–280). Wiesbaden: Springer Fachmedien Wiesbaden.
- Erikson, R. (2016). Is it enough to be bright? Parental background, cognitive ability and educational attainment. *European Societies*, 18(2), 117–135.
- Erikson, R., Goldthorpe, J. H., Jackson, M., Yaish, M. & Cox, D. R. (2005). On class differentials in educational attainment. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27), 9730–9733.
- Erikson, R. & Jonsson, J. O. (Hrsg.) (1996a). *Can education be equalized? The Swedish case in comparative perspective*. Boulder, CO: Westview Press.
- Erikson, R. & Jonsson, J. O. (1996b). Explaining class inequality in education. In R. Erikson & J. O. Jonsson (Hrsg.), *Can education be equalized? The Swedish case in comparative perspective* (S. 1–63). Boulder, CO: Westview Press.
- Esser, H. (2016). Bildungssysteme und ethnische Bildungsungleichheiten. In C. Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf* (S. 331–396). Wiesbaden: Springer Fachmedien Wiesbaden.
- Fend, H., Berger, F. & Grob, U. (Hrsg.) (2009). *Lebensverläufe, Lebensbewältigung, Lebensglück. Ergebnisse der Life-Studie*. Wiesbaden: Verlag für Sozialwissenschaften.
- Hadjar, A. & Becker, R. (2017). Erwartete und unerwartete Folgen der Bildungsexpansion in Deutschland. In R. Becker (Hrsg.), *Lehrbuch der Bildungssoziologie* (S. 211–232). Wiesbaden: Springer Fachmedien Wiesbaden.
- Helbig, M. & Nikolai, R. (2015). *Die Unvergleichbaren*. Bad Heilbrunn: Julius Klinkhardt.
- Hillmert, S. & Jacob, M. (2003). Social inequality in higher education: Is vocational education a pathway to or away from university? *European Sociological Review*, 19(3), 319–334.
- Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E.-M., McElvany, N., Stubbe, T. C. & Valtin, R. (Hrsg.) (2017). *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Knoppick, H., Becker, M., Neumann, M., Maaz, K. & Baumert, J. (2016). Das subjektive Erleben des Übergangs in die weiterführende Schule. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 48(3), 129–143.
- Kohn, M. L. & Slomczynski, K. M. (1990). *Social structure and self-direction. A comparative analysis of the United States and Poland*. Oxford: Basil Blackwell.
- Lenski, G. E. (1966). *Power and privilege. A theory of social stratification*. New York, NY: UNC Press Books.
- Lucas, S. R. (2001). Effectively maintained inequality. Education transitions, track mobility, and social background effects. *American Journal of Sociology*, 106(6), 1642–1690.
- Lucas, S. R. (2009). Stratification theory, socioeconomic background, and educational attainment. *Rationality and Society*, 21(4), 459–511.
- Meyer, J. W., Ramirez, F. O., Rubinson, R. & Boli-Bennett, J. (1977). The world educational revolution, 1950–1970. *Sociology of Education*, 50(4), 242–258.

- Müller, W. & Haun, D. (1994). Bildungsungleichheit im sozialen Wandel. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 46(1), 1–42.
- Raftery, A. E. & Hout, M. (1993). Maximally maintained inequality. Expansion, reform, and opportunity in Irish education, 1921–75. *Sociology of Education*, 66(1), 41–62.
- Reiss, K., Sälzer, C., Schiepe-Tiska, A., Klieme, E. & Köller, O. (Hrsg.) (2016). *PISA 2015. Eine Studie zwischen Kontinuität und Innovation*. Münster: Waxmann.
- Shavit, Y. & Blossfeld, H.-P. (1993). *Persistent inequality. Changing educational attainment in thirteen countries*. Boulder, CO: Westview Press.
- Tillmann, K.-J. (2010). Der Schritt in die Zweigliedrigkeit. Motive und bundesweiter Vergleich – Teil II. *Schulverwaltung. Nordrhein-Westfalen*, 21(3), 93–95.
- Treiman, D. J. (1970). Industrialization and social stratification. *Sociological Inquiry*, 40(2), 207–234.
- Weber, M. (1976). *Wirtschaft und Gesellschaft*. Tübingen: J. C. B. Mohr.
- Weinert, S., Artelt, C., Prenzel, M., Senkbeil, M., Ehmke, T. & Carstensen, C. H. (2011). Development of competencies across the life span. *Zeitschrift für Erziehungswissenschaft*, 14(S2), 67–86.

Ludger Woessmann

Facing the life-cycle trade-off between vocational and general education in apprenticeship systems: An economics-of-education perspective

Abstract

This paper discusses the role of vocational education, and in particular apprenticeship education, in preparing students for the labor market, with a particular focus on a life-cycle perspective in changing economies. The basic idea is that vocational education may facilitate entry into the labor market but hurt employment opportunities later in life because of limited adaptability to changing economic environments. We summarize evidence on the changing effects of vocational education over the life cycle from the international adult achievement tests IALS and PIAAC and country-specific evidence. We then discuss policy implications for elements of future-oriented education systems, especially apprenticeship programs.

Keywords

Vocational education; Apprenticeship; Labor market; Life cycle

Die Abwägung berufsspezifischer und allgemeiner Bildung über den Lebenszyklus: Eine bildungsökonomische Perspektive

Zusammenfassung

Dieser Beitrag befasst sich mit der Rolle berufsspezifischer Bildung, insbesondere im Rahmen der dualen Berufsausbildung, in der Vorbereitung von Schülern auf den Arbeitsmarkt unter besonderer Berücksichtigung einer Lebenszyklusperspektive in einer sich verändernden Wirtschaft. Die Grundidee besteht darin, dass berufsspezifische Bildung den Einstieg in den Arbeitsmarkt erleichtert, im späteren Leben aber aufgrund von begrenzter Anpassungsfähigkeit an das sich wandelnde wirtschaftliche Umfeld die Beschäftigungsmöglichkeiten behindern kann. Wir fassen Evidenz über die sich verändernden Effekte berufsspezifischer

Prof. Dr. Ludger Woessmann, University of Munich and ifo Center for the Economics of Education, Poschingerstraße 5, 81679 Munich, Germany
E-mail: woessmann@ifo.de

Bildung über den Lebenslauf aus den internationalen Erwachsenentests IALS und PIAAC sowie länderspezifische Ergebnisse zusammen. Darauf aufbauend werden Politikimplikationen für Elemente eines zukunftsorientierten Bildungssystems diskutiert, mit einem besonderen Fokus auf die duale Berufsbildung.

Schlagworte

Berufsbildung; duale Ausbildung; Arbeitsmarkt; Lebenszyklus

1. Introduction

Vocational education programs have attracted particular interest in many countries in recent years (see, e.g., Zimmermann et al., 2013 and Hanushek, Schwerdt, Woessmann, & Zhang, 2017b for references). Advocates praise apprenticeship systems that are particularly common in German-speaking countries for the way they facilitate the youth's transition from school to work, thereby apparently combating youth unemployment. Other countries have school-based vocational programs that also prepare for a specific occupation, but without the close link to work experience within firms. Still other countries – such as the United States – have mostly abandoned vocational programs, relying on general education programs in the hope of providing the basis for later learning on the job.

An aspect that is of particular relevance to understand the relative merits of vocational and general education programs is economic change. In a static economy, the occupation-specific skills obtained in vocational education programs will be just as relevant towards the end of one's working life as they were at the beginning of one's career. But in reality, our economies are constantly changing. Two major trends are globalization and technological change. First, the integration of the former Eastern Bloc and China into the world economy has meant that many sectors of Western economies have been relocated to Eastern Europe and East Asia. This process of globalization has been further facilitated by technological developments in global communication and logistics. Second, technological progress has been particularly pronounced in information and communication technologies. Broadly speaking, computerization, automation, industrial robots, and digitalization have led to skill-biased and routine-biased technological change, replacing many tasks that specific occupations used to perform by machines, robots, and computer networks (e.g., Katz & Autor, 1999; Autor, Levy, & Murnane, 2003; Akerman, Gaarder, & Mogstad, 2015).¹

In a constantly changing knowledge-based economy, the facilitated labor-market entry due to vocational skills may have to be weighed against reduced later employability when specific skills become obsolete and workers lack the ability to adjust to the changed economic environment (Krueger & Kumar, 2004; Hanushek

¹ This paper was originally prepared as a background report for Finland's Economic Policy Council. I would like to thank Franziska Hampf and an anonymous referee for constructive comments.

et al., 2017b). Much of the recent discussion about the disappearance of formerly middle-class jobs in many Western societies due to technological advance is related exactly to the theme of adapting to changed conditions. Therefore, conveying skills that allow workers to adapt to changing environments must be a central aspect in evaluating the extent to which vocational and general education programs prepare students for lifetime opportunities on the labor market.

This paper discusses the role of vocational education, and in particular apprenticeship education, in preparing students for the labor market, with a specific focus on a life-cycle perspective in changing economies. The next section outlines the basic idea that vocational education facilitates entry into the labor market but may hurt employment opportunities later in life because of limited adaptability to changing economic environments. Section 3 presents evidence on the changing effects of vocational education over the life cycle, summarizing findings from the international adult achievement tests IALS and PIAAC as well as additional country-specific evidence. On this basis, section 4 discusses policy implications for elements of future-oriented education systems, and in particular apprenticeship programs.

2. Vocational education: The basic trade-off over the life-cycle

To highlight the trade-off involved in different types of education programs, it helps to distinguish two types of skills.² On the one hand, there are *general* skills such as broad knowledge and basic skills in language, communication, math, science, and any other cognitive, social, or personal skills that are valuable on the labor market in many different occupations. These skills serve as the foundation for further learning on the job. They can thus be used to enter different jobs. Moreover, at a later age, general skills that facilitate continuous learning such as basic cognitive skills, transversal skills, adaptability, creativity, problem solving, and critical thinking skills may provide a useful basis to adapt to changing labor-market requirements and facilitate the learning of a different set of job-related skills.

On the other hand, there are *job-specific* skills that are required only in a specific occupation. Learning these job-related skills prepares students well to work in the specific occupation, but they are less relevant in other occupations. Being equipped with the relevant job-specific skills makes students productive right when they enter the labor market, as they have the skills that employers require now. But if demand for these skills dissipates on the labor market, these skills become mostly obsolete.

As a stylized depiction, we can distinguish two types of education programs: general education programs and vocational education programs (see Table 1). For

2 See, e.g., Psacharopoulos (1987), Ryan (2001), and Zimmermann et al. (2013) for general discussion of aspects of vocational and general education curricula.

the argument here, the assessment of specific education programs depends on the extent to which the education program conveys general and/or job-specific skills. We will refer to programs that mostly convey general skills as general education programs and to programs that mostly convey job-specific skills as vocational education programs.

Table 1: Stylized depiction of types of education programs

	General education	School-based vocational education	Apprenticeship-type vocational education
Focus on skill type	General skills (language, communication, math, science, ...)	Occupation-/industry-specific skills	Occupation-/firm-specific skills
Entry into labor market	Not directly prepared, needs substantial learning on the job	Job-specific skills facilitate labor-market entry, but limits due to missing experience	Easy entry as education included direct labor-market experience
Labor market at later age	General skill base facilitates adaptation to changing economy	Job-specific skills may become obsolete in changing economy; low adaptability due to limited general skill base	Job-specific skills may become obsolete in changing economy; low adaptability due to focus on firm and limited general skill base

The indicated characteristics of the two types of skills imply that a basic trade-off exists between general and vocational education programs when it comes to their implications for labor markets over the life-cycle. Vocational education programs have the particular advantage of helping young people master the transition from school to work because the job-specific skills directly prepare students for the tasks that firms require them to perform (e.g., Shavit & Müller, 1998; Ryan, 2001; Zimmermann et al., 2013). As a consequence, they may result in reduced youth unemployment.

But there is another side to the same coin: If the skills conveyed by vocational programs are particularly useful in a specific occupation, they will be subject to increasing risk of becoming obsolete when the structure of the economy – and hence the specific occupation – changes. Furthermore, if the acquired general skill base is limited, vocationally educated people may find it hard to learn different job-specific skills. That is, over the life-cycle, vocational programs may lead to lower adaptability to technological and structural change (cf. Krueger & Kumar, 2004). As a consequence, they may come with the disadvantage of reduced employment opportunities at older ages. The main point is that there is a basic trade-off between vocational and general education in that any relative labor-market advantage of vocational over general education decreases with age.

Among the vocational education programs, there are two main stylized subtypes. On the one hand, there are school-based vocational programs that focus on skills that prepare for work in a specific occupation. On the other hand, there are

apprenticeship-type vocational programs that partly take place within firms so that industry is directly involved in the educational process. These combined school and work-based apprenticeship programs tend to provide a high intensity of industry-based vocational experience (cf. Wolter & Ryan, 2011). In addition, part of the skills acquired during the work experience will be specific to the particular firm in which they are obtained, rather than to the occupation or industry in general. This may further facilitate the school-to-work transition, but lead to additional skill obsolescence as soon as the worker is required to change firm. A particular advantage of apprenticeship systems over school-based systems is that they usually do not educate too many youth in areas without demand in the economy. Because apprenticeship places have to be offered by employers, there are limits to the extent to which youth are educated in occupations for which there is not enough demand on the labor market – limits that may be much less strict in school-based vocational systems.

3. Evidence on the effects of vocational education over the life-cycle

In recent years, an increasing body of evidence confirms the basic trade-off between vocational and general education programs over the life-cycle.

3.1 Initial evidence from IALS

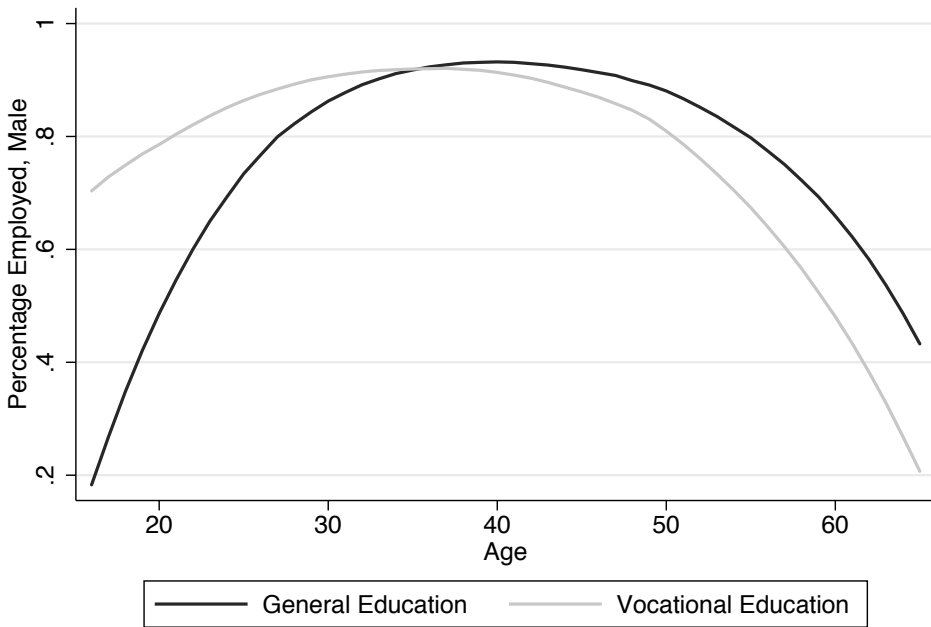
To test the labor-market effects of vocational and general education over the life-cycle, Hanushek et al. (2017b) use the data of the International Adult Literacy Survey (IALS). Coordinated by the Organisation for Economic Co-operation and Development (OECD) in the mid-1990s, IALS provides data for 15,000 individuals aged 16 to 65 years in 18 countries. The basic analysis compares the labor-market experience of people who attended vocational and general education programs at different ages.³ In the *prime age* (roughly age 30 to 45), people with the different types of education have very similar (and high) employment rates, indicating that there are no fundamental differences in labor-market attachment between these two groups (see Figure 1). However, there are strong differences both at labor-market entry and at older ages.

The key finding is the following trade-off in employment patterns over the life-cycle: At the beginning, individuals with a job-specific vocational education have better employment chances than those who completed general education. At older ages, it is the other way around, with individuals with a general education having

3 The analysis is restricted to males who historically show stable aggregate labor-force participation patterns during prime age. Cohort-specific selection into work by females raises concerns for this type of analysis.

better employment chances. This age-employment pattern is consistent with vocational education improving the transition from school to work but reducing adaptability of older workers to economic change. On the one hand, the skills generated by vocational education appear to facilitate the start into the labor market. On the other hand, vocationally educated people are at a higher risk of losing their jobs later in life if demand for their occupation-specific skills fades over time due to technological and structural change. The initial advantage of vocational education thus seems to turn into a disadvantage over the life-cycle. Short-term benefits have to be traded off against long-term costs.

Figure 1: Employment over the life-cycle with vocational and general education in IALS



Notes. Sample includes males who completed at least secondary education and are currently not students in the three “apprenticeship countries” (Denmark, Germany, and Switzerland). Smoothed scatterplot using locally weighted regressions (Stata *lowess*). Source: Hanushek et al. (2017b).

The life-cycle employment pattern – both the initial advantage and the later disadvantage – is particularly pronounced in those countries with extensive apprenticeship programs that combine school and work-based training – Denmark, Germany, and Switzerland in the IALS data. In these countries, the emphasis on industry-based education seems to provide the strongest treatment intensity of vocationalization. The results are robust to conditioning on observed background characteristics such as years of schooling, literacy scores, and parental education and occupation.

A key element of the empirical approach – called a “difference-in-differences” approach as it compares two differences in the employment rates: between vocational and general education and between young and old workers – is to confirm that today’s older workers provide a good approximation for what we can expect today’s younger workers to look like a few decades from now. This is a standard requirement of approaches that use cross-sectional data for life-cycle comparisons. To test this, Hanushek et al. (2017b) provide deepened analyses that make use of specific aspects of the IALS data, in particular the literacy and numeracy test produced by IALS. Among others, they match individuals within each country so that each individual with a vocational education is compared to an individual with a general education who is very similar in terms of skill scores as well as age, years of schooling, and parental education levels. Results are very similar in this sample of observationally comparable individuals, indicating that results are not driven by differences in these characteristics between vocationally and generally educated individuals. In their regression analyses, they can also hold constant any change in the overall shares of individuals entering vocational or general programs in a country over time. Furthermore, they show that while both literacy scores and parental education predict whether an individual has chosen a vocational or general education program, these predictions do not change significantly with the age of the individual. In other words, older people with a specific education type seem to be a good proxy for future expectations of younger people with the same education type.

The trade-off between education types over the life-cycle is visible not only in employment patterns, but also in income. Furthermore, at least in apprenticeship countries, there is some indication that as people get older, generally-educated individuals become increasingly likely to participate in adult career-related education. This increasing participation in adult education may be one mechanism by which generally-educated individuals maintain their employability given their updated knowledge and skills. The initial general education may decrease the costs of subsequent educational investments, and the updated skills may raise adaptability to technological change.

It is hard to come up with a full assessment of how life-time earnings compare between vocational and general education, which would allow for a judgement of whether the advantage of early employment is sufficient to make up for the later period of less employment for those with vocational education. But back-of-the-envelope calculations suggest that the present values of earnings (discounted at 3 %) is larger for the vocationally educated in Switzerland, but larger for the generally educated in Denmark and Germany. This cross-country pattern is consistent with the idea that those with general education are more capable to adapt to changing economic demands, as Denmark and Germany had noticeably higher growth rates of their economies over the relevant period than Switzerland.

3.2 New evidence from PIAAC

A significant drawback of the IALS analysis is that the results relate to the period of the mid-1990s. Across countries, labor markets have seen substantial transformations since then. As indicated above, globalization and technological change such as automation and digitalization have led to significant structural changes in Western economies (e.g., Autor, Dorn, & Hanson, 2015). These changes may have made the obsolescence of occupation-specific skills over the life-cycle even more pronounced in recent decades. However, there are also trends that may have weakened the differences in life-cycle implications across education types. In particular, reforms of the social security systems such as reduced options of generous early retirement schemes may have dampened the incidence of reduced employment at older ages,⁴ thereby lowering the scope for differential employment patterns by education type at older ages. Furthermore, education programs may have changed over time, with an altered extent to which the curricula of vocational programs contain general material.

To revisit the differential effects of vocational education in a more recent economic environment, Hampf and Woessmann (2017) use the data of the Programme for the International Assessment of Adult Competencies (PIAAC), conducted by the OECD in 2011/12. The dataset covers 29,000 adults in 16 countries. Apart from the larger sample size, the PIAAC data also provide a much richer testing of skills than IALS and further measures of family background, allowing for richer analyses to confirm that results are not driven by individual differences in literacy and numeracy skills and in family environments during youth.

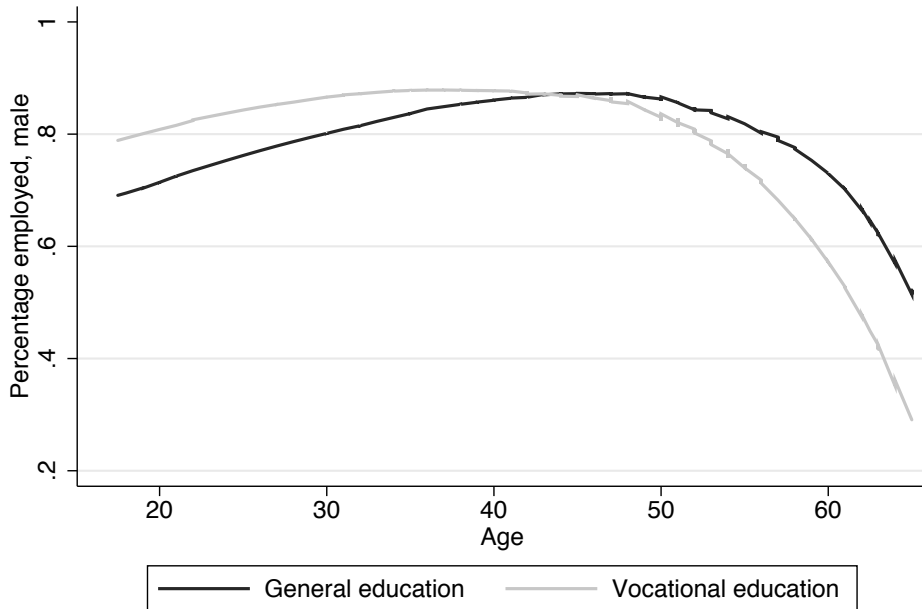
As is apparent from Figure 2, they find the same basic employment pattern of vocational and general education over the life-cycle: While vocationally educated individuals initially have better employment opportunities than generally educated individuals, this pattern turns around at older ages. Again, results are particularly strong in countries that have extensive apprenticeship systems – Austria, Denmark, and Germany in the case of the PIAAC data. As in the earlier study, results are robust to controlling for observed background characteristics including years of schooling, literacy and numeracy scores, parental education, and the share of the respective country cohort that has obtained a general education.

The estimates for the early 2010s are surprisingly similar in size to the prior estimates for the mid-1990s, although they tend to indicate a slightly earlier cross-over age by which individuals with a general education have higher employment probability than individuals with a vocational education. In the recent data, the cross-over age is around 50, and somewhat earlier around 45 in the apprenticeship countries. Thus, also after the significant transformation of labor markets due to globalization, digitalization, and pension reforms, vocational education seems to

4 For example, in Germany the share of those retiring before age 61 among all retirees has declined from 56 % in 1995 to 25 % in 2012 after reforms made the terms of early retirement less generous and raised the entitlement age for early retirement after a one-year unemployment spell from 60 to 63 years.

facilitate the school-to-work transition, but at the same time reduce the adaptability of older workers to changing environments.

Figure 2: Employment over the life-cycle with vocational and general education in PIAAC



Notes. Sample includes males who completed at least secondary education and are currently not students in the three “apprenticeship countries” (Austria, Denmark, and Germany), based on a matched sample that uses propensity-score matching to ensure common support between persons with a vocational and a general education in each country. Smoothed scatterplot using locally weighted regressions (Stata *lowess*). Source: Hampf and Woessmann (2017).

3.3 Further country-specific evidence

Several complementary country-specific investigations reinforce the general finding of a trade-off of vocational education over the employment life-cycle. To begin with, Hanushek et al. (2017b) provide two additional sets of analyses that strengthen the interpretation that the distinct age pattern reflects depreciated skills rather than other forces inducing people to leave employment. Using the much larger sample of the German Microcensus data, they show that the differential age-employment pattern by education type also holds when restricting the analysis to variation within tightly defined occupation groups and when excluding occupations in which brawn is important. This finding indicates that the differential movement out of employment is not simply a matter of physical wear and tear of people in specific vocationally intensive occupations.

In further analysis, they use the large administrative database of Austrian Social Security to identify job losses that occur due to plant closures. After such involuntary displacements, for workers of young ages the relative employment rates of displaced blue-collar workers (used as a proxy for vocational education in these data) are above those of white-collar workers. By contrast, when aged above 50, the relative employment rates of displaced blue-collar workers are below those of white-collar workers. Again, this pattern is consistent with an overall life-cycle trade-off between vocational and general education. More importantly, while the general results might in principle be partly affected by differences in unobserved retirement preferences between different types of workers, the exogenous nature of the employment shock in the plant closure analysis removes such concerns.

A number of additional recent country-specific studies show a similar age pattern of labor-market outcomes by education type over the life-cycle.⁵ Cörvers, Heijke, Kriechel, and Pfeifer (2011) provide life-cycle earnings patterns in vocational and general education for Germany, the Netherlands, and Great Britain. Consistent with the pattern described above, they show that earnings of vocationally educated individuals are higher at labor-market entry, but are overtaken by those of generally educated individuals with increasing age. The pattern is particularly pronounced in Germany, the country with a widespread apprenticeship system.

Relatedly, Weber (2014) estimates human capital depreciation rates for Switzerland and finds that depreciation rates are higher for vocational education than for academic education. Analyzing training programs for the unemployed in Sweden, Stenberg and Westerlund (2015) find that in the short run, specific training is related to higher earnings, but that the earnings of those who attended general training converge after a few years. By contrast, Hall (2016) is an exception that does not find a significant pattern based on the pilot of a Swedish reform in 1988-1993 that extended upper-secondary vocational programs by one year and increased their general content.

Two additional studies are of particular interest because they have longitudinal data to observe individual labor-market outcomes by education type over the life-cycle. Focusing on two-year programs at the upper secondary school level in Sweden, Golsteyn, and Stenberg (2017) show that vocational education is related to higher short-term earnings whereas general education is related to higher long-term earnings. Similarly, using British data, Brunello and Rocco (2017) find evidence of a trade-off between short-term advantages and long-term disadvantages of vocational education over the life-cycle for wages of individuals with lower secondary vocational education. The fact that these studies confirm a trade-off of labor-market outcomes by education type over the life-cycle with longitudinal data in Britain and Sweden indicates that the age differences analyzed by the studies above

5 In a contemporaneous study similar to Hampf and Woessmann (2017), Forster, Bol, and van de Werfhorst (2016) also find similar patterns with the PIAAC data. Brunello and Rocco (2015) use the PIAAC data to look at the effects of vocational education on adult skills and wages.

reflect actual age effects rather than cohort effects that are specific to different education types.

4. Policy implications for future-oriented apprenticeship systems

The consistent evidence on a life-cycle trade-off of vocational education provides the basis for a number of policy implications about how education systems can prepare students for lifetime work. In the following conclusions, we will place a particular focus on apprenticeship systems, their advantages and limitations, future reform requirements, and possibilities to emulate them from other countries.

4.1 Some general policy implications

At the most basic level, the findings indicate that in dynamic economies, policy needs to consider the full working life-cycle. They raise caution about policies that focus just on the current employment situation and ignore the dynamics of growing economies. Focusing just on the labor-market entry phase and youth unemployment is too narrow, as it ignores important countervailing effects later in life.

The crucial point is that no-one knows which specific skills will be demanded in the economy in, say, 30 years from now. The only thing that we can be sure of is that the economy will look very different from today's due to continuous technological and structural change. But today's secondary-school graduates will not even be 50 years old at that time, and they will want to strive on the labor market for at least another 15 years. This requires a strong educational foundation that provides workers with the ability to adapt as demands change.

As a consequence, neither vocational education nor general education are uniformly *good* or *bad*. There are pros and cons of both types of education programs. An advantage of the focus of vocational education programs on job-specific skills is that it may facilitate the transition from school to work. But focusing on job-specific skills also entails the risk of reducing the adaptability of workers to changing economic conditions. By contrast, general education programs may make it harder for graduates to find their way into the labor market, but foster their capability to adapt to change later in life. This insight implicates a number of further conclusions for the design of future-oriented education systems.

4.2 Securing a basis for adaptability

A key factor for education policy is to recognize the importance of the ability to adapt to changing economic conditions (cf. Nelson & Phelps, 1966; Welch, 1970;

Schultz, 1975 for early contributions). Education systems need to promote the kind of general skills that pay off in dealing with new economic processes. Recent cross-country evidence based on the PIAAC data suggests that labor markets strongly reward cognitive skills, and that these returns are particularly high in countries that have grown rapidly (see Hanushek, Schwerdt, Wiederhold, & Woessmann, 2015, 2017a). General skills are important to be able to cope with dynamic economic change.

This insight points to the importance of a high-quality school system. Many general skills that form the basis for further learning and adaptation are basic skills that are learned during childhood and adolescence. It is very hard to compensate failures of the primary and secondary school system later on. In particular, the lack of basic general skills cannot simply be replaced by learning job-specific skills in an apprenticeship. It is the strong general skill base that is required to strive on the labor market when things change.

One of the most significant labor-market problems facing many Western societies today is the number of workers whose middle-class jobs are slipping away and who are not prepared to adjust to a rapidly moving labor market. Societies are struggling to find ways of dealing with this problem. But expanding programs that focus on job-specific skills cannot substitute for the failure of the school system to provide the required general skill base. Failing to provide basic skills to the next generation would lock in the current middle-class malaise for the future and would likely make the long-run skill problems worse, as it would reproduce the current skill mismatch for future generations.

4.3 Providing the institutional frameworks required in apprenticeship systems

Beyond securing the general skill base, the hope of a smoothed school-to-work transition and ensuing reduced youth unemployment has enticed many countries to consider emulating the type of apprenticeship systems most common in German-speaking countries that combine learning at school and in the workplace. It is important to be aware, however, that the implementation of a functioning apprenticeship system requires many specific details of a strong institutional and regulatory framework that supports its success (cf., e.g., Wolter & Ryan, 2011).

A key element is that these labor markets put a strong emphasis on certification. For this, the degrees obtained from an apprenticeship need to send credible signals about students' acquired skills to potential employers. This requires not only a broad prevalence and publicity of the apprenticeship programs, but also clear certification requirements. An employer who considers hiring an applicant who received the apprenticeship training in another firm needs to be sure that the apprenticeship degree credibly certifies that the applicant has obtained the most important skills that will be required for performing the job. This is usually

achieved by a strong involvement of employers' associations in curriculum development and, in particular, final examinations.

More generally, successful apprenticeship systems tend to require strong involvement of both social partners – employers' associations and employees' unions – as well as the state. Apart from employers, the unions usually contribute to defining the apprenticeship curriculum and training requirements for firms and enforce restrictions and regulations on the labor market that facilitate the implementation of apprenticeships. There is also strong government involvement, as the state tends to coordinate the apprenticeship regulations with the social partners. Even more, the state tends to run and pay for the schools that apprentices have to attend for usually 1–2 days per week – the second part of the “dual” vocational system apart from the work-based training. In that sense, the financing of apprenticeship education is split between employers (who pay apprentice wages), apprentices (who accept lower wages during apprenticeship), and the state (which pays for schools).

Cost-benefit analyses of the German and Swiss apprenticeship systems indicate that the willingness of firms to train apprentices also depends on the extent to which they entail net benefits (production motive) or net costs (investment motive) during the apprenticeship phase (Muehleemann, Pfeifer, Walden, Wenzelmann, & Wolter, 2010). That is, the more likely it is that firms can recoup their training costs already during the apprenticeship phase and the more likely they are to retain apprentices in whom they have invested, the more likely it is that apprenticeship positions are in fact offered by the firms. Simulations of different scenarios of implementing apprenticeship programs in Spain based on the Swiss experience suggest that the expected net costs of offering apprenticeships for Spanish firms strongly differ by the specific occupation, training scenario, and firm size (Muehleemann & Wolter, 2017). For any country that considers emulating an apprenticeship system, this suggests that finding a particular balance between benefits and costs for each stakeholder – employers, apprentices, and the state – is an intricate part of a successful apprenticeship system.

The consideration of stakeholder tasks, financial structures, and certification requirements indicates that apprenticeship systems require institutional and regulatory frameworks that often took decades to emerge and consolidate. The German-speaking countries can build on a specific tradition that allows for the functioning of their apprenticeship systems. Emulating these systems requires the implementation of institutions and regulations that similarly make the apprenticeship system worthwhile both for school graduates and for employers.

4.4 “Generalizing” apprenticeship education

For countries that have an apprenticeship system in place or are considering implementing one, it is important to take into consideration what is required to make graduates fit for employment over their full life-cycles. The presented evidence sug-

gests that reducing the early specialization of students on specific occupational skills may be conducive to their long-run prospects on the labor market. Based on observations of the German apprenticeship system, it may be worthwhile to consider designs and reforms that reduce the early specialization of apprentices. This could entail initiatives to lower the number of specific apprenticeships, expand the share of general educational content, and modularize apprenticeship components.

Countries with extensive apprenticeship systems differ substantially in the number of specific apprenticeships offered. While Germany has substantially more than 300 specific apprenticeship occupations, Switzerland gets by with much fewer specific apprenticeships, despite an otherwise quite similar system. For example, Germany has more than 30 specific apprenticeships in the commercial area, from special programs for commercial apprentices working in the courier and mail business to special programs for commercial apprentices working in the logistics business. By contrast, Switzerland has moved away from such an especially fine-grained system and offers one apprenticeship for all commercial trainees. They follow the same school curriculum for the first two years of their apprenticeship and specialize in their specific sector only during the third year. From the point of view of adaptability to changing economic environments, the latter system seems more likely to equip apprentices with skills that make them employable over their full life-cycle which often requires changing sectors. More generally, lowering the number of specific apprenticeships is likely to raise the chances that graduates obtain skills that are demanded in broader sections of the labor market.

Another means to ensure that apprenticeship graduates also obtain general skills that facilitate adaptability to change is to ascertain that the curriculum of the vocational schools contains a sufficient share of general educational content. Apart from teaching job-specific skills, these schools should also convey a basis of general skills that go beyond narrow occupational areas and facilitate later learning. Such skills may help graduates also later on when the requirements of the economy have changed. The “generalization” of skills obtained during an apprenticeship might also entail the modularization of apprenticeship components, so that apprentices in different occupations in part take the same training modules. Such a setup could facilitate the movement of graduates into different occupations later on.

Thus, there are a number of options to reduce the early specialization of apprentices and increase their general skills when implementing an apprenticeship system or reforming it to make graduates fit for the future.

4.5 Strengthening lifelong learning

Finally, an important aspect for the long-term employability of vocationally educated workers is a strong focus on lifelong learning. Considering the specificity of the skills obtained during an apprenticeship, graduates of vocational programs are in particular need to retrain when the economy changes. Unfortunately, quite to the contrary, data from several apprenticeship countries suggest that apprenticeship

graduates are less likely to participate in adult education than graduates from general education programs as they become older (Hanushek et al., 2017b). This may be partly due to the fact that general education programs develop skills that facilitate later learning throughout one's career. In addition, as long as an employee stays with the same company, the company is likely to offer the kind of job-specific courses that update the worker's skills in the particular profession. However, companies obviously do not have an incentive to provide the kind of training courses that would prepare their employees to work in a different profession, something that is increasingly needed in changing economies.

As a consequence, it may be worth considering the establishment of a system for lifelong learning that does not only update workers' skills within their occupation but also conveys skills that facilitate their flexibility if changing labor-market conditions require occupational change. That is, it is important to strengthen lifelong learning for graduates of vocational education programs on a broader scale. These requirements will become increasingly severe as we move more into being a knowledge economy and as digitization and automation of routinized job tasks require adaptability to change.

References

- Akerman, A., Gaarder, I., & Mogstad, M. (2015). The skill complementarity of broadband internet. *Quarterly Journal of Economics*, 130(4), 1781–1824.
- Autor, D. H., Dorn, D., & Hanson, G. H. (2015). Untangling trade and technology: Evidence from local labour markets. *Economic Journal*, 125(584), 621–646.
- Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4), 1279–1333.
- Brunello, G., & Rocco, L. (2015). *The effects of vocational education on adult skills and wages: What can we learn from PIAAC?* OECD Social, Employment and Migration Working Papers 168. Paris, France: OECD.
- Brunello, G., & Rocco, L. (2017). The labor market effects of academic and vocational education over the life cycle: Evidence from two British cohorts. *Journal of Human Capital*, 11(1), 106–166.
- Cörvers, F., Heijke, H., Kriechel, B., & Pfeifer, H. (2011). *High and steady or low and rising? Life-cycle earnings patterns in vocational and general education*. ROA Research Memorandum ROA-RM-2011/7. Maastricht, Netherlands: Research Centre for Education and the Labour Market.
- Forster, A. G., Bol, T., & van de Werfhorst, H. G. (2016). Vocational education and employment over the life cycle. *Sociological Science*, 3, 473–494.
- Golsteyn, B. H. H., & Stenberg, A. (2017). Earnings over the life course: General versus vocational education. *Journal of Human Capital*, 11(2), 167–212.
- Hall, C. (2016). Does more general education reduce the risk of future unemployment? Evidence from an expansion of vocational upper secondary education. *Economics of Education Review*, 52, 251–271.
- Hampf, F., & Woessmann, L. (2017). Vocational vs. general education and employment over the life cycle: New evidence from PIAAC. *CESifo Economic Studies*, 63(3), 255–269.

- Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, 73, 103–130.
- Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2017a). Coping with change: International differences in the returns to skills. *Economic Letters*, 153, 15–19.
- Hanushek, E. A., Schwerdt, G., Woessmann, L., & Zhang, L. (2017b). General education, vocational education, and labor-market outcomes over the life-cycle. *Journal of Human Resources*, 52(1), 48–87.
- Katz, L. F., & Autor, D. H. (1999). *Changes in the wage structure and earnings inequality*. In O. C. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Volume 3, Part A, pp. 1463–1558). Amsterdam, Netherlands: Elsevier.
- Krueger, D., & Kumar, K. B. (2004). Skill-specific rather than general education: A reason for US-Europe growth differences? *Journal of Economic Growth*, 9(2), 167–207.
- Muehlemann, S., Pfeifer, H., Walden, G., Wenzelmann, F., & Wolter, S. C. (2010). The financing of apprenticeship training in the light of labor market regulations. *Labour Economics*, 17(5), 799–809.
- Muehlemann, S., & Wolter, S. C. (2017). Can Spanish firms offer dual apprenticeships without making a net investment? Empirical evidence based on ex ante simulations of different training scenarios. *Evidence-based HRM*, 5(1), 107–118.
- Nelson, R. R., & Phelps, E. (1966). Investment in humans, technology diffusion and economic growth. *American Economic Review*, 56(2), 69–75.
- Psacharopoulos, G. (1987). To vocationalise or not to vocationalise: That is the curriculum question. *International Review of Education*, 33(2), 187–211.
- Ryan, P. (2001). The school-to-work transition: A cross-national perspective. *Journal of Economic Literature*, 39(1), 34–92.
- Schultz, T. W. (1975). The value of the ability to deal with disequilibria. *Journal of Economic Literature*, 13(3), 827–846.
- Shavit, Y. & Müller, W. (Eds.). (1998). *From school to work: A comparative study of educational qualifications and occupational destinations*. Oxford, England: Clarendon Press.
- Stenberg, A., & Westerlund, O. (2015). The long-term earnings consequences of general vs. specific training of the unemployed. *IZA Journal of European Labor Studies*, 4(1), 1–26.
- Weber, S. (2014). Human capital depreciation and education level. *International Journal of Manpower*, 35(5), 613–642.
- Welch, F. (1970). Education in production. *Journal of Political Economy*, 78(1), 35–59.
- Wolter, S. C., & Ryan, P. (2011). Apprenticeship. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Volume 3, pp. 521–576). Amsterdam, Netherlands: Elsevier.
- Zimmermann, K. F., Biavaschi, C., Eichhorst, W., Giulietti, C., Kendzia, M. J., Muravyev, A., Pieters, J., Rodríguez-Planas, N., & Schmidl, R. (2013). Youth unemployment and vocational training. *Foundation and Trends in Microeconomics*, 9(1-2), 1–157.

Nils Berkemeyer, Björn Hermstein, Sebastian Meißner &
Ina Semper

Kritische Schulsystementwicklungsforschung: Ein normativ-analytischer Forschungsansatz der schulischen Ungleichheitsforschung

Zusammenfassung

Soziale Ungleichheiten im Schulsystem sind ein zentraler Untersuchungsgegenstand der Ungleichheitsforschung und des Bildungsmonitorings. Die gegenwärtigen Forschungszugänge zeichnen sich, trotz ihrer methodischen Elaborierung, durch eine normative Unterbestimmtheit aus. Vor diesem Hintergrund plädiert der Beitrag für eine gesellschafts- und gerechtigkeits-theoretisch fundierte Weiterentwicklung der gegenwärtigen Schulentwicklungsforschung hin zu einer Kritischen Schulsystementwicklungsforschung, die sich als ein Programm zur Integration von empirischer Wirkungsforschung, Systemmonitoring sowie Schultheorie auf Basis ethisch-moralischer Referenzkategorien versteht. Eine Kritische Schulsystementwicklungsforschung (KSEF) kann zukünftig einen Rahmen bieten, um empirische Beobachtungskategorien einerseits an theoretisch gewonnenen, ethisch-moralischen Bewertungsmaßstäben zu diskutieren, andererseits, um sie auf Basis gerechtigkeits-theoretischer Implikationen weiterzuentwickeln. Am Beispiel des Nationalen Bildungsberichts werden die analytischen Potenziale einer Kritischen Schulsystementwicklungsforschung in einer ersten Exploration hinsichtlich der theoriegeleiteten Bewertung und Weiterentwicklung schulsystemischer Beobachtungskategorien aufgezeigt. Es zeichnen sich abschließend erste Konturen einer Ethik des Schulsystems ab, die Schulsysteme stärker als bislang als ethisch und moralisch umfassend zu reflektierende Gegenstände grundlegt.

Prof. Dr. Nils Berkemeyer (corresponding author) · Sebastian Meißner M.A. · Ina Semper M.A., Institut für Erziehungswissenschaft, Lehrstuhl für Schulpädagogik und Schulentwicklung, Friedrich-Schiller-Universität, Fürstengraben 11, 07743 Jena, Deutschland
E-Mail: nils.berkemeyer@uni-jena.de
sebastian.meissner@uni-jena.de
ina.semper@uni-jena.de

Björn Hermstein M.A., Institut für Schulentwicklungsforschung (IFS) der TU Dortmund, Vögelpothsweg 78, 44227 Dortmund, Deutschland
E-Mail: bjoern.hermstein@tu-dortmund.de

Schlagworte

Soziale Ungleichheit; Gerechtigkeitstheorien; Kritische Schulsystementwicklungsforschung; Bildungsmonitoring; Normativität

Critical educational system development research: A normative-analytical research approach to school-based inequality research

Abstract

It is a main subject of educational monitoring and inequality research to explore social inequalities. But despite of elaborated research methods it should be noted that these studies are usually not informed by normative reflections. So this article argues for a further development from educational system development research to an integrated theory of inequality research and justice-based school theory: critical educational system development research. Theory-based valuation standards may be used to discuss results of educational inequality research as well as to develop criteria of empirical observation by ethical principles. We will illustrate the critical potential of this programme by the example of the national report on education ("Education in Germany").

Keywords

Social inequality; Theories of justice; Critical educational system development research; Educational monitoring; Normativity

1. Einleitung und Problemskizze

Die schulische Ungleichheitsforschung wird aktuell durch Ansätze dominiert, die empirisch (Becker, 2017; Becker & Lauterbach, 2016; Becker & Schulze, 2013; Solga & Becker, 2012; Maaz, Baumert & Trautwein, 2011), nicht aber normativ hinreichend elaboriert sind. Dies trifft insbesondere für die schulentwicklungsrelevanten Elemente des Schulsystemmonitorings zu, die von der Kultusministerkonferenz der Länder (KMK) im Anschluss an die ersten PISA-Befunde beschlossen wurden (KMK, 2002). Die normative Unterbestimmtheit bezieht sich dabei erstens auf die Bewertung der beobachteten Ungleichheiten, zweitens auf die für die Schulentwicklung konstitutive Frage nach Strategien der Veränderung, sowie drittens auf die Beobachtungskategorien, die bei der Feststellung sozialer Ungleichheit Verwendung finden und die ihrerseits begründungsbedürftig sind. Arbeiten, die soziale Ungleichheiten gerechtigkeits theoretisch interpretieren, finden sich nur vereinzelt (Wigger, 2015; Stojanov, 2011; Brenner, 2010; Giesinger, 2007). Eine solche Arbeitsteilung zwischen beobachtender empirischer Bildungsforschung und einigen wenigen bewertenden Beiträgen, die der Bildungsphilosophie oder einer gesellschaftstheoretisch reflektierten Schultheorie zugeordnet werden können,

ist nicht untypisch. Sie kann als eine strukturelle methodologische Schwäche der Erziehungswissenschaft beschrieben werden, die aber auch für andere Disziplinen zu konstatieren ist. Es ist leicht zu sehen, dass eine produktive Bezugnahme der beiden Diskurse sinnvoll ist. Während es relativ offensichtlich ist, dass soziale Ungleichheiten nicht grundsätzlich ungerecht sein müssen, besteht die Herausforderung darin, genau diesen Zusammenhang zu spezifizieren und Überlegungen anzustellen, die zeigen, welche Ungleichheiten im Schulsystem vor welchen Hintergrundannahmen als akzeptabel beurteilt werden können. Darüber hinaus ist es auch einleuchtend, dass die Beobachtungskategorien, die bei der Feststellung sozialer Ungleichheit Verwendung finden, ihrerseits begründungsbedürftig sind. Diskussionswürdig ist aber auch die Frage, was unter einem gerechten Schulsystem insgesamt verstanden werden kann und mit welchen Strategien dieses zu erreichen ist.

Die Schulentwicklungsforschung sollte die dargestellte Lücke ausfüllen, ohne gleichsam genuine Ungleichheitsforschung sein zu können. Sie hat bereits Erfahrungen in Bezug auf Integrationsversuche von Paralleldiskursen. Um die 2000er Jahre gab es eine intensive Diskussion um die Zusammenführung von Forschungen zu *school effectiveness* und *school improvement*. Ziel war es u. a., belastbare Wirkungszusammenhänge mit Schulentwicklungsstrategien zu koppeln (Bonsen, Bos & Rolff, 2008; Klieme, 2016). Projekte wie *Selbstständige Schule in Nordrhein-Westfalen* (Holtappels, Klemm & Rolff, 2008) und *Schulen im Team* (Berkemeyer, Bos, Manitius & Müthing, 2008) sind in dieser Diskussionslinie zu verstehen. Die Herausforderung bei der Integration beider Zugänge bestand und besteht nicht zuletzt in der Verbindung von Empirie und Theorie, also zum einen in der Anreicherung der empirischen Befunde zu Schuleffektivitätsfaktoren durch Theorien der Schulentwicklung und zum anderen in einer durch Empirie weitergehenden Theoretisierung von Prozessen der Schulentwicklung. Forschungen dieses Typs sind vor allem evaluative Analysen zu und innerhalb von Programmen der Schulentwicklung, die den normativen Zielen und den normativen Implikationen der verwendeten Change-Strategien weitgehend unkritisch gegenüberstehen. Vielmehr geht es meist darum, Aspekte innerhalb des Programms zu verbessern. Somit ist dieser Forschungstypus in gewisser Weise affirmativ in Bezug auf die bestehenden Ordnungen des Schulsystems (Bellmann & Müller, 2011; Berkemeyer, 2017).

Die Schulentwicklungsforschung, so soll hier argumentiert werden, sollte jedoch eine Teildisziplin ausbilden, die die Integrationsleistung von Forschungen zu *school effectiveness* und *school improvement* zu ihrem Ausgangspunkt macht und auf dieser Grundlage gesellschaftstheoretisch weiterentwickelt. Die Herausforderung wird dabei künftig sein, die Schulentwicklungsforschung zu einer kritischen Disziplin weiterzuentwickeln bzw. ihren kritischen Gehalt wiederzugewinnen (Berkemeyer, 2016; Berkemeyer & Hermstein, 2018), ohne dabei aber die erzielten Fortschritte einem rein akademisch geführten theoretisch-normativen Spezialdiskurs zu opfern.

Um diese mögliche Entwicklung zu beschreiben, wird zunächst daran erinnert, dass Schulentwicklungsforschung, will sie als eine Variante einer gesellschaftstheoretisch angelegten Schultheorie eine kritische Funktion ausüben, auf moralisch-

ethische Referenzkategorien angewiesen ist. Als eine solche Basis für Beurteilungen von empirischen Befunden zur sozialen Ungleichheit im Schulsystem werden nachfolgend gerechtigkeits-theoretische Konzepte eingeführt, die als wertbezogene Deutungsangebote die kulturtheoretische Seite gesellschaftstheoretischen Denkens betonen (diskutiert werden im Kontext von Gesellschaftstheorie weiterhin Differenzierungs- und Ungleichheitstheorien; Schimank, 2015). Exemplarisch zeigen wir die von uns bislang nur behaupteten, analytischen Potenziale anhand des Nationalen Bildungsberichts als einem diskursmächtigen Beispiel des Bildungsmonitorings auf. Wir schließen mit einem Blick auf zukünftige Arbeitsaufgaben einer Kritischen Schulsystementwicklungsforschung (KSEF) sowie Perspektiven der Weiterentwicklung in Richtung einer Ethik des Schulsystems. Damit greifen wir auch eine aktuelle Forderung auf, die „mehr intellektuelle Beweglichkeit im bildungspolitischen und empirisch-bildungswissenschaftlichen Disparitätendiskurs [anmahnt]“ (Cortina & Pant, 2018, S. 78).

2. Schulsystementwicklungsforschung und moralische Urteile

In Abgrenzung zur Schulentwicklungsforschung nutzen wir im Folgenden den Begriff der Schulsystementwicklungsforschung, die über die Betrachtung der Einzelschule hinausgeht und sich auf das Schulsystem und dessen Kopplungssysteme (Luhmann, 2004, S. 209ff.) – Bildungspolitik und -verwaltung, Recht, Arbeitsmarkt und die relevanten Lebenswelten wie Familie – als Ganzes bezieht und damit die konstitutive Wechselwirkung zwischen Schule und Gesellschaft zentral stellt. Hans-Günter Rolff und Klaus-Jürgen Tillmann haben diese Zusammenhänge und Implikationen bereits 1980 in einem für die Schulsystementwicklungsforschung zentralen, programmatischen Aufsatz dargestellt (Rolff & Tillmann, 1980). Unter Rückgriff auf die Curriculumtheorie, die Sozialisationsforschung und die Institutionenanalyse sollte die Schule als Ganzes analysiert werden. Insbesondere die Institutionenanalyse hatte hierbei eine herausragende Bedeutung: „Sie vermag Struktur- und Handlungszusammenhänge des innerschulischen Geschehens aufzuklären, seien es die Qualifikations- oder Legitimationsfunktion, die Statusallokation über Bildungsabschlüsse oder der Beitrag zur sozialen Reproduktion insgesamt“ (Rolff & Tillmann, 1980, S. 251).

Ziel sollte es sein, die Entwicklungen und Veränderungen von Schule und Gesellschaft methodisch sauber zu erfassen, miteinander ins Verhältnis zu setzen und so „Aufschlüsse über das wechselseitige Implikationsverhältnis“ (ebd.) zu erlangen. Die Aufgabe, Schule und Gesellschaft aufeinander zu beziehen, bleibt sicherlich von zentraler Bedeutung. Was Rolff und Tillmann aber noch nicht konzeptionell angelegt hatten, ist die Tatsache, dass eine Institutionenanalyse unweigerlich auch den Bereich der politischen Ethik berührt. Politische Ethik zeichnet sich als Bereichsethik dadurch aus, dass der „Beurteilung politischer

Institutionen und/oder politischen Handelns ein Primat oder jedenfalls eine gewisse Unabhängigkeit [zuerkannt wird]. Die klassische Frage ist dabei die nach der Gerechtigkeit politischer Ordnungen und politischer Institutionen“ (Nida-Rümelin, 2016, S. 12).

Mit Ernst Tugendhat kann auch davon gesprochen werden, dass es in der politischen Moral darum geht, über das „Gut- oder Schlechtsein eines Staates“ (Tugendhat, 2012, S. 337) zu befinden. Das ethische Urteil ist dabei zwar nicht empirisch begründbar (ebd., 1. Vorlesung), aber es ist auch nicht willkürlich (Pfordten & Nida-Rümelin, 2015; Nida-Rümelin, 2002, 2006). Nur durch diese Integration einer ethischen und moralischen Perspektive kann eine Institutionenanalyse ihren eigentlichen Wert erhalten (siehe hierzu in Bezug auf soziale Ungleichheit im Schulsystem Hügli, 1999, S. 179–184). Hierfür lassen sich zumindest zwei zentrale Gründe anführen:

- 1) Es ist abwegig, Schulsystementwicklungsforschung zu betreiben und in Bezug auf die Bewertung der empirisch gewonnenen Aussagen auf eine Arbeitsteilung zwischen Wissenschaft einerseits und Politik andererseits zu setzen und dadurch das kritische Potenzial aus der Forschung zu suspendieren. Wissenschaft ist verpflichtet, die Normativität der Themenwahl, die Selektivität der Beschreibung, auch in methodischer Hinsicht, darzulegen (Dahrendorf, 1962) und die so gewonnenen Befunde in einen reflektierten Deutungshorizont zu stellen, der keineswegs zu Vorschriften führt, aber mögliche Lesarten und Grundlagen einer ethischen Bewertung zur Verfügung stellt – hier gilt Rortys (1988) Vorrang der Demokratie vor der Philosophie. Es geht also darum, zu einem Aussagentypus zu kommen, der folgender Form folgt: Die Veränderungen von Schule S und Gesellschaft G werden unter der Perspektive von Werten W_1 - W_n wie folgt gewertet. W_1 ist W_2 dabei aus folgenden Gründen zur Beurteilung der Veränderungen von S und G vorzuziehen. Solch qualifizierte Gründe haben dann zwar keinen Vorrang, aber sie sind in der Deliberation schwieriger zu verwerfen.
- 2) Eine kritische, auf gesellschaftstheoretischen Arbeiten gründende Analyse reicht zur Beurteilung der beobachteten Sachverhalte allein nicht aus (vgl. Tugendhat, 2012, S. 17ff.). Beschreibungen von Ungleichheitszuständen beispielsweise reichen demnach nicht aus, um ein System normativ in Frage zu stellen. Empirie und normatives Urteil dürfen nicht vermischt werden! Gerade darum ist aber eine explizite normative Position notwendig, und dies muss eine in der Ethik begründete Position sein (siehe hierzu die instruktive Reflexion zum Werturteilsstreit bei Dahrendorf, 1962). Kritik ist damit konstitutiver Bestandteil einer Schulsystementwicklungsforschung und keineswegs ein von Empirie oder Theorie der Schule zu trennender Sachverhalt.

Eine kritische Schulsystementwicklungsforschung, wie sie sich hier bereits andeutet, fokussiert demnach im Besonderen die Veränderungsprozesse von Schule und Gesellschaft, und verbindet sie mit ethischen und gesellschaftstheoretischen Überlegungen einer politischen Philosophie, die nach einer gerechten

Gesellschaftsordnung und damit verbundenen gerechten Institutionen (Rawls, 1979, 2006) sowie an Achtung (Tugendhat, 2012, S. 79ff.) und Anerkennung (Honneth, 2003) orientierten Praktiken fragt, die ihrerseits als Praktiken der bzw. innerhalb der Demokratie unter einem Rechtfertigungsprimat stehen müssen (Forst, 2015). Eine recht verstandene Schulsystementwicklungsforschung muss diese Perspektive auch deshalb einnehmen, weil es zu prüfen gilt, inwieweit Schule zur Sicherung des Staatsziels beiträgt, das in der Realisierung und Wahrung eines demokratischen und sozialen Bundesstaates besteht, wie in Art. 20 GG festgeschrieben (Richter, 1995; Reichenbach, 2018; Berkemeyer, im Ersch.). Dies ist insbesondere für die Fragen der Beobachtung (Monitoring) und der Evaluation von immenser Bedeutung, da sich erst aus einer umfassenden und sachgemäßen Erfassung des Zusammenhangs von Schule, Staat und Gesellschaft die Beobachtungsparameter und die Zielgrößen einer Evaluation ableiten lassen. Darum ist es notwendig, Schulsystementwicklungsforschung auch als eine kritische Disziplin aufzufassen, die sich dadurch auszeichnet, die normativen Leerstellen durch eine Verschränkung der Diskurse um soziale Ungleichheit und soziale Gerechtigkeit zu füllen.¹

3. Welche Gerechtigkeit? Widersprüche oder eine Frage der Perspektive und Beobachtung?

Mit der Neuzeit wird Gerechtigkeit zunehmend als Ordnungsstiftung durch den Menschen und nicht mehr als Abbild einer göttlichen Ordnung gedacht (Heidenreich, 2010, S. 79). Damit stellt sich die Frage, ob Gerechtigkeit lediglich eine Sache der Aushandlung ist, wie in Vertragstheorien der Gerechtigkeit (Rawls, 1979) oder Verfahrenstheorien (Habermas, 1992), oder ob es möglich ist, Prinzipien der Gerechtigkeit zu erkennen (moralischer Realismus, moralischer Naturalismus). Solche und andere Fragen der Metaethik (Birnbacher, 2013; Stahl, 2013; Rüter, 2015) müssen hier zurückgestellt werden, ohne dass sie damit zurückgewiesen werden sollen. Für diesen Beitrag genügt es, auf unterschiedliche Gerechtigkeitsprinzipien, die sich grundsätzlich zwischen den normativen Polen Freiheit und Gleichheit bewegen (Schink, 2017; Höffe, 2015b; Rawls, 2006; Honneth, 1993), und ihre möglichen Anwendungen hinzuweisen. Dabei geht es nicht um eine voreilige Festlegung auf eine bestimmte Theorie, sondern vielmehr darum, die analytische Leistungsfähigkeit spezifischer Ansätze für eine kritische Schulsystementwicklungsforschung herauszustellen. Analysiert werden muss das Schulsystem dabei mindestens in Bezug auf Gesellschaft, seine institutionelle

1 Jenseits von gerechtigkeitstheoretischen Perspektiven könnten wir auch alternative Zugänge zur moralischen Beurteilung der empirischen Befunde zum Schulsystem heranziehen. Neben utilitaristischen Positionen seien auch solche der Tugendethik erwähnt. Allerdings erweisen sich beide Perspektiven für eine pädagogisch orientierte kritische Schulsystementwicklungsforschung als nicht zielführend (vgl. zu den jeweiligen Ansätzen und ihren Vor- und Nachteilen Nida-Rümelin, 2015; Höffe, 2015a, 2013; Hübner, 2014).

Verfasstheit, Regeln und Verfahren, seine Praktiken und die durch das Schulsystem moderierten Ergebnisse.

3.1 Worauf kommt es an? Verteilung von Grundgütern oder Gewährleistung von personalem Grundvermögen als leitendes Prinzip der Gerechtigkeit?

John Rawls, der eine linksliberale, auf egalitaristischen Überzeugungen basierende, vertragstheoretische Gerechtigkeitstheorie entwirft (Rawls, 1979, 2006), die als Zäsur im Nachdenken über soziale Gerechtigkeit gilt (Höffe, 2013; Kersting, 2008), entwickelt zwei basale Prinzipien einer demokratisch legitimierbaren Verfahrensgerechtigkeit:

- 1) Jede Person hat den gleichen unabdingbaren Anspruch auf ein völlig adäquates System gleicher Grundfreiheiten, das mit demselben System von Freiheiten für alle vereinbar ist.
- 2) Soziale und ökonomische Ungleichheiten müssen zwei Bedingungen erfüllen: Erstens müssen sie mit Ämtern und Positionen verbunden sein, die unter Bedingungen fairer Chancengleichheit allen offenstehen; und zweitens müssen sie den am wenigsten begünstigten Angehörigen der Gesellschaft den größten Vorteil bringen (Differenzprinzip). (Rawls, 2006, S. 78)

Das Differenzprinzip scheint für die Diskussion über gerechte Schulsysteme besonders bedeutsam: Rawls folgend, müsste die Institution Schule so gestaltet sein, dass vor allem die Schülerinnen und Schüler mit geringen Lernkapazitäten (Risikogruppe) den größten Nutzen haben. Denn für Rawls steht, wie für viele andere Gerechtigkeitstheoretikerinnen und -theoretiker auch, fest, dass Freiheit immer eine Herausforderung für die Gleichheit darstellt. Sie gefährdet diese und ist zugleich doch auch notwendig (Nida-Rümelin, 2006, S. 138; Menke, 1995; Nagel, 2016). Auf Schule übertragen, muss die Freiheit der Potenzialentfaltung gewährleistet (nicht erzwungen) und zugleich darauf geachtet werden, dass die Ungleichheit nicht soweit anwächst, dass eine solidarische Gemeinschaft gefährdet wird. Aktuell dürfte dies, angesichts des Befunds differentieller Lernmilieus (Baumert, Stanat & Watermann, 2006) sowie der selektiven Folgeeffekte von Schulreformen (Hermstein, Berkemeyer, Bos & Semper, 2019), sogar gegenläufig sein: Die Institution begünstigt die ohnehin schon überlegenen Schülerinnen und Schüler, die zudem in organisatorischen Einheiten (Gymnasien) lernen, die teurer sind. Damit stellt sich insgesamt die Frage nach der Verteilung des Gutes Bildung, wobei die Konzeption von Bildung als Gut selbst nicht unproblematisch ist (siehe hierzu ausführlich Giesinger, 2011; zum Grundgüterbegriff allgemein Rawls, 2006, S. 259ff.). Kompensatorische Verteilungsmodelle kommen auch in Deutschland zunehmend zum Einsatz, ohne dass ihre Wirkung schon hinreichend untersucht und die Art der gewählten Verteilungsschlüssel hinreichend begründet worden sind (Fickermann & Maritzen, 2014; Groot-Wilken, Isaac & Schräpler, 2016). Von ei-

ner Verteilung nach dem Leistungsprinzip, was insbesondere im Schulsystem häufig als Verteilungsprinzip gehandelt wird (vgl. auch Nerowski, 2018), rät Rawls insgesamt ab.²

Amartya Sen (2010) entwickelt seine Theorie der Gerechtigkeit vor dem Hintergrund der Rawls'schen Überlegungen, kritisiert diesen aber mitunter heftig. Eine wesentliche Kritik besteht darin, dass eine juristische Fassung von Institutionengerechtigkeit, wie sie auch von Dworkin (2011, 2012) oder Gauthier (1999) vertreten werden, nicht hinreichend sei. Vielmehr komme es ihm darauf an, dass die Menschen faktisch ein gutes Leben führen können, was er als Realisierung faktischer Freiheit beschreibt (Sen, 2010, S. 331ff.). Um dies zu garantieren, reiche es eben nicht aus, formal gerechte Institutionen vorzuhalten, die eine Verteilung von Grundgütern gewährleisten, vielmehr müssten die Menschen auch tatsächlich befähigt werden (Grundvermögen), ihre Freiheiten zu nutzen, womit er an einen positiven Freiheitsbegriff anschließt. Menschen müssten demnach nicht nur frei von zum Beispiel Zwang sein (negative Freiheit; siehe hierzu Berlin, 1995; Taylor, 1992), sondern auch frei, um etwas tun zu können. Eben dieser Aspekt seines Ansatzes ist vor allem in der Sozialpädagogik stark rezipiert worden (Ziegler, 2011), aber auch für die Schule relevant, denn selbst wenn die Schule rechtlich gerecht verfasst wäre (was formaljuristisch nicht unwahrscheinlich ist), müsste sie hinsichtlich ihrer Befähigungspotenziale befragt werden und zeigen, was sie dazu beiträgt, dass Menschen ein besseres Leben führen können bzw. welche Fähigkeiten zur Nutzung menschlicher Freiheiten durch die Teilnahme an Schule ermöglicht und realisiert werden, wobei die Bedingungen, unter denen dieser Erwerb erfolgt, selbst gerecht und fair sein müssen (Sen, 2010, S. 38ff.). Gerechte Verteilung ist in diesem Sinne eine Funktion der Nutzung faktischer Freiheit, die durch die Befähigung der jeweiligen Subjekte garantiert werden soll. Insbesondere bei Angehörigen der sogenannten Risikogruppe, die sich auf Kompetenzstufe 1 oder niedriger befinden, scheint eine solche Befähigung nicht zu gelingen, sodass Subjekte zwar in einem negativen Sinne frei sind, aber eben nicht frei, um etwa ihr eigenes Leben umfänglich zu gestalten. In theoretischer Hinsicht sei abschließend darauf hingewiesen, dass Rawls die Kritik Sens aufgreift und beansprucht, diese weitgehend aufzulösen (Rawls, 2006).

3.2 Unterschiedliche Gerechtigkeitsprinzipien für unterschiedliche gesellschaftliche Sphären?

Nicht wenige Gerechtigkeitstheorien kritisieren die Vorstellung universell gültiger Gerechtigkeitsprinzipien, die häufig in einer kantianisch-liberalen Tradition stehen (Höffe, 2015a). Ein Argument der Kritiker hierfür findet sich im diagnostizierten Wertpluralismus, der sich einerseits auf die jeweiligen Gemeinschaften

2 Sowohl das PISA-Konsortium als auch der Aktionsrat Bildung beziehen sich zustimmend auf das Leistungsprinzip (kritisch dazu Wigger, 2015, S. 78; kritisch zum Leistungsprinzip allgemein Heid, 2012; Berkemeyer, im Ersch.).

(Taylor, 2012) und andererseits auf die jeweiligen gesellschaftlichen Teilsysteme bzw. gesellschaftliche Sphären (Walzer, 2006) bzw. Vergesellschaftungszonen (Miller, 2008) bezieht.³ Solche Theorien stehen in einer hegelianischen Tradition, die Gesellschaft in ihrer gewordenen Sittlichkeit und damit Historizität begreift (Honneth, 2010). Zudem werden stärkere Bezüge zu Fragen der Macht und ihres begrenzenden Einflusses auf Gerechtigkeit hergestellt, beispielsweise, wenn Walzer unter Rückgriff auf Pascal die Frage von Gerechtigkeit und Herrschaft diskutiert (Walzer, 2006, S. 47ff.). Während Walzer und Miller distributive Gerechtigkeitstheorien vorlegen, betont Honneth die Bedeutung sozialer Freiheit als Gerechtigkeit durch Anerkennung.

3.2.1 Komplexe Gleichheit

Walzer (2006) schlägt ein pluralistisches Modell komplexer Gleichheit vor, welches sich dadurch auszeichnet, dass es elf *Distributionssphären* unterscheidet, die je durch eines von drei Verteilungsprinzipien gerecht geregelt werden: a) freier Austausch, b) Verdienst und c) Bedürfnis. Sodann ergibt sich für ihn die Forderung, dass Vorteile in einer Sphäre nicht in eine andere übersetzt werden dürfen bzw., dass sie nicht zu Vorteilen oder Begünstigungen führen dürfen (ebd., S. 48). Entsprechend werden im Konzept *komplexer Gleichheiten* (Walzer, 2006) sphärenspezifische Gerechtigkeitsprinzipien nach dem Gleichheitsgrundsatz formuliert.

„Das System der komplexen Gleichheit ist das Gegenteil von Tyrannei. Es erzeugt ein Netz von Beziehungen, das Dominanz und Vorherrschaft verhindert. Formal gesprochen bedeutet komplexe Gleichheit, dass die Position eines Bürgers in einer bestimmten Sphäre oder hinsichtlich eines bestimmten sozialen Guts nicht unterhöhlt werden kann durch seine Stellung in einer anderen Sphäre oder hinsichtlich eines anderen sozialen Guts.“ (ebd., S. 49)

Für das Erziehungssystem sieht Walzer zwei relevante Verteilungsprinzipien: im Bereich der Elementarerziehung das der Bedürfnisse, in den weitergehenden Bildungsprozessen das der Leistung. Argumentationen, die auf das Bedürfnisprinzip rekurren, werden häufig auch mit Ansätzen in Verbindung gebracht, die ein Mindestniveau an Qualifikation oder Bildung fordern (Giesinger, 2007; Anderson, 2000).

3 Zur Logik des Differenzarguments (Pluralismus) siehe auch Demmerling (1995) und Klenner (1995).

3.2.2 Grundsätze sozialer Gerechtigkeit

David Miller (2008) legt in seinem Buch *Grundsätze sozialer Gerechtigkeit* eine an der Verteilungsgerechtigkeit orientierte Theorie vor. Mit dem Begriff der *sozialen Gerechtigkeit* will er aber übliche Ansätze der Verteilungsgerechtigkeit erweitern:

„Eine Konzeption von sozialer Gerechtigkeit, die um die Frage herum konstruiert ist, wie die Grundstruktur einer Gesellschaft Vorteile und Nachteile an die Gesellschaftsmitglieder verteilt, braucht solange nicht übermäßig [verteilungstheoretisch] restriktiv zu sein, wie sowohl „Vorteile und Nachteile“ als auch „Grundstruktur“ weit genug gefasst werden.“ (Miller, 2008, S. 59)

Insbesondere mit dem Hinweis auf die Grundstruktur, also die Institutionen und Praktiken der Gesellschaft, sei es möglich, auch gesellschaftstheoretische Kategorien wie Macht hinreichend zu berücksichtigen (siehe hierzu auch Honneth im Vorwort zu Miller, 2008). Zugleich grenzt er die Reichweite der Theorie ein, indem er an die jeweiligen Kontexte der Gerechtigkeit, also die je spezifischen lokalen, regionalen und nationalstaatlichen sowie kulturellen Rahmenbedingungen und geschichtlichen Hintergründe erinnert. Die Reichweite seiner Theorie, wie insgesamt die Reichweite von Verteilungsansätzen – so Miller – sei durch drei Prämissen umrissen: Erstens geht eine solche Theorie von einer Gruppe von Menschen aus, die zweitens orientiert an und durch Institutionen agiert, die drittens ihrerseits durch eine Agentur verändert werden können. Dies lässt solche Theorieansätze insbesondere auch für eine Theorie der Schulentwicklung relevant werden.

Miller (2008) geht im Unterschied zu Walzer nun lediglich von drei Vergesellschaftungsformen aus, zu denen je ein spezifisches Verteilungsprinzip gehört. Die erste Vergesellschaftungsform wird als Gemeinschaft bezeichnet und durch den Wert der Solidarität bestimmt, so dass das Bedürfnis zum entscheidenden Verteilungsprinzip wird. Die zweite Vergesellschaftungsform nennt er Zweckverband, der sich am Wert der Leistung orientiert und entsprechend die Verteilung von Gütern nach Verdienst vornimmt. Die dritte Form der Vergesellschaftung ist die Staatsbürgerschaft, die am Recht bzw. an gerechtfertigten Ansprüchen orientiert ist und Güter nach dem Verteilungsprinzip der Gleichheit distribuiert.

In der Erziehungswissenschaft ist Miller bislang leider kaum rezipiert worden, obwohl er im Rahmen der gegenwärtigen erziehungswissenschaftlichen Gerechtigkeitsdiskussion zu grundlegenden Einsichten führen könnte, insbesondere durch seine sehr differenzierte Darstellung des von ihm gegen Rawls verteidigten Leistungsprinzips. So hebt er beispielsweise hervor, dass das Leistungsprinzip nur dann gerecht zur Anwendung kommen kann, wenn es mit einem Äquivalenzprinzip einhergeht (z. B. gleicher Lohn für gleiche Arbeit oder gleiche Noten für gleiche erbrachte Schulleistungen). Im Unterschied zu Walzer ist allerdings etwas unklar, wie Schule in Bezug auf die Verteilungsregeln zu denken ist. Herrscht hier ein

Prinzip vor oder sind alle Prinzipien anzuwenden? Oder ist eine Anwendung der Prinzipien nach Alter bzw. Entwicklungsstufe sinnvoll? Solche Fragen ergeben sich auch bei der letzten hier vorzustellenden Theorie.

3.2.3 Soziale Freiheit durch Anerkennung

Der häufig als Anerkennungstheorie der Gerechtigkeit rezipierte Ansatz von Axel Honneth (2011) steht explizit in der Traditionslinie der hegelschen Rechtsphilosophie und differenziert entsprechend drei Sphären der Moderne (persönliche Beziehungen, marktwirtschaftliches Handeln und demokratische Öffentlichkeit, Honneth, 2011, S. 229).⁴ Wenngleich auf diese Weise ein plurales Modell von Gerechtigkeit konzipiert zu sein scheint, kann Honneth auch anders gelesen werden, da es letztlich wie bei Rawls darum geht, Freiheit zu gewährleisten. Allerdings ist der Freiheitsbegriff bei Honneth ein hegelianischer, während sich Rawls auf Kant stützt. Dies ist insofern nicht nur von theoretischer Bedeutung, als dass im hegelschen Konzept der sozialen Freiheit die Institutionen nicht nur mittelbar und formal als Verfahren zur Sicherstellung von Gerechtigkeit dienen, sondern als Sozialisationsinstanzen die Dialektik der sozialen Freiheit erst ermöglichen, weil in den Sphären die Ideale der Gesellschaft, also die Maßstäbe für Gerechtigkeit, bereits verankert, keineswegs jedoch bereits vollständig oder auch nur teilweise realisiert sind (so ist z.B. Gleichberechtigung als Ideal im Grundgesetz, im Sinne negativer Freiheit, verankert, in den Institutionen jedoch unterschiedlich weit realisiert). „Als „gerecht“ muss mithin gelten, was in den verschiedenen sozialen Sphären dazu angetan ist, einen angemessenen Umgang im Sinne der ihnen [den Sphären] tatsächlich zugeordneten Rolle in der ethischen Aufgabenteilung einer Gesellschaft zu fördern“ (ebd., S. 21). Dies meint, dass Gerechtigkeit in der Gesellschaft dort entsteht, wo sphärenspezifisch Anerkennung durch die die Institutionen orientierenden Grundwerte ermöglicht wird, wobei hier auch von den Anerkennungsmodi der Liebe, des Rechts und der Solidarität gesprochen wird, die jeweils Selbstvertrauen, Selbstwert und Selbstachtung ermöglichen (Honneth, 2014, S. 309; Zurn, 2015). Diese Zustände des Selbst sind aber als Resultate von Anerkennungsprozessen der durch die in Institutionen verankerten idealen Werte zu verstehen. Erläuterungsbedürftig ist womöglich vor allem die Positionierung Honneths zum Markt. Er sieht hier im Unterschied zu Miller nicht nur einen Zweckverband, sondern eine auf Kooperation angelegte Form der Befriedigung von gesellschaftlichen Bedürfnissen, bei der Leistung zwar eine Rolle spielt, aber als

4 Miller hatte ebenfalls diese drei Bereiche identifiziert und als Zonen der Vergesellschaftung markiert, ohne auf Hegel zurückzugreifen. Für die Theoriekonstruktion ist es sicherlich notwendig zu klären, wie es zu genau diesen drei Sphären kommt und nicht wie bei Walzer beispielsweise zu elf Sphären. Dabei wäre dann auch herauszuarbeiten, welche gesellschaftlichen Institutionen als Mischsphären zu bezeichnen wären, in denen also unterschiedliche Verteilungs- oder Anerkennungsprinzipien zum Tragen kommen. Kein Ansatz kann hierauf bislang eine Antwort geben. Insbesondere für Schule ist dies aber eine naheliegende Annahme.

Kooperationsbeitrag in einer solidarisch gedachten Gemeinschaft konzipiert wird. Nur so – und dies meint eben die Orientierung an einem Ideal innerhalb einer Sphäre – kann soziale Freiheit auch im Bereich des Marktes zum Tragen kommen (Zurn, 2015, S. 181).

Das Bildungs- und Erziehungssystem wird von Honneth nicht systematisch in den Blick genommen, wenngleich es sicherlich zu den freiheitsverbürgenden Institutionen der Gesellschaft gezählt werden muss (Hernandez, Herzog & Rebelo, 2017). Diese fehlende Aufmerksamkeit in den Arbeiten Honneths, die einer breiten Rezeption seines Ansatzes in der Erziehungswissenschaft nicht abträglich gewesen ist (exemplarisch Balzer, 2014; Helsper et al., 2006; Stojanov, 2006), verwundert allerdings, weil Honneth selbst die Bedeutung der Praktiken der Reproduktion einer Gesellschaft als Kriterium für die normativ-rekonstruktive Auseinandersetzung mit der Gesellschaft ausweist (Honneth, 2011, S. 23). In einem Interview zu schulischer Bildung und Anerkennung (Hernandez et al., 2017) gibt er dann aber doch einige wenige Hinweise. Dort rekonstruiert er Schule als Kampfplatz, auf dem Lehrkräfte, Schülerinnen und Schüler sowie Eltern „um die Autorität der Festlegung von Leistungskriterien, von Unterrichtszielen und sogar von grundsätzlichen Aufgaben der Schule als solche“ (ebd., S. 119) ringen. Honneth sieht dabei zwei im Schulsystem miteinander konkurrierende Werte institutionalisiert, deren Verhältnis nicht näher definiert wird: Gleichheit und Leistung (ebd., S. 120), was zur Folge hat, dass Schulen als „institutionelles Scharnier“ zwischen familiärer Sozialisation und demokratischem Rechtsstaat sowie zwischen Familien und Wirtschaftssystem fungieren (ebd.). Welches Prinzip im schulischen Kontext Vorrang hat, sei Gegenstand des „politisch-moralischen“ (ebd.) Diskurses über die Aufgaben schulischer Erziehung. Schulische Anerkennungsverhältnisse zeichnen sich demnach durch eine enge Verwobenheit rechtlicher und sozialer Anerkennung aus (Helsper, Sandring & Wiezorek, 2005). Es stellt sich jedoch die Frage, auch kritisch gegen die doch mitunter vereinfachenden Ausführungen Honneths zur Schule gewandt, wie sich soziale Freiheit in Schule realisieren kann. Dies wäre normativ mit Honneth dann der Fall, wenn die Individuen „an Institutionen partizipieren können, deren normative Praktiken ein Verhältnis der wechselseitigen Anerkennung sicherstellen“ (Honneth, 2011, S. 89). Soziale Freiheit verlangt eine Wirklichkeit, die uns bei der Realisierung unserer selbst gesetzten Ziele unterstützt (ebd., S. 90). Für die Schule bleibt die analytische Herausforderung, die Honneth im „Recht der Freiheit“ noch allgemein formuliert:

„Es ist, mit anderen Worten, der ethische Bezug auf die Idee der Freiheit, der es für eine Theorie der Gerechtigkeit erforderlich macht, den rein formalen Rahmen zu verlassen und die Grenze zur gesellschaftlichen Materie zu überschreiten; denn zu erläutern, was es für den einzelnen heißt, über individuelle Freiheit zu verfügen, impliziert notwendigerweise, die existierenden Institutionen zu benennen, in denen er in der normativ geregelten Interaktion mit anderen die Erfahrung der Anerkennung machen kann.“ (ebd., S. 124)

3.3 Analyseperspektiven und -dimensionen – Eine Zwischenbilanz

Was, so stellt sich die Frage, ist aber nun gewonnen, da wir die Unterschiedlichkeit gerechtigkeits-theoretischer Konzepte zumindest in Ansätzen aufgezeigt und auch bereits erste Anwendungen auf das Schulsystem erprobt haben? Die Frage gewinnt insbesondere an Brisanz, wenn mit Wigger (2015, S. 89) davon auszugehen ist, dass insbesondere die erziehungswissenschaftlichen Rezeptionen selektiv auf Gerechtigkeits-theorien zugreifen, um je spezifische Funktionen des Schulsystems zu kritisieren oder zu stützen. Ein solch selektiver Zugriff, so instruktiv er für den Moment auch sein mag, kann nicht dauerhaft überzeugen, da sich immer auch andere Theorien finden lassen, die alternative, womöglich auch gegenteilige Erörterungen und Analysen erlauben. Darum lassen sich künftig mindestens drei Strategien nennen, wie Gerechtigkeits-theorien angewendet werden können:

3.3.1 Reflexive Nutzung eines Ansatzes

Aufgrund des behaupteten Mehrwerts gerechtigkeits-theoretischer Reflexionen in Bezug auf das Schulsystem muss nicht zwingend von einer einseitigen Nutzung abgesehen werden, wenn diese ihre Prämissen, ihren Standpunkt und die begrenzte, mittlere Reichweite ihrer Aussagen deutlich macht. So wäre eine normative Rekonstruktion der Kämpfe um Anerkennung im Schulsystem analog der Studie von Honneth (2011) gewiss lohnenswert.

Ebenso sinnvoll wäre die Nutzung spezifischer Gerechtigkeits-theorien zur kritischen Analyse von Programmen zur Steigerung der Chancengerechtigkeit im Bildungswesen sowie der Beobachtung und Reduzierung von Ungleichheiten. Dabei könnten vor allem Begriffe geprüft und Unklarheiten oder Widersprüche aufgezeigt werden. Dies würde auch für eine Analyse von normativen Urteilen im Kontext der Diskussion von Forschungsbefunden oder der davorliegenden Konstruktion von Kategorien der Ungleichheitserfassung gelten.

3.3.2 Theorie-Triangulation

Im Sinne der Grundidee der Triangulation, die darin besteht, eine Mehrperspektivität zu entfalten (Flick, 2011), können die diversen Ansätze genutzt werden, um das je Unterschiedliche in Bezug auf das zu analysierende Problem oder Phänomen, hier Gerechtigkeit im Schulsystem, offenzulegen. Eine solche Position würde jeder Theorie eine begrenzte Reichweite zusprechen und ihren Nutzen hinsichtlich einer möglichst umfassenden Erschließung des Gegenstandes durch eine Triangulation zu erweitern suchen, wodurch die Stärken der jeweiligen Theorien besonders zum Tragen kommen können und etwaige blinde Flecken reduziert würden. Streitigkeiten in Bezug auf Argumentationsfiguren innerhalb der Theorien,

Passung und Widerspruchsfreiheit könnten ausgeblendet, zumindest aber zurückgestellt werden. Für die Frage aber, welche Ungleichheit den herrschenden Gerechtigkeitsvorstellungen entsprechend als akzeptabel gelten darf, ist eine theoriangreifende Erfassung zwingend angezeigt, da diese Frage ein wesentliches Problem demokratischer Gesellschaften anspricht und nur mittels Transparenz, Deliberation und Rechtfertigung legitimationsfähig abgearbeitet werden kann.

3.3.3 Theorie-Integration zwischen Neuentwurf und Eklektizismus

Man muss den Eklektizismus nicht scheuen (Zierer, 2009), und so darf man es durchaus wagen, gerechtigkeits-theoretische Überlegungen zu integrieren. Wigger (2015) sieht diese Möglichkeit als gangbaren Weg und weist auf den Chancenspiegel hin, der unter dem Begriff der Teilhabe die Ansätze von Rawls, Sen und Honneth zu kombinieren versucht (Berkemeyer, Bos, Hermstein, Abendroth & Semper, 2017). Horster (2015) schlägt ähnliches vor, wenn er als gemeinsamen Kern der Konzepte von Gleichheit, Menschenwürde und Anerkennung (Egalitaristen, Anderson und Hegel bzw. Honneth) formuliert: „Eine unbegründete Ungleichbehandlung verletzt die Menschenwürde und zugleich das Gebot der wechselseitigen Anerkennung“ (ebd., S. 46). Im Übrigen könnte auch die von Allmendinger (1999) entwickelte Position zum Problem der Bildungsarmut, die in ihrer Konsequenz Bildungspolitik als Sozialpolitik zu begreifen sucht, als solch ein integrierender Ansatz gewertet werden (Hermstein et al., 2019). Es ist selbstredend, dass hierbei nicht wenige Probleme eingekauft werden. Es könnte sich aber dennoch lohnen, um der Gefahr einer einseitigen Analyse aus dem Weg zu gehen. Die unterschiedlichen Theorien könnten auch zu einem eigenständigen Entwurf sozialer Gerechtigkeit mit einem spezifischen Fokus auf Schule und Gesellschaft führen. Vorarbeiten hierzu liegen bei Miller und Walzer wie auch in den Überlegungen Honneths vor.

4. Kritische Schulsystementwicklungsforschung – Skizze eines Forschungsprogramms

Vor dem Hintergrund der skizzierten reflektorischen und deliberativen Potenziale normativer Ansätze, wird nun eine Rückbindung an die Fragen der Schulentwicklungsforschung, wie sie in Kapitel zwei aufgeworfen worden sind, in Form eines Programms einer kritischen Schulsystementwicklungsforschung unternommen (Berkemeyer, 2016; Berkemeyer & Hermstein, 2018; Berkemeyer & Mende, 2018), wobei vor allem deutlich werden soll, wie die gewonnenen normativen Elemente eingesetzt werden können. Die meisten Aspekte des Programms bieten dabei Potenziale einer reflexiven Nutzung.

4.1 Analyse der Ungleichheitsbeobachtung

Schulentwicklungsrelevante Formen der Ungleichheitsbeobachtungen finden sich vor allem in Formaten des nationalen Bildungsmonitorings (Hermstein & Manitius, 2015; Hermstein, Berkemeyer & Abendroth, 2018). In diesen bildungspolitisch wirksamen Berichten wird definiert, was als Ungleichheit gewertet wird und wie diese Ungleichheiten zu bewerten sind. Da diese Berichte aber selbst als Produkt politischer Aktivität zu werten sind (Rürup, Fuchs & Weishaupt, 2010), sie also ein machtvoll Instrument der Repräsentation des Schulsystems darstellen (angelehnt an Barlösius, 2005), sind die hierin ausgedrückten Relevanzsetzungen hochgradig selektiv und das Ergebnis von Konstruktions- und Aushandlungsleistungen (Tippelt, 2009; Brüsemeister & Niedlich, 2012). Cortina und Pant (2018) folgend, ist mit Blick auf die empirische Bildungsforschung insgesamt davon auszugehen, dass Kategorien der Beobachtung zum einen unvollständig und zum anderen nicht hinreichend begründet sind, weshalb auch die Bewertung der beobachteten Ungleichheiten erschwert wird. Kritische Schulsystementwicklungsforschung (KSEF) kann hierauf aufmerksam machen, Beobachtungskategorien analysieren und auf ihre normativen Grundlagen hin befragen. Wir werden diese Dimension einer KSEF in Kapitel fünf noch einmal konkretisieren.

4.2 Gerechtigkeitstheoretische Weiterentwicklung von Ungleichheitsbeobachtungen

Im Unterschied zu den wenig systematisch wirkenden Vorschlägen einer erweiterten Ungleichheitsbeobachtung bei Cortina und Pant (2018), kann die KSEF unter Rückgriff auf Gerechtigkeitstheorien systematischer begründete Kategorien der Beobachtung vorschlagen und zugleich Hinweise zur Bewertung der Beobachtungen liefern. Mit Blick auf Rawls' Differenzprinzip könnten beispielsweise Befunde zur Leistung von Schülerinnen und Schülern stärker in Bezug auf Lernzuwächse bei den im System am schlechtesten gestellten Schülerinnen und Schülern im Vergleich zu den anderen Schülerinnen und Schülern analysiert werden.

Unter Rückgriff auf gerechtigkeitstheoretische Überlegungen im Anschluss an Honneth (2011) oder Miller (2008) kann zudem gesehen werden, dass Fragen der Bedürfnisgerechtigkeit so gut wie gar nicht beobachtet werden. Was unternimmt das jeweilige Schulsystem, um die Bedürfnisse der Kinder und Jugendlichen zu fördern? Hier könnte in einer ersten Annäherung beispielsweise geschaut werden, welche Mittel für zusätzliche Fördermaßnahmen bereitgestellt werden oder auch wie viele Beratungsstunden im System zur Verfügung stehen.

4.3 Bewertung von Schulentwicklungsprogrammen und Zielbestimmungen

Eine kritische Schulsystementwicklungsforschung greift den von Forst (2015) entwickelten Zusammenhang von Legitimation, Demokratie und Gerechtigkeit auf, um Programme der Schulentwicklung zu reflektieren und in Bezug auf ihre reformerische Leistung hinsichtlich sozialer Gerechtigkeit zu analysieren. So hat auch Miller seine Theorie als Korrektiv politischen und sozialen Handelns in pragmatischer Absicht verstanden (Miller, 2008, S. 26 ff.). Programme der Schulentwicklung (z. B. Schulinspektion, Vergleichsarbeiten, aber auch Projekte zur Erprobung von Schulautonomie oder die Umstellung auf inklusive Beschulung, u. a. m.) müssen deutlich machen können, wodurch sie legitimiert sind, was sie zur Demokratisierung (in einem sehr weiten Verständnis, was die Weiterentwicklung der Institutionen in Bezug auf ihre demokratisch formulierten Ziele ihrer Leistungen einschließt) beitragen und ob sie Grundprinzipien der Gerechtigkeit einhalten. Die Bewertung von Schulentwicklungsprogrammen kann auf diese Weise Einsichten aus der klassischen Evaluationsforschung mit Elementen einer kritischen Gesellschaftstheorie und ethischen Beurteilungsmaximen kombinieren. In einer solchen Perspektive wird erkennbar, dass das Schulsystem bislang noch zu wenig auf das „Recht auf Rechtfertigung“ abstellt (dies wird sich spätestens mit der Aufnahme der Kinderrechte in das Grundgesetz ändern müssen). Wie Schulentwicklungsprogramme gegenüber Kindern als Träger von Rechten begründet werden können, wird somit zur leitenden Fragestellung. Im Falle von Inklusion führt dies beispielsweise zu der Frage, ob die Wirkvariable *time on task* durch inklusive Beschulung geschwächt wird, ob andere, durch Inklusion entstehende, Lerngelegenheiten dies kompensieren und wie eine Abwägung von Individualrechten an dieser Stelle aussehen kann. Aus solchen Analysen lassen sich schließlich auch weitergehende Zielbestimmungen von Schulentwicklungsprogrammen ableiten, die die Achtung vor den Kindern und ihren Rechten zu berücksichtigen haben.

4.4 Pathologien des Schulsystems analysieren und Strategien ihrer Reduzierung

Insbesondere unter Rückgriff auf Honneth (1994, 2000, 2007), aber auch durch eine von uns vorgeschlagene Theorie-Triangulation, kann eine Institutionenanalyse mit dem Ziel der Aufdeckung von Sozialpathologien des Schulsystems als Aufgabe der KSEF herausgestellt werden. Sozialpathologien können dabei mit Honneth als strukturelle Störungen bzw. strukturelle Fehlentwicklungen verstanden werden, die die Selbstverwirklichung des Menschen behindern oder gar verhindern (Honneth, 1994, S. 51, 2010, S. 103 ff. sowie 202 ff.). Formen institutioneller Diskriminierung können in diesem Zusammenhang genannt werden und sind in Teilen auch bereits gut untersucht. So etwa bei Gomolla und Radtke (2009)

als auf Ebene der Organisation Einzelschule institutionalisierte Praxen der Benachteiligung von Schülerinnen und Schülern mit Migrationshintergrund oder als auf Schulsystemebene zur Diskriminierung führende *tradierte* Differenzen in den Schulangebotsstrukturen (Terpoorten, 2014) und damit verbundenen Leistungserbringungen der Schulsysteme (Berkemeyer et al., 2017).

Schulentwicklungsprogramme gewinnen an Legitimität, wenn sie in der Lage sind, zur Bearbeitung solcher Pathologien beizutragen. Die bereits analysierten Pathologien des Schulsystems zeigen, dass insbesondere Kinder und Jugendliche mit ökonomischen Risikolagen, also von Armut betroffene Kinder und Jugendliche, davon benachteiligt sind. KSEF beobachtet entsprechend die Programme der Schulentwicklung hinsichtlich ihrer Potenziale, eben diese Pathologien zu reduzieren, bzw. entwickelt eigenständige Strategien, die wiederum unter Berücksichtigung der in Abschnitt 4.3 genannten Aspekte erfolgen muss.

KSEF versteht sich keineswegs als Gegenprogramm zu einer empirischen, vor allem auf Fragen der Wirksamkeit ausgerichteten Bildungsforschung, sondern als ein Programm zur Integration von empirischer Wirkungsforschung, Systemmonitoring und normativen Überlegungen zur Verfasstheit der Institution Schule. Wilfried Bos, dem dieser Aufsatz gewidmet ist, hat mit seinen Arbeiten den Weg bereitet, um dieses Programm zu entwerfen. Die Leistungen, die mit einem solchen Programm möglich werden können, sind insbesondere in der Rückgewinnung einer kritischen Gesellschaftstheorie für die Fragen der Schulentwicklung zu sehen, die ihrerseits nicht auf ethische Reflexionen und Positionierungen in Bezug auf gerechte Institutionen und gerechte Praktiken in diesen verstanden werden kann. Zudem kann er zu einer Überwindung von wenig produktiven Debatten führen, in denen Empiriker und Theoretiker gleichermaßen dogmatisch erscheinen.

5. Kritische Schulsystementwicklungsforschung – Eine erste Anwendung am Beispiel des Nationalen Bildungsberichts

Nachfolgend soll exemplarisch gezeigt werden, wie eine KSEF vorgehen könnte und welchen Nutzen dies birgt. Dafür wird mit dem Nationalen Bildungsbericht eine prominente Ungleichheitsbeobachtung (siehe Abschnitt 5.1) an wenigen Beispielen analysiert, wobei wir selektive und triangulative Verfahren nutzen. Zudem soll an einem Beispiel gezeigt werden, wie eine gerechtigkeits-theoretisch fundierte Weiterentwicklung von Indikatoren (siehe Abschnitt 5.2) aussehen könnte. Dabei sei nochmals daran erinnert, dass es dem Ansatz der KSEF nicht darum geht, Ungleichheitsbeobachtungen als grundsätzlich unangemessen herauszustellen. Im Gegenteil: Es geht darum, die Prämissen und impliziten Normen und Werte besser zu reflektieren, um die Interpretationen der beobachteten Daten auf Kohärenz und Stimmigkeit zu prüfen.

5.1 Analyse der Ungleichheitsbeobachtungen (Nationaler Bildungsbericht)

Zunächst sei darauf hingewiesen, dass allein die Tatsache, dass Ungleichheiten beobachtet und als soziale Ungleichheiten klassifiziert werden, noch keine Aussagen über deren Legitimität erlaubt. Allerdings beinhaltet die Rede von sozialer Ungleichheit bereits eine „moralische Codierung“ (Schmidt, 2000, S. 385), impliziert sie doch, dass es sich, im Gegensatz zum normativen Leitbild der Gleichheit, um ein rechtfertigungs- und korrekturwürdiges soziales Problem handelt (ebd. S. 385; Zinn, 2008, S. 159). Ohne Referenz auf normative Grundüberzeugungen und Gerechtigkeitspostulate kann demnach kein Maßstab zur Bestimmung von Gleichheit und Ungleichheit entwickelt werden (Barlössius, 2004, S. 10). Obwohl die Wahl der untersuchten Ungleichheitsmerkmale vor dem Hintergrund gesamtgesellschaftlicher Wandlungstendenzen erfolgt und dadurch begründet wird, handelt es sich dennoch um einen selektiven Zugriff, der die historische, gesellschaftliche und politische Bedingtheit von Bildungsmonitoringinstrumenten verdeutlicht – ein Aspekt, der einer (gerechtigkeits-)theoretischen und normativen Reflexion bedarf.

Zur Analyse sozialer Ungleichheiten rekurriert der Nationale Bildungsbericht auf die in der empirischen Bildungsforschung häufig genutzten Merkmale der sozialen Herkunft, des Geschlechts, der nationalen und ethnischen Herkunft sowie regionaler Angebotsdisparitäten (Klieme et al., 2007, S. 130), ohne diese Auswahl eigens zu begründen oder gerechtigkeits-theoretisch zu reflektieren, obwohl implizite Bezüge zu Gerechtigkeitsprinzipien hergestellt werden. Soziale Ungleichheiten werden vor dem Hintergrund des Gleichheitsideals bewertet, wodurch grundlegende Fragen der Verteilungsgerechtigkeit in den Blick geraten, wenn es beispielsweise um ungleiche Zugänge zum Gymnasium bzw. insgesamt um schichtspezifische Bildungsbeteiligungsunterschiede geht.

Der normative Rahmen des Nationalen Bildungsberichts geht gleichzeitig über Fragen der Verteilungsgerechtigkeit hinaus. Dies wird besonders bei der Bestimmung des eigenen Bildungsverständnisses deutlich, welches neben der Ziel-dimension der gesellschaftlichen Teilhabe und Chancengleichheit auch die Dimension der Humanressourcen sowie die individuelle Regulationsfähigkeit umfasst.

Insbesondere der zuletzt genannte Aspekt der individuellen Regulationsfähigkeit, d. h. der individuellen Fähigkeit, „die eigene Biografie, das Verhältnis zur Umwelt und das Leben in der Gemeinschaft selbstständig zu planen und zu gestalten“ (Autorengruppe Bildungsberichterstattung, 2018, S. 1), verweist implizit auf den Befähigungsansatz von Sen. Gegenwärtig ist der nationale Bildungsbericht allerdings noch nicht in der Lage, „umfassende Chancen“ (Sen, 2010, S. 259) abzubilden. Sen folgend, würde es bezogen auf das Schulsystem nicht ausreichen, dieses nur hinsichtlich seiner Ergebnisse zu beurteilen, z. B., welche Kompetenzen Schülerinnen und Schüler erreichen, vielmehr müssten auch die Prozesse, die zu diesem Output geführt haben, in den Blick genommen werden (ebd., S. 259f.). Der Fokus auf Prozesse findet auch in die aktuelle Ausgabe des Bildungsberichts

Eingang. Es wurden bspw. zusätzliche Indikatoren aufgenommen, die ausgewählte Prozessmerkmale wie die Lernumwelten in Schule und Unterricht in den Blick nehmen (Autorengruppe Bildungsberichterstattung, 2018, S. 107ff.). Die Indikatoren zu den innerschulischen Rahmenbedingungen wie auch zu den Aspekten des Unterrichtsgeschehens bieten zugleich Anknüpfungspunkte für Anerkennungstheoretische Zugänge, wenn beispielsweise nach einem wertschätzenden Unterrichtsklima gefragt wird.

Mit Blick auf die Prozessebene wird im aktuellen Bildungsbericht zudem ein neuer Indikator erhoben, der die Kompositionsprofile von Schulen und damit die Herausforderungen und Probleme differenzieller Lern- und Entwicklungsmilieus in den Blick nimmt (ebd., S. 90f.). Mit Sens Befähigungsansatz wären Prozesse zunehmender schulischer Segregation als problematisch anzusehen, da die mit der jeweiligen Schulkomposition einhergehenden unterschiedlichen Lernzuwächse Schülerinnen und Schüler in je unterschiedlichem Maße befähigen, an den verschiedenen gesellschaftlichen Teilsystemen teilzuhaben.

Mit Blick auf die Zieldimension der Humanressourcen lässt sich zudem attestieren, dass der Nationale Bildungsbericht ebenfalls auf Annahmen einer konsequentialistischen Ethik (Utilitarismus) rekurriert. Dies wird im jüngsten Bericht von 2018 besonders deutlich, wenn nach den Wirkungen und Erträgen von Bildung gefragt wird. Trotz der positiven Bilanz, dass Bildung *wirkt* und (monetäre) Erträge zeitigt, bleibt bislang zum einen unklar, ob die Erträge hätten auch höher ausfallen können und zum anderen ist nicht geklärt, über welche Verfahren und Bedingungen die Bildungsrenditen zu Stande gekommen sind (kritisch zur Theorie und Empirie der Bildungsrenditenforschung siehe Meißner, 2015).

Auftrag einer KSEF kann es zukünftig sein, ein systematisches Analyseinstrumentarium auf Basis zentraler, für demokratische Gesellschaften unabdingbarer Gerechtigkeitskonzeptionen zu entwickeln und zu erproben, welches zur Bewertung normativer Implikationen empirischer Forschung ebenso genutzt werden kann wie zu einer theoriebasierten Weiterentwicklung von empirischen Beobachtungskategorien.

5.2 Gerechtigkeitstheoretische Weiterentwicklung von Indikatoren zur Ungleichheitsbeobachtung

Döbert (2010) folgend, werden unter Indikatoren allgemein „quantitativ erfassbare Größen (wörtlich: „Anzeiger“) verstanden, die als Stellvertretergrößen für komplexere, in der Regel mehrdimensionale Gefüge einen möglichst einfachen und verständlichen Statusbericht über die Qualität eines Zustandes liefern sollen“ (ebd., S. 15).

Die Qualität des Schulsystems ist als kontingentes Konzept zu begreifen, das einen Einigungsprozess darüber impliziert, an welchen gewollten Zielgrößen sich die Schulsystementwicklung orientieren soll. Die nationale Bildungsberichterstattung gibt als Quelle der Orientierung ihrer diesbezüglichen Indikatorenauswahl national

und international verfolgte Ziele von Bildung (Benchmarks) aus (Autorengruppe Bildungsberichterstattung, 2018). Solche Absichten können auf relativ simplen Imperativen aufbauen, wie etwa die Mehrung von Hochschulreifezertifikaten oder die Reduzierung der Klassenwiederholungsquote, die dann entsprechend durch Kennwerte abgebildet werden können. Zielstellungen hingegen, die auf komplexeren Ideen fußen, wie z. B. die Förderung von Chancengleichheit (wie in der nationalen Bildungsberichterstattung) oder die Gewährleistung von Gerechtigkeit, bedürfen zunächst einer Vergewisserung über den zielbezogenen Gehalt des formulierten Anspruchs, damit anschließend eine beobachtungsleitende Operationalisierung vorgenommen werden kann, die eine interpretative Beurteilung des so festgestellten Realisierungsgrades von Zielstellungen ermöglicht.

Ausgangspunkt der kritischen Analyse sind also stets normative Ansprüche des Wollens. Wie schon in Abschnitt 4.2 angesprochen, können Elemente gerechtigkeits-theoretischer Konzepte für die Fragen, auf welches Verständnis von Chancengerechtigkeit hingearbeitet werden soll und welche Ungleichheiten im Schulsystem vor diesem Hintergrund als akzeptabel oder gar gewünscht gelten können, leitend sein. Um dies zu konkretisieren und eine gerechtigkeits-theoretisch fundierte Weiterentwicklung von Indikatoren anzudeuten, soll hier exemplarisch auf das sog. Differenzprinzip von John Rawls (siehe Abschnitt 3.1) Bezug genommen werden. Dieses legt nahe, dass soziale Ungleichheiten als gerecht zu bewerten sind, insofern diese unter fairen Bedingungen entstanden sind und den am wenigsten begünstigten Gesellschaftsangehörigen den größten Vorteil bringen. Was bedeutet dies nun für eine auf (statistischen) Kennziffern fußende Indikatorenkonstruktion?

Beispielsweise können Kennziffern aus dem Bereich der Kompetenzen von Schülerinnen und Schülern direkt auf dieses Differenzprinzip hin ausgerichtet werden. Dazu gilt es zunächst, Leistungswerte gemäß ihren Differenzen auf der empirisch realisierten Kompetenzskala einer Messung so anzuordnen, dass beispielsweise die 10 % der besten (oder am meisten begünstigten) den 10 % der schwächsten (oder am wenigsten begünstigten) Schülerinnen und Schüler gegenübergestellt werden können. So sind Minimum-Maximum-Differenzen datenbasiert herauszuarbeiten, die noch weiter durch die Hinzunahme von Hintergrundmerkmalen (soziale Herkunft der Schülerinnen und Schüler, Schulformzugehörigkeit, Nutzung von Ganztagsangeboten usw.) ausdifferenziert werden können. Allerdings kann auf Basis der festgestellten Differenzen kein Rückschluss darüber gezogen werden, ob oder in welcher Weise die am meisten benachteiligten Schülerinnen und Schüler vom schulischen Bildungsangebot profitieren. Anzureichern wären die Darstellungen somit durch kriteriale oder auch zeitliche Maßstäbe. Wenn Längsschnittdatensätze vorliegen (wie sie bspw. im Rahmen der National Educational Panel Study (NEPS) bereitgehalten werden), könnten die Eingangsdifferenzen von Beginn bis zum Ende der Sekundarstufe I abgebildet und auch hinsichtlich (kriterialer) fachbezogener Kompetenzstufen dargestellt werden. Steigen die zu Beginn der Sekundarstufe festgestellten Differenzen zwischen den Extremgruppen über die Klassenstufen hinweg an, womöglich signifikant begünst-

tigt durch Separationen in unterschiedliche Schulformen, kann unter Rückgriff auf Rawls die Diagnose gestellt werden, dass der institutionelle Umgang mit kompetenzbezogenen Ungleichheiten den am wenigsten begünstigten Lernenden offenbar nicht am meisten nutzt, weil sich ihr Abstand zu den am meisten Begünstigten sukzessive vergrößert. Dies wäre als Verletzung des normativen Anspruchs der Realisierung von Gerechtigkeit im Rawls'schen Sinne zu werten. Tiefer gegliederte Kennziffern, durch die die Extremgruppen noch auf ihre schulischen Lernbedingungen hin unterschieden werden können (Lernen in leistungsheterogenen Klassen oder aber in stark differenziellen Lern- und Entwicklungsmilieus), würden einen solchen, gerechtigkeits-theoretisch fundierten Indikator zu Schülerkompetenzen weiter um institutionenbezogene Beurteilungsgrundlagen anreichern.

Wie diese kursorischen Darstellungen andeuten, motivieren explizite normative Prinzipien potentiell die Fortentwicklung von empirischen Stellvertretergrößen, um im Sinne des normativen Konzepts zu fundierteren Entscheidungsgrundlagen zu gelangen. Theorie gestaltet somit Empirie. Ein theortriangulatives Vorgehen würde dann diesen Prozess weiter befruchten und eine vergleichende Vorgehensweise implizieren. Das Verständnis von Bedürfnisgerechtigkeit nach Walzer (2006) etwa regt dazu an, Kennziffern der erbrachten Leistungen gegenüber den differenzenspezifischen Maßen höher zu gewichten, um zu anderen Schlussfolgerungen zu kommen. Dem Befähigungskonzept nach Sen (2010) entsprechend, würde man das Erreichen von Mindestkompetenzen als prioritären Faktor betrachten, da damit der Anspruch an die Ausbildung individueller Freiheit in den Vordergrund rückt. Die Interpretation der empirisch realisierten Beobachtungen ermöglicht somit eine kritische Auseinandersetzung um gemeinsame bzw. geteilte Vorstellungen von gerechten Schulsystemen.

6. Auf dem Weg zu einer Ethik des Schulsystems?

Es wird sich künftig zeigen müssen, ob es sinnvoll ist, eine Teilbereichsethik bzw. angewandte Ethik, wie sie mittlerweile für einige Gesellschaftsbereiche vorliegt (Nida-Rümelin, 2002) für das Schulsystem zu formulieren (siehe hierzu Abschnitt 3.3), sei es auch nur als Übergangsphänomen, um pragmatisch über soziale Gerechtigkeit im Schulsystem zu sprechen. Im Unterschied zu den bisherigen Ausführungen müsste eine solche Ethik nicht nur Fragen der sozialen Gerechtigkeit aufgreifen, sondern auch Fragen nach moralischem Schülerinnen- und Schülerverhalten sowie Lehrerinnen- und Lehrerverhalten aufgreifen. Auch die Organisation der Schule müsste eine stärkere Berücksichtigung finden, als dies bislang der Fall ist.

Eine Ethik des Schulsystems müsste entsprechend Fragen der Gestaltung des Systems, der professionellen Handlungsweisen sowie der Verpflichtungen und Rechte der Mitglieder umfassen. Eine Ethik des Schulsystems dürfte sich dabei vermutlich nicht allein auf die Frage der Gerechtigkeit stützen, sondern wei-

tere Prinzipien der ethischen Reflexion, wie Nützlichkeit oder Tugendhaftigkeit, in Rechnung stellen, aber womöglich auch Kategorien wie Gemeinwohl und Freundschaft (Höffe, 2015a, S. 118ff.). Schließlich müsste die oftmals stiefmütterlich behandelte Frage des Rechts im Schulsystem ergänzend Beachtung finden. Dies gilt insbesondere vor dem Hintergrund, dass die Kinderrechte noch in dieser Legislaturperiode Aufnahme in das Grundgesetz finden sollen. Damit wäre endgültig eine Notwendigkeit konstituiert, die die Umstellung einer Rechtfertigung durch die Schülerinnen und Schüler zur Teilnahme am Unterricht durch Leistung auf eine Rechtfertigung der Institution in Bezug auf die zwangsweise Beschulung der Schülerinnen und Schüler einfordert. Die Schule müsste dann in ihrer gegebenen Verfasstheit erst einmal zeigen, was sie für die, mit Honneth gesprochen, Realisierung sozialer Freiheit beiträgt. Eine KSEF könnte hierzu einen wichtigen Beitrag leisten.

Literatur

- Allmendinger, J. (1999). Bildungsarmut: Zur Verschränkung von Bildungs- und Sozialpolitik. *Soziale Welt*, 50(1), 35–50.
- Anderson, E. (2000). Warum eigentlich Gerechtigkeit? In A. Krebs (Hrsg.), *Gleichheit oder Gerechtigkeit: Texte der neuen Egalitarismuskritik* (S. 117–171). Frankfurt a. M.: Suhrkamp.
- Autorengruppe Bildungsberichterstattung (2018). *Bildung in Deutschland 2018: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: wbv.
- Balzer, N. (2014). *Spuren der Anerkennung: Studien zu einer sozial- und erziehungswissenschaftlichen Kategorie*. Wiesbaden: Springer VS.
- Barlösius, E. (2004). *Kämpfe um soziale Ungleichheit: Machttheoretische Perspektiven*. Wiesbaden: Springer VS.
- Barlösius, E. (2005). *Die Macht der Repräsentation: Common Sense über soziale Ungleichheiten*. Wiesbaden: Springer VS.
- Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differentieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit: Vertiefende Analysen im Rahmen von PISA 2000* (S. 95–188). Springer VS.
- Becker, R. (2017). Bildungssoziologie – Was sie ist, was sie will, was sie kann. In R. Becker (Hrsg.), *Lehrbuch der Bildungssoziologie* (S. 9–36). Wiesbaden: Springer VS.
- 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* (S. 3–53). Wiesbaden: Springer VS.
- Becker, R. & Schulze, A. (2013). Kontextuelle Perspektiven ungleicher Bildungschancen – eine Einführung. In R. Becker & A. Schulze (Hrsg.), *Bildungskontexte* (S. 1–30). Wiesbaden: Springer VS.
- Bellmann, J. & Müller, T. (2011). *Wissen, was wirkt. Kritik einer evidenzbasierten Pädagogik*. Wiesbaden: Springer VS.

- Berkemeyer, N. (2016). Chancengerechtigkeit: Gewollt ungewollt? In B. Jungkamp & M. John-Ohnesorg (Hrsg.), *Soziale Herkunft und Bildungserfolg* (S. 25–31). Berlin: Friedrich-Ebert-Stiftung.
- Berkemeyer, N. (2017). Anthropologie „Neuer Steuerung“: Ein Essay über implizite und explizite Menschenbilder und deren Potenziale für die Steuerungsforschung im Schulsystem. In J. Standop, E. D. Röhrig & R. Winkels (Hrsg.), *Menschenbilder in Schule und Unterricht* (S. 181–194). Weinheim: Beltz Juventa.
- Berkemeyer, N. (im Ersch.). Gerechtigkeit als Maßstab für das System Schule. In R. Koerrenz & N. Berkemeyer (Hrsg.), *Das Schulsystem auf dem Prüfstand*. Weinheim: Beltz Juventa.
- Berkemeyer, N., Bos, W., Hermstein, B., Abendroth, S. & Semper, I. (2017). *Chancenspiegel – eine Zwischenbilanz: Zur Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme seit 2002*. Gütersloh: Bertelsmann Stiftung.
- Berkemeyer, N., Bos, W., Manitius, V. & Müthing, K. (Hrsg.) (2008). *Unterrichtsentwicklung in Netzwerken: Konzeptionen, Befunde, Perspektiven*. Münster: Waxmann.
- Berkemeyer, N. & Hermstein, B. (2018). Schulentwicklung(-sforschung) – Quo vadis? In K. Drossel & B. Eickelmann (Hrsg.), *Does “What works” work? Bildungspolitik, Bildungsadministration und Bildungsforschung im Dialog* (S. 13–36). Münster: Waxmann.
- Berkemeyer, N. & Mende, L. (2018). *Die Handlungsfelder des Lehrerberufs: Eine Einführung*. Stuttgart: UTB.
- Berlin, I. (1995). *Freiheit: Vier Versuche*. Frankfurt a. M.: Fischer.
- Birnbacher, D. (2013). *Analytische Einführung in die Ethik*. Berlin: De Gruyter.
- Bonsen, M., Bos, W. & Rolff, H.-G. (2008). Zur Fusion von Schuleffektivitäts- und Schulentwicklungsforschung. In H.-G. Rolff (Hrsg.), *Jahrbuch der Schulentwicklung: Daten, Beispiele und Perspektiven* (S. 11–39). Weinheim: Beltz Juventa.
- Brenner, P. J. (2010). *Bildungsgerechtigkeit*. Stuttgart: Kohlhammer.
- Brüsemeister, T. & Niedlich, S. (2012). Bildungsmonitoring zwischen Berichterstattung und Steuerungsanspruch – Entwicklungslinien und akteurstheoretische Implikationen. In A. Wacker, U. Maier & J. Wissinger (Hrsg.), *Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung: Empirische Befunde und forschungsmethodische Implikationen* (S. 131–153). Wiesbaden: Springer VS.
- Cortina, K. S. & Pant, H. A. (2018). Ignorierte Differenzen, illegitime Disparitäten: Die zunehmende Betriebsblindheit im Disparitätendiskurs der empirischen Bildungsforschung. *Zeitschrift für Pädagogik*, 64(1), 71–79.
- Dahrendorf, R. (1962). *Gesellschaft und Freiheit: Zur soziologischen Analyse der Gegenwart*. München: R. Piper & Co.
- Demmerling, C. (1995). Differenz und Gleichheit. Zur Anatomie eines Arguments. In C. Demmerling & T. Rentsch (Hrsg.), *Die Gegenwart der Gerechtigkeit: Diskurse zwischen Recht, praktischer Philosophie und Politik* (S. 122–134). Berlin: Akademie.
- Döbert, H. (2010). Indikatorenentwicklung im Rahmen der Bildungsberichterstattung in Deutschland. In BMBF – Bundesministerium für Bildung und Forschung (Hrsg.), *Indikatorenentwicklung für den nationalen Bildungsbericht „Bildung in Deutschland“* (S. 9–22). Berlin.
- Dworkin, R. (2011). *Was ist Gleichheit?* Berlin: Suhrkamp.
- Dworkin, R. (2012). *Gerechtigkeit für Igel*. Berlin: Suhrkamp.
- Fickermann, D. & Maritzen, N. (Hrsg.) (2014). *Grundlagen für eine daten- und theoriegestützte Schulentwicklung: Konzeption und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)*. Münster: Waxmann.
- Flick, U. (2011). *Triangulation: Eine Einführung*. Wiesbaden: Springer VS.
- Forst, R. (2015). *Normativität und Macht: Zur Analyse sozialer Rechtfertigungsordnungen*. Berlin: Suhrkamp.

- Gauthier, D. P. (1999). *Morals by agreement*. Oxford: Clarendon Press.
- Giesinger, J. (2007). Was heißt Bildungsgerechtigkeit? *Zeitschrift für Pädagogik*, 53(3), 362–381.
- Giesinger, J. (2011). Bildung als öffentliches Gut und das Problem der Gerechtigkeit. *Zeitschrift für Pädagogik*, 57(3), 421–437.
- Gomolla, M. & Radtke, F.-O. (2009). *Institutionelle Diskriminierung: Die Herstellung ethnischer Differenz in der Schule*. Wiesbaden: Springer VS.
- Groot-Wilken, B., Isaac, K. & Schräpler, J.-P. (Hrsg.) (2016). *Sozialindices für Schulen: Hintergründe, Methoden und Anwendung*. Münster: Waxmann.
- Habermas, J. (1992). *Faktizität und Geltung: Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Frankfurt a. M.: Suhrkamp.
- Heid, H. (2012). Der Beitrag des Leistungsprinzips zur Rechtfertigung sozialer Ungerechtigkeit. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 81(1), 22–34.
- Heidenreich, F. (2010). *Theorien der Gerechtigkeit: Eine Einführung*. Stuttgart: Budrich.
- Helsper, W., Krüger, H.-H., Fritzsche, S., Sandring, S., Wiezorek, C., Böhm-Kasper, O. & Pfaff, N. (2006). *Unpolitische Jugend? Eine Studie zum Verhältnis von Schule, Anerkennung und Politik*. Wiesbaden: Springer VS.
- Helsper, W., Sandring, S. & Wiezorek, C. (2005). Anerkennung in pädagogischen Beziehungen – Ein Problemaufriss. In W. Heitmeyer & P. Imbusch (Hrsg.), *Integrationspotenziale einer modernen Gesellschaft* (S. 179–206). Wiesbaden: Springer VS.
- Hermstein, B., Berkemeyer, N. & Abendroth, S. (2018). Indikatorengeschütztes Bildungsmonitoring zwischen Institutionsgestaltung und Verständigung: Kritisch-konstruktive Überlegungen. In K. Drossel & B. Eickelmann (Hrsg.), *Does "What works" work? Bildungspolitik, Bildungsadministration und Bildungsforschung im Dialog* (S. 175–190). Münster: Waxmann.
- Hermstein, B., Berkemeyer, N., Bos, W. & Semper, I. (2019). Schulreform und Bildungsarmut. In G. Quenzel & K. Hurrelmann (Hrsg.), *Handbuch Bildungsarmut* (S. 771–789). Wiesbaden: Springer VS.
- Hermstein, B. & Manitiuis, V. (2015). Bildungsberichterstattung als diskursive Ordnung: Begründung der Annahme von sozialer Gerechtigkeit als Leitkategorie. In V. Manitiuis, B. Hermstein, N. Berkemeyer & W. Bos (Hrsg.), *Zur Gerechtigkeit von Schule: Theorien, Konzepte, Analysen* (S. 270–292). Münster: Waxmann.
- Hernandez, F. J., Herzog, B. & Rebelo, M. (2017). Bildung und Anerkennungstheorie: Ein Gespräch mit Axel Honneth. *Vierteljahresschrift für wissenschaftliche Pädagogik*, 74(1), 111–123.
- Höffe, O. (Hrsg.) (2013). *Einführung in die utilitaristische Ethik: Klassische und zeitgenössische Texte*. Tübingen: A. Francke.
- Höffe, O. (2015a). *Gerechtigkeit: Eine philosophische Einführung*. München: Beck.
- Höffe, O. (2015b). *Kritik der Freiheit*. München: Beck.
- Holtappels, H. G., Klemm, K. & Rolff, H.-G. (Hrsg.) (2008). *Schulentwicklung durch Gestaltungsautonomie: Ergebnisse der Begleitforschung zum Modellvorhaben „Selbstständige Schule“ in Nordrhein-Westfalen*. Münster: Waxmann.
- Honneth, A. (Hrsg.) (1993). *Kommunitarismus.: Eine Debatte über die moralischen Grundlagen moderner Gesellschaften*. Frankfurt a. M.: Campus.
- Honneth, A. (1994). *Pathologien des Sozialen. Die Aufgaben der Sozialphilosophie*. Frankfurt a. M.: Fischer.
- Honneth, A. (2000). *Das Andere der Gerechtigkeit: Aufsätze zur praktischen Philosophie*. Frankfurt a. M.: Suhrkamp.
- Honneth, A. (2003). *Kampf um Anerkennung: Zur moralischen Grammatik sozialer Konflikte*. Frankfurt a. M.: Suhrkamp.

- Honneth, A. (2007). *Pathologien der Vernunft: Geschichte und Gegenwart der kritischen Theorie*. Frankfurt a. M.: Suhrkamp.
- Honneth, A. (2010). *Das Ich im Wir: Studien zur Anerkennungstheorie*. Berlin: Suhrkamp.
- Honneth, A. (2011). *Das Recht der Freiheit: Grundriß einer demokratischen Sittlichkeit*. Berlin: Suhrkamp.
- Honneth, A. (2014). *Kampf um Anerkennung: Zur moralischen Grammatik sozialer Konflikte; mit einem neuen Nachwort*. Frankfurt a. M.: Suhrkamp.
- Horster, D. (2015). Bildungsgerechtigkeit aus sozialphilosophischer Sicht. In V. Manitius, B. Hermstein, N. Berkemeyer & W. Bos (Hrsg.), *Zur Gerechtigkeit von Schule: Theorien, Konzepte, Analysen* (S. 42–50). Münster: Waxmann.
- Hübner, D. (2014). *Einführung in die philosophische Ethik*. Göttingen: Vandenhoeck & Ruprecht.
- Hügli, A. (1999). *Philosophie und Pädagogik*. Darmstadt: wbg.
- Kersting, W. (2008). *John Rawls zur Einführung*. Hamburg: Junius.
- Klenner, H. (1995). Über die vier Arten von Gerechtigkeitstheorien gegenwärtiger Rechtsphilosophie. In C. Demmerling & T. Rentsch (Hrsg.), *Die Gegenwart der Gerechtigkeit: Diskurse zwischen Recht, praktischer Philosophie und Politik*. Berlin: Akademie.
- Klieme, E. (2016). Schulqualität, Schuleffektivität und Schulentwicklung. Welche Erkenntnis eröffnet empirische Forschung? In U. Steffens & T. Bargel (Hrsg.), *Schulqualität – Bilanz und Perspektiven: Grundlagen der Qualität von Schule 1* (S. 45–64). Münster: Waxmann.
- Klieme, E., Avenarius, H., Baethge, M., Döbert, H., Hetmeier, H.-W., Meister-Scheufelen, G., Rauschenbach T. & Wolter, A. (2007). Grundkonzeption der Bildungsberichterstattung für Deutschland. In H.-H. Krüger, T. Rauschenbach & U. Sander (Hrsg.), *Bildungs- und Sozialberichterstattung* (S. 129–145). Wiesbaden: Springer VS.
- KMK (2002). *PISA 2000 – Zentrale Handlungsfelder: Zusammenfassende Darstellung der laufenden und geplanten Maßnahmen in den Ländern*. (Stand: 07.10.2002) (Beschluss der 299. Kultusministerkonferenz vom 17./18.10.2002).
- Luhmann, N. (2004). *Schriften zur Pädagogik*. Frankfurt a. M.: Suhrkamp.
- Maaz, K., Baumert, J. & Trautwein, U. (2011). Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? In H.-H. Krüger, U. Rabe-Kleberg, R.-T. Kramer & J. Budde (Hrsg.), *Bildungsungleichheit revisited: Bildung und soziale Ungleichheit vom Kindergarten bis zur Hochschule*. Wiesbaden: Springer VS.
- Meißner, S. (2015). Bildungsrenditen: Neoklassische Probleme, beschäftigungstheoretische Lösungen. In H.-P. Müller & T. Reitz (Hrsg.), *Bildung und Klassenbildung: Kritische Perspektiven auf eine Leitinstitution der Gegenwart* (S. 179–205). Weinheim: Beltz Juventa.
- Menke, C. (1995). Despotie, Individualismus, Vereinheitlichung: Tocqueville über Freiheit und Gleichheit. In C. Demmerling & T. Rentsch (Hrsg.), *Die Gegenwart der Gerechtigkeit: Diskurse zwischen Recht, praktischer Philosophie und Politik* (S. 142–154). Berlin: Akademie.
- Miller, D. (2008). *Grundsätze sozialer Gerechtigkeit*. Frankfurt a. M.: Campus.
- Nagel, T. (2016). *Eine Abhandlung über Gleichheit und Parteilichkeit*. Berlin: Suhrkamp.
- Nerowski, C. (2018). Leistung als Kriterium von Bildungsgerechtigkeit. *Zeitschrift für Erziehungswissenschaft*, 21(3), 441–464.
- Nida-Rümelin, J. (2002). *Ethische Essays*. Frankfurt a. M.: Suhrkamp.
- Nida-Rümelin, J. (2006). *Demokratie und Wahrheit*. München: Beck.
- Nida-Rümelin, J. (2015). *Die Optimierungsfalle: Philosophie einer humanen Ökonomie*. München: btb.

- Nida-Rümelin, J. (2016). *Demokratie als Kooperation*. Frankfurt a. M.: Suhrkamp.
- Pfordten, D. v. D. & Nida-Rümelin, J. (Hrsg.) (2015). *Moralischer Realismus? Zur kohärentistischen Metaethik Julian Nida-Rümelins*. Münster: Mentis.
- Rawls, J. (1979). *Eine Theorie der Gerechtigkeit*. Frankfurt a. M.: Suhrkamp.
- Rawls, J. (2006). *Gerechtigkeit als Fairneß: Ein Neuentwurf*. Frankfurt a. M.: Suhrkamp.
- Reichenbach, R. (2018). *Ethik der Bildung und Erziehung: Essays zur pädagogischen Ethik*. Paderborn: Ferdinand Schöningh.
- Richter, I. (1995). Verfassungsrechtliche Grundlagen des Bildungswesens. In M. Baethge & K. Nevermann (Hrsg.), *Organisation, Recht und Ökonomie des Bildungswesens* (S. 226–243). Stuttgart: Klett.
- Rolff, H.-G. & Tillmann, K.-J. (1980). Schulentwicklungsforschung: Theoretischer Rahmen und Forschungsperspektive. In H.-G. Rolff, G. Hansen, K. Klemm & K.-J. Tillmann (Hrsg.), *Jahrbuch der Schulentwicklung Band 8* (S. 237–264). Weinheim: Beltz.
- Rorty, R. (1988). *Solidarität oder Objektivität? 3 philosophische Essays*. Stuttgart: Reclam.
- Rürup, M., Fuchs, H.-W. & Weishaupt, H. (2010). Bildungsberichterstattung – Bildungsmonitoring. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 377–399). Wiesbaden: Springer VS.
- Rüther, M. (2015). *Metaethik zur Einführung*. Hamburg: Junius.
- Schimank, U. (2015). Grundriss einer integrativen Theorie der modernen Gesellschaft. *Zeitschrift für theoretische Soziologie* 4(2), 236–268.
- Schink, P. (Hrsg.) (2017). *Freiheit: Zeitgenössische Texte zu einer philosophischen Kontroverse*. Berlin: Suhrkamp.
- Schmidt, V.H. (2000). Ungleichheit, Exklusion und Gerechtigkeit. *Soziale Welt*, 51(4), 383–400.
- Sen, A. (2010). *Die Idee der Gerechtigkeit*. Bonn: Bundeszentrale für Politische Bildung.
- Solga, H. & Becker, R. (2012). Soziologische Bildungsforschung – eine kritische Bestandsaufnahme. In R. Becker & H. Solga (Hrsg.), *Soziologische Bildungsforschung* (S. 7–43). Wiesbaden: Springer VS.
- Stahl, T. (2013). *Einführung in die Metaethik*. Stuttgart: Reclam.
- Stojanov, K. (2006). *Bildung und Anerkennung: Soziale Voraussetzungen von Selbst-Entwicklung und Welt-Erschließung*. Springer VS.
- Stojanov, K. (2011). *Bildungsgerechtigkeit: Rekonstruktionen eines umkämpften Begriffs*. Wiesbaden: Springer VS.
- Taylor, C. (1992). *Negative Freiheit? Zur Kritik des neuzeitlichen Individualismus*. Frankfurt a. M.: Suhrkamp.
- Taylor, C. (2012). *Multikulturalismus und die Politik der Anerkennung*. Frankfurt a. M.: Suhrkamp.
- Terpoorten, T. (2014). *Räumliche Konfiguration der Bildungschancen*. Bochum: ZEFIR.
- Tippelt, R. (2009). Steuerung durch Indikatoren!? – Methodologische und theoretische Reflexionen zur deutschen und internationalen Bildungsberichterstattung. In R. Tippelt (Hrsg.), *Steuerung durch Indikatoren: Methodologische und theoretische Reflektionen zur deutschen und internationalen Bildungsberichterstattung* (S. 7–15). Opladen: Budrich.
- Tugendhat, E. (2012). *Vorlesungen über Ethik*. Frankfurt a. M.: Suhrkamp.
- Walzer, M. (2006). *Sphären der Gerechtigkeit: Ein Plädoyer für Pluralität und Gleichheit*. Frankfurt a. M.: Campus.
- Wigger, L. (2015). Bildung und Gerechtigkeit – Eine Kritik des Diskurses um Bildungsgerechtigkeit aus bildungstheoretischer Sicht. In V. Manitiuis, B. Hermstein, N. Berkemeyer & W. Bos (Hrsg.), *Zur Gerechtigkeit von Schule: Theorien, Konzepte, Analysen* (S. 72–92). Münster: Waxmann.

- Ziegler, H. (2011). Soziale Arbeit und das gute Leben – Capabilities als sozialpädagogische Kategorie. In C. Sedmak, B. Babic, R. Bauer & C. Posch (Hrsg.), *Der Capability-Approach in sozialwissenschaftlichen Kontexten* (S. 117–137). Wiesbaden: Springer VS.
- Zierer, K. (2009). Eklektik in der Pädagogik. *Zeitschrift für Pädagogik*, 55(6), 928–944.
- Zinn, J. O. (2008). Normativität in der Ungleichheitsforschung. In J. Ahrens, R. Beer, U. H. Bittlingmayer & J. Gerdes (Hrsg.), *Beschreiben und/oder Bewerten I. Normativität in sozialwissenschaftlichen Forschungsfeldern* (S. 159–183). Berlin: LIT.
- Zurn, C. F. (2015). *Axel Honneth: A critical theory of the social*. Cambridge: Polity.

Ingrid Gogolin

Lernende mit Migrationshintergrund im deutschen Schulsystem und ihre Förderung – Forschungstraditionen und aktuelle Entwicklungen

Zusammenfassung

Seit Beginn der 2000er Jahre, und erneut zunehmend seit der jüngsten Neuzuwanderung ab ca. 2015, richtet sich breite Aufmerksamkeit der Erziehungswissenschaft und Bildungsforschung in Deutschland auf Lernende aus Familien, die hier nicht altansässig, sondern zugewandert sind. Insbesondere die international vergleichenden Large-Scale-Assessment-Studien wie Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS) oder Progress in International Reading Literacy Study (PIRLS) haben das Interesse der politischen und öffentlichen Sphäre darauf gelenkt, dass migrationsbedingte sprachliche und kulturelle Heterogenität in der Schülerschaft nicht als Problem einer zu vernachlässigenden Randgruppe aufzufassen ist, sondern als Grundproblem von Erziehung und Bildung. Vertreten und begründet wird diese Perspektive in der deutschsprachigen Forschung über Migration und ihre Folgen für Bildung und Erziehung indes seit den 1980er Jahren. In meinem Beitrag werfe ich zum Zwecke der Einordnung von Perspektiven zunächst einen Blick zurück auf die damit begründete Tradition und ihre allmähliche Entwicklung seit dem Zweiten Weltkrieg. Sodann stelle ich einige Ergebnisse dar, über die beim gegenwärtigen Stand der Forschung Konsens besteht. Gleichwohl sind durchaus nicht alle Phänomene, die Folgen der Migration für Bildung und Erziehung anzeigen, inzwischen verstanden und erklärt. Mit der Schilderung von zwei Studien, die sich solchen ungeklärten Fragen zuwenden, schließt mein Beitrag ab.

Schlagworte

Folgen der Arbeitsmigration für Bildung und Erziehung (FABER)

Prof. Dr. Dr. h.c. mult. Ingrid Gogolin, Universität Hamburg, Fakultät für Erziehungswissenschaft – Diversity in Education Research, Von-Melle-Park 8, 20146 Hamburg, Deutschland
E-Mail: gogolin@uni-hamburg.de

Learners with a migration background in the German school system and their support – Tradition and current trends

Abstract

Since the beginning of the 2000s, and again increasingly after the recent immigration since 2015, a broader focus of educational science and research in Germany has been on learners from families who have immigrated to the country. The international comparative large scale assessment studies such as Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS) or Progress in International Reading Literacy Study (PIRLS) in particular have drawn the attention of the political and public sphere to the fact that migration-related linguistic and cultural heterogeneity in the student body cannot be perceived as a problem of a marginal group to be neglected, but as a fundamental problem of education and training. This perspective is represented and justified in German-language research on migration and its consequences for education since about the 1980s. In my contribution, I first take a look back at this tradition and its pre-history since World War II – just for the purpose of clarifying perspectives. In the second part, I present key findings on which there is consensus at the current state of research. Nevertheless, not all phenomena that indicate the consequences of migration for education are understood and explained to this day. My contribution concludes with the description of two studies addressing examples of such unanswered questions.

Keywords

Consequences of labour migration for education and formation

Vorbemerkung

Die Forschung über Lernende mit Migrationshintergrund im deutschen Bildungssystem ist thematisch reichhaltiger, als das in meinem folgenden Beitrag zum Ausdruck kommen kann. Je nach fachlicher Perspektive und Interesse, könnten unterschiedliche Facetten dieser Forschung beleuchtet werden. Mit Bezug auf das Handeln in Erziehungs- und Bildungsinstitutionen könnten etwa Fragen der sprachlichen Entwicklung und Bildung (Gogolin, 2017), des interkulturellen Lernens (Baros & Kempf, 2014) oder des fachlichen Lernens (Wessel & Prediger, 2018) im Vordergrund stehen. In meinem folgenden Beitrag jedoch konzentriere ich mich auf einen Überblick über Forschung, die sich mit Ursachen für soziale Ungleichheit befasst, die im Bildungssystem zu beobachten ist. Grund für diese Konzentration ist es, dass ich das Interesse an der Klärung entsprechender Fragen mit Wilfried Bos teile, den zu ehren schließlich der Anlass für diese Publikation überhaupt ist.

1. Hinführungen – Folgen der Arbeitsmigration für Bildung und Erziehung

Die Geschichte der Beschäftigung von Pädagogik und Bildungspolitik mit Kindern und Jugendlichen, die als „Fremde“ „integriert“ werden müssen, beginnt mit der Etablierung von Bildungssystemen nach heutigem Verständnis überhaupt (Krüger-Potratz, 2000). Aus pragmatischen Gründen jedoch konzentriere ich den folgenden groben Rückblick auf die Phase seit dem Wiederaufbau von Bildungssystem, Bildungspolitik und universitärer Pädagogik in Deutschland seit dem Ende des Zweiten Weltkriegs.¹ Nachgezeichnet wird die allmähliche Hinwendung zur Erkenntnis, dass sprachliche und kulturelle Diversität, die durch internationale Migrationen verstärkt wird, nicht als vorübergehendes Sonderphänomen für Bildung und Erziehung gelten kann, sondern eines der grundlegenden Probleme von Bildung in differenzierten Gesellschaften ist.

Die Kriegsfolgenwanderung seit 1945 – *Displaced Persons*, Evakuierte, Flüchtlinge und Vertriebene – hätte jederzeit Anlass gegeben, sich auch mit der Integration von „fremden“ Kindern und Jugendlichen in das wiederentstehende deutsche Bildungssystem zu befassen (Oltmer, 2017).² Aber die Integrationsprobleme der Kinder „fremder Volksgruppen“, „heimatloser Ausländer“ oder „Staatenloser“ waren nicht Gegenstand des Interesses der universitären Pädagogik bzw. der Erziehungswissenschaft, die sich nach 1949 etablierte. Kurzfristige Aufmerksamkeit erregten in jener Zeit beispielsweise „brown babies“: ca. 4600 Kinder von deutschen Müttern und (zumeist) US-amerikanischen („schwarzen“) Vätern, vor allem im amerikanischen Sektor. Ihre Einschulung Anfang der 1950er Jahre entfachte Debatten über „rassische Differenz“. Diskutiert wurde beispielsweise die Frage, ob man die „Mischlingskinder“ nicht zu ihrem eigenen Besten in die Heimatländer ihrer Väter schicken müsse, weil sie möglicherweise „seelisch“ und physisch das Klima in Deutschland nicht verträgen (Lemke Muniz de Faria, 2002). Die in den Nachkriegsjahren mit Fragen der Internationalität und Internationalisierung besonders befasste Vergleichende Erziehungswissenschaft sah keinen Anlass für eine Auseinandersetzung mit den Folgen der „Internationalisierung im Innern“, denn es wurden keine wesentlichen kulturellen, sprachlichen oder pädagogischen Unterschiede im Lande gesehen. Solche Unterschiede, so hieß es, seien in anderen Regionen virulent und daher relevant für erziehungswissenschaftliche Forschung, etwa in der Schweiz mit ihren unterschiedlichsprachigen Landesteilen – aber nicht in Deutschland (Schneider, 1961).

Abzusehen war indes schon zur Zeit solcher Publikationen, dass infolge der Gastarbeiter-Anwerbung auch mit Migrantenkindern in den Schulen zu rechnen war. Die dadurch hervorgerufene Diversitätlinie weckte aber nicht das Interesse der Erziehungswissenschaft, die sich in diesen Jahren intensiv mit anderen po-

1 Die folgenden Ausführungen stützen sich auf die sehr viel ausführlichere Darstellung in Gogolin und Krüger-Potratz (2019), insbes. Kapitel 1 und 3.

2 Dies gilt für die Zeit der deutschen Teilung vor allem für das im Westen der Republik etablierte System (Krüger-Potratz, Hansen & Jasper, 1991).

tentiellen Ursachen für Benachteiligung von Lernenden befasste: Geschlecht, regionale Herkunft, Religionszugehörigkeit (Cortina, Baumert, Leschinsky & Meyer, 2005). Im bekannten Bild des „katholischen Arbeitermädchens vom Lande“ versammelten sich die Merkmale, die als verantwortlich für eine benachteiligende Lebenssituation von Heranwachsenden identifiziert wurden (vgl. dazu auch Helbig & Schneider, 2014). Seit den 1970er Jahren wurden insbesondere in Arbeiten über Sprache und Bildung auch Themen wie die Mittelschicht-Orientierung der Schule und mit dieser sowie mit der Herkunft aus dialektsprechenden Familien verbundene „Sprachbarrieren“ diskutiert, die die Chancen auf Bildungserfolg beeinträchtigen. In sozialwissenschaftlicher Hinsicht wurde an Basil Bernsteins Konzept der Benachteiligung von Kindern aus dem Arbeitermilieu durch schulspezifischen Sprachgebrauch angeknüpft (Bernstein, 1977). Sprachwissenschaftliche Aufmerksamkeit richtete sich besonders auf Zusammenhänge zwischen Dialekt und sozialer Ungleichheit, die durch die sprachlichen Normalerwartungen der Schule noch verstärkt werden (Ammon, 1972). Anklänge an die Möglichkeit, dass Kinder von Migranten vom Problem der Divergenz zwischen schulischen Sprachanforderungen und sprachlicher Herkunft betroffen sein könnten, fanden sich hier zunächst nicht (Auernheimer, 2003).

Die mit Erziehung und Bildung befasste Wissenschaft jener Jahre war also recht zurückhaltend dabei, sich mit Zuwanderung und ihren Folgen zu befassen. Anders war dies in der Bildungspraxis. Aus dieser ist eine Fülle von Publikationen zu verzeichnen, die alltägliche Erfahrungen aus Schulen, aber auch aus sozialpädagogischen Handlungsfeldern vorstellen. Besprochen wurden in der Regel die Schwierigkeiten, die Zugewanderte für die pädagogische Praxis im weiteren Sinne bereiteten. Als wesentliche Ursache für solche Schwierigkeiten wurden „kulturelle Fremdheit“ sowie „sprachliche Defizite“ herausgestellt.³ Maßnahmen, die als erfolgversprechend beschrieben wurden, richteten sich folgerichtig auf die Kompensation solcher Defizite (vgl. z. B. Müller, 1974).

Unter der Bezeichnung „Ausländerpädagogik“ entwickelten sich in den 1970er Jahren zunächst vor allem schulpolitische, schulpädagogische Arbeiten (Hohmann, 1976). Der Grundgedanke der „Fremdheit“ als Ursache für Benachteiligung der Migranten im Bildungssystem des Aufnahmelandes spielte auch hier eine Rolle. Ob der staatsbürgerliche Status eines Heranwachsenden eine bildungsrelevante Kategorie sei, wurde aber schon rasch kritisch diskutiert (Ruhloff, 1983). Das Ringen um Begriffe ist ein starkes Moment der seinerzeitigen erziehungswissenschaftlichen Befassung mit Folgen der Migration (Roth, 2010). Hierin kommt das Interesse an deduktiver Theoriebildung zum Ausdruck, womit darüber hinaus-

3 Dies sind Redeweisen im Übrigen, die sich bis heute halten. Nach wie vor wird vielfach von „sprachlichen Defiziten“ gesprochen und geschrieben, obwohl nur eine Einzelsprache – z. B. die deutsche – der Gegenstand der Betrachtung ist. Mit der Unschärfe der Bezeichnung ist gewiss nicht immer, aber doch auch nicht selten signalisiert, dass die sprachliche Konstitution zwei- oder mehrsprachiger Personen nur dem Ausschnitt beachtenswert sei, den die Majoritätssprache ausmacht.

gelangt werden sollte, Oberflächenphänomene der Migration als grundlegende Gesetzmäßigkeiten misszuverstehen.

Die Bezugnahme auf internationale Entwicklungen ist ein weiteres Merkmal der frühen auf Migration bezogenen Arbeiten. Dadurch angeregt, dass Arbeitskräfteanwerbung bis Mitte der 1970er Jahre eine politische Praxis in allen prosperierenden west- und nordeuropäischen Staaten war, entwickelte sich ein internationaler pädagogischer Diskurs- und Forschungsraum zu Fragen der Integration der Kinder und Jugendlichen, die quasi eine Begleiterscheinung der Anwerbepolitik waren. Diese Entwicklung wurde unterstützt durch internationale Forschungsförderung, insbesondere durch die Europäische Gemeinschaft (später Europäische Union) und kleinere Geldgeber wie den Europarat. Diese Förderung ermöglichte z. B. vergleichende Untersuchungen zur Frage, welche Merkmale des jeweils aufnehmenden Bildungssystems zur Bildungsbenachteiligung von Migranten beitragen bzw. welche unterstützenden Maßnahmen geeignet sein könnten, Nachteile abzubauen (z. B. Boos-Nünning et al., 1983).

Unterstützt durch die Orientierung an einer gesellschaftlichen Problemlage, kam es auch bereits früh zu interdisziplinären Forschungsk Kooperationen mit Nachbarwissenschaften wie der Soziologie oder den Sprachwissenschaften, und damit auch zur vertieften Auseinandersetzung mit zuvor lediglich aus erziehungswissenschaftlicher Sicht reflektierten Begriffen. So wurden etwa Vorstellungen von „Herkunftsgruppen“, „ethnischen Gruppen“ (Barth, 1969) oder „Minderheiten“ (Heckmann, 1982) unter der Frage danach beleuchtet, welche Bedeutung ihnen im Kontext des Problems der Bildungsbeteiligung von Kindern und Jugendlichen aus Migrantenfamilien zukommt (Reich & Wittek, 1984). Mit diesen Reflexionen rückten Fragen nach Zusammenhängen zwischen den Machtverhältnissen in einer Gesellschaft, der Herkunft aus einer Migrantenfamilie und den Bildungschancen der Kinder stärker ins Blickfeld von Untersuchungen, wobei auf die Verquickung von Migration und sozialer Lage bzw. das Verfügen über Ressourcen im Einwanderungsland aufmerksam gemacht wurde (Nauck, 1985, 1988).

Die Forschung richtete sich zudem nicht nur auf Themen, die das Handeln in Erziehungs- oder Bildungseinrichtungen betrafen, sondern auch auf darüber hinausgehende bildungsrelevante Kontexte. Früh realisiert wurden etwa Untersuchungen zu Erziehungsvorstellungen und Sozialisationspraktiken (Nauck & Özel, 1986; Herwartz-Emden, 1986) oder zu Bildungsaspirationen in Migrantenfamilien, verbunden mit Erhebungen zu ihrer Vertrautheit mit den allgemeinen und beruflichen Bildungsmöglichkeiten in Deutschland (Yakut, Reich, Neumann & Boos-Nünning, 1986). Auch das Problem eventueller geschlechtsspezifischer Erziehungsvorstellungen und -praktiken wurde in Studien als Hintergrundfaktor berücksichtigt (Neumann, 1980) oder explizit adressiert (Herwartz-Emden, 1991).

Diese Entwicklung kann zusammenfassend als Hinwendung zu einer theoretisch argumentierenden und (entsprechend der gegebenen Möglichkeiten) datengestützten erziehungswissenschaftlichen Bildungsforschung im Kontext von Migration gekennzeichnet werden. In empirischen Analysen der Bildungsbeteiligung auf der Basis öffentlich zugänglicher amtlicher Daten (Klemm, 1979) wurde der Umstand,

dass lediglich die Kategorie der Staatsbürgerschaft als Hinweis auf Migration zur Verfügung stand, kritisiert, und es wurden mögliche Differenzierungen – etwa mit Bezug zur Bedeutung der sprachlichen Herkunft für Bildungserfolg – diskutiert (Boos-Nünning & Gogolin, 1988). Es etablierte sich die Bezeichnung „Interkulturelle Pädagogik“ für die Befassung mit Fragen der Erziehung und Bildung im Kontext von Migration (Krüger-Potratz, 2005) – eine Nomenklatur allerdings, die ebenfalls von vornherein auch umstritten war (Hamburger et al., 1981). Die Forschung entwickelte sich zunehmend interdisziplinär und auf internationale Grundlagen bezogen, insbesondere auf Arbeiten, die in den „klassischen“ Migrationskontexten wie USA und Kanada entstanden.

Auf der Grundlage des erreichten Standes der Theorie und Forschung gelang 1990 die Einwerbung eines DFG-Forschungsschwerpunktprogramms mit dem Titel „Folgen der Arbeitsmigration für Bildung und Erziehung – FABER“ (Reich, Gogolin, Krüger-Potratz & Neumann, 1990). Dieses war seinerzeit das erste erziehungswissenschaftliche Schwerpunktprogramm seit mehr als einem Jahrzehnt (Reich & Merckens, 1993). In der in diesem Programm geförderten Forschung konnten sowohl weitere theoretische Klärungen zum Problem der Erziehung und Bildung im Migrationskontext erfolgen als auch eine Konsolidierung methodischer Ansprüche an die Forschung (Diefenbach & Nauck, 2000; Herwartz-Emden & Westphal, 2000). Die Projekte im Schwerpunktprogramm stützten sich auf ein geteiltes Grundverständnis, nach dem Migration und die in ihrer Folge zunehmende innergesellschaftliche soziale, sprachliche, ethnische und kulturelle Heterogenität nicht als ein vorübergehendes Sonderproblem aufzufassen sei, sondern als ein Grundmerkmal von differenzierten Gesellschaften. Die Konsequenzen dieser Konstellation betreffen nicht allein die Gewanderten, sondern zugleich auch die Nichtgewanderten in einer Gesellschaft – sei dies die Aufnahme-, sei es die Herkunftsgesellschaft. Die Forschung im Schwerpunktprogramm war daher mit dem Anspruch verbunden, den Blick nicht allein auf die Zielgruppe „Migranten“ zu richten, sondern auf die Migrationsgesellschaft insgesamt. Migration und ihre Folgen für Bildung und Erziehung sind nach diesem Verständnis ein „Fall“ sozialer Ungleichheit mit spezifischen Merkmalen (Gogolin, 2016; Gogolin & Nauck, 2000). Das Programm FABER, so heißt es in einer Rückschau auf die Entwicklung, hat „das Profil der interkulturellen Bildungsforschung nachhaltig geprägt“ (Mehringer & Herwartz-Emden, 2018, S. 195).

2. Weiterführungen – Internationale Schulleistungsvergleiche

Die Forschung über Migrationsfolgen für Gesellschaft und Individuum, insbesondere mit dem Bezug auf Bildung und Erziehung, hat sich also bis zum Ende des 20. Jahrhunderts in Deutschland konsolidiert, historisch und theoretisch vergewissert, inhaltlich ausdifferenziert, disziplinübergreifend und international vernetzt. Sie

ist gleichwohl konzentriert geblieben auf einen eher engen Kreis von Spezialisten, die sich bisweilen selbst als „mit Außenseitern befasste Außenseiter“ der wissenschaftlichen Gemeinschaft gesehen haben – denen diese Rolle aber durchaus auch zugeschrieben wurde. Viele der wissenschaftlichen Publikationen, etwa aus den an FABER beteiligten Projekten, haben in Bildungsverwaltung und -praxis ein Echo gefunden, weniger aber in jenen Bereichen der Erziehungswissenschaft und Bildungsforschung, die sich selbst als „allgemein“ verstehen (Krüger-Potratz, 1999).

Eine Änderung der Aufmerksamkeit gegenüber dem Problem der Bildungsbenachteiligung von Migranten stellte sich mit der Realisierung von internationalen Schulleistungsvergleichsstudien wie *Trends in International Mathematics and Science Study* (TIMSS), *Programme for International Student Assessment* (PISA) und *Progress in International Reading Literacy Study* (PIRLS) ein; hieran hat Wilfried Bos erheblichen Anteil (Baumert, Bos & Lehmann, 2000; Bos, 2003). Durch internationale ökonomische (wie die *Organisation for Economic Co-operation and Development* (OECD)) oder wissenschaftliche „Treiber“ (wie die *International Association for the Evaluation of Educational Achievement* (IEA)) kamen die inzwischen routinisierten internationalen Schulleistungsvergleichsstudien auf die Tagesordnung, eingeschlossen die Prüfung von Faktoren, die jenseits der Gestaltung des Bildungssystems die Leistung von Lernenden beeinflussen können. Zu den Letzteren – definiert als Risikofaktoren – gehören Merkmale der Lebensumstände, darunter auch solche, die mit Migration verbunden sind oder sein könnten.

Für die Identifikation der davon betroffenen Kinder und Jugendlichen wurden Indikatoren gewählt (keineswegs aber neu erfunden), denen eher als dem Staatsbürgerschaftskriterium ein messbarer Einfluss auf Bildungserfolg zuzuschreiben sein kann. Auch diese beruhen auf dem Grundgedanken einer Abschätzung der „Fremdheit“ von Heranwachsenden für ein Bildungssystem (oder umgekehrt: eines Bildungssystems für die Heranwachsenden) und der Möglichkeit ihrer Überwindung. Als Grundlage für die Möglichkeit der Überwindung von „Fremdheit“ wird „Zeit“ angenommen. Sie wird geschätzt als Spanne, die seit einer Migration vergangen ist, und verfeinert durch die Berücksichtigung der Migrationskette: den Hinweis darauf, ob Migration im eigenen Erleben der getesteten Person stattfand oder im Erleben eines oder mehrerer direkter Vorfahren. Als ein Hinweis auf „Fremdheit“ in der Lebenspraxis wurde der Sprachgebrauch in der Familie ausgewählt. Das (überwiegende) Sprechen einer anderen als der Majoritätssprache im häuslichen Umfeld galt (und gilt vielfach) als Anzeichen für unzureichende Adaption an die Lebensverhältnisse und schulischen Voraussetzungen im Einwanderungsland und somit als Risiko für Bildungserfolg.

Ungeachtet des Umstands, dass diese ersten, recht holzschnittartigen Deutungen der Differenzierung bedürfen, sind mit den internationalen Schulleistungsvergleichsstudien wichtige Beiträge zur Forschung über Folgen der Migration für Bildung und Erziehung geleistet worden. Ihre Bedeutung besteht nicht zuletzt darin, dass sie sich auf das „Argument der großen Zahl“ stützen können. Zu den Gewinnen gehört erstens, dass die Aufmerksamkeit der fachlichen

und allgemeinen Öffentlichkeit für die Größenordnung der Migration geweckt wurde. Durch die neue Zählung war eine Gruppe von erheblicher Größe „entstanden“. Junge Menschen „mit Migrationshintergrund“ konnten danach nicht mehr als zu vernachlässigende Randerscheinung im Bildungssystem betrachtet werden, sondern es galt nun als empirisch untermauert, dass erhebliche Teile der an Bildung teilhabenden Bevölkerung aus Familien mit Migrationserfahrung stammen.

Ein zweiter wesentlicher Beitrag ist darin zu sehen, dass mit den Studien das Konzept der „Literalität“ als Übertragung aus dem Englischen in die deutschsprachige empirische Bildungsforschung Einzug genommen hat. Im deutschsprachigen Diskurs wurde es mit dem einheimischen Begriff der Bildung verknüpft und interpretiert als „Grundbildung“. Angesprochen ist hier das Verfügen über grundlegende Kulturfähigkeiten, zunächst mit Augenmerk auf Zugang zu den Welten der Schrift, der Zahl und zu den Gesetzmäßigkeiten der Natur als Voraussetzung für gesellschaftliche Teilhabe (Baumert, 2001; für die Grundschule vgl. auch Bos, 2003).

Mit Literalität in diesem Verständnis ist zugleich die Perspektive der Messbarkeit der entsprechenden Fähigkeiten verbunden, die im Begriff der Kompetenz zum Ausdruck kommt. Für das komplexe Problem der Messung schriftsprachlicher Fähigkeiten wurde bekanntlich die Lösung einer Konzentration auf Lesekompetenz gefunden. Die gemessene Lesekompetenz wurde auch unter Bezeichnungen wie „Sprachkompetenz“ (Baumert, 2001), „deutsche Sprachkompetenz“ (Müller & Stanat, 2006; Stanat, 2006) oder „Sprachbeherrschung“ (Baumert & Schümer, 2002, S. 375) präsentiert, wodurch die Assoziation geweckt wurde, es handele sich um Maße mit einem allgemeinen, umfassenden Aussagegehalt über sprachliche Fähigkeiten (im Deutschen) statt um – zweifellos bildungsrelevante – Teilfähigkeiten, deren Aussagekraft für „Sprache insgesamt“ jedoch unklar ist (Grabowski, 2014).

Ungeachtet dessen gelang es mit der Fokussierung auf Literalität und der Messung der Lesekompetenz, die fachliche und politische Aufmerksamkeit für die grundlegende Bedeutung sprachlicher Fähigkeiten für Lernen und Bildung überhaupt (wieder-)zugewinnen. In einer historischen und bildungstheoretischen Perspektive stand diese Bedeutung kaum je in Zweifel (Tenorth, 2016). Auch im Kontext interkultureller Pädagogik war sie präsent – sowohl mit Blick auf die Frage der Vermittlung des Deutschen an Lernende, die diese Sprache nicht selbstverständlich aus der Familie mitbringen, als auch im Kontext der Aufarbeitung des Verhältnisses von Bildung und Sprache im nationalen Bildungssystem überhaupt (Gogolin, 2008; Krüger-Potratz, 1994). Im Allgemeinen jedoch richtete sich seit den 1980er Jahren nur recht geringe Aufmerksamkeit auf die grundlegende Bedeutung von Sprache für Bildung, und es gehört zu den bedeutenden Meriten der internationalen Schulleistungsvergleichsforschung, dieses Thema zurück auf die Tagesordnung gebracht zu haben – nicht nur in der Erziehungswissenschaft und Bildungsforschung, sondern auch in Bildungspolitik und Praxis. In der Forschung schlägt sich das Interesse am Thema sowohl in Arbeiten zur Gestaltung von sprachlicher Bildung nieder (Becker-Mrotzek & Roth, 2017; Kilian, Duarte,

Gogolin, Schnoor & Lagemann, 2016; McElvany, Ohle, El-Khechem, Hardy & Cinar, 2017) als auch in solchen, die sich der Klärung von möglichen sprachlichen Ursachen für Bildungsungleichheit zuwenden (Kempert et al., 2016).

3. Geklärtes und zu Klärendes

Soweit der Rückblick auf zwei bedeutende Forschungstraditionen zu Fragen der Bildung und Erziehung im Migrationskontext. Festzuhalten ist, dass erhebliche Gewinne im Hinblick auf Datenbestände und Methoden erzielt wurden, die eine differenzierte Lagebeschreibung ermöglichen. Auch gelingt es deutlich besser, der Frage nach Erklärungen für das Phänomen des im Durchschnitt geringeren Bildungserfolgs der Schülerschaft mit Migrationshintergrund im Vergleich zu altansässigen Mitschüler(inne)n nachzugehen, als das in den 1970er bis zum Ende der 1990er Jahre der Fall war. Dies ist nicht zuletzt der inzwischen erreichten Leistungsfähigkeit methodischer Zugriffsweisen zu verdanken. Diese ermöglichen es, ein Spektrum von potentiellen Einflussfaktoren in Analysen einzubeziehen und ihre Verschränkungen oder Wechselwirkungen komplex zu modellieren. Auch die Möglichkeiten der Verknüpfung von qualitativen und quantitativen Perspektiven sind reifer als noch in den 1990er Jahren. Für viele Fragen hilfreich ist ferner, dass inzwischen auf amtliche Daten (Mikrozensus) sowie auf Befragungsdaten (etwa aus dem Nationalen Bildungspanel) zurückgegriffen werden kann, die das Phänomen Migration genauer fassen als im Merkmal der Staatsangehörigkeit – auch das verdankt sich nicht zuletzt den Debatten im Anschluss an die internationalen Schulleistungsvergleichsstudien.

Inhaltliche Übereinstimmung zur Frage, welche Faktoren – ungeachtet der Unterschiede zwischen Migranten – für Bildungsbenachteiligung mitverantwortlich sind, besteht in Bezug auf soziale Herkunftseffekte. In der Mehrzahl der Fälle ist Migration mit einem Verlust an sozialem und kulturellem Kapital verbunden; zugleich verbessern sich die ökonomischen Lebensumstände zwar im Vergleich zur vorherigen Existenz in der Herkunftsgesellschaft, aber in der Aufnahmegesellschaft werden eher Positionen in unteren Einkommensbereichen eingenommen.⁴ Der Handlungsspielraum für Investitionen in Bildung ist daher in Migrantenfamilien durchschnittlich geringer als in altansässigen Familien (Nauck, 2011; Haag, Kocaj, Jansen & Kuhl, 2017). Benachteiligende Effekte machen sich schon früh in der Bildungskarriere bemerkbar und verstärken sich in ihrem Verlauf bis in den tertiären Bereich hinein (Diehl, Hunkler & Kristen, 2016). Im Hinblick auf Wünsche oder Erwartungen, die Migranteltern und ihre Kinder an Bildung herantragen, wären nach dem Stand der Forschung eher begünstigende als benachteiligende Effekte auf Bildungserfolg zu erwarten (Becker, 2010). In Bezug auf das Geschlecht

4 Dies gilt insbesondere für den Kontext von Arbeitsmigration, wie sie etwa in europäischen Staaten vorfindlich ist, die nicht über eine systematische Einwanderungspolitik verfügen, sondern eine „nachholende Integrationspolitik“ betreiben (wie die Bundesrepublik Deutschland; vgl. Bade, 2007).

als Einflussfaktor widersprechen die vorliegenden Ergebnisse auch eher den verbreiteten Erwartungen, als dass sie ihnen entsprechen. Wenn Anzeichen für eine geschlechtsspezifische Benachteiligung bei Migranten gefunden werden, so betreffen sie das männliche Geschlecht (Hadjar & Hupka-Brunner, 2013). Mit unumstrittenen Resultaten wie diesen ist ein weiter Bereich von möglichen Ursachen für Bildungsbenachteiligung aus der Perspektive sozialer Ungleichheitsforschung aufgeklärt. Allerdings wäre es ungut, sich mit dem erreichten Stand der Klärung zufriedenzugeben. Zwei Beispiele dafür, dass Beunruhigung über das Erreichte noch angebracht ist, möchte ich abschließend kurz vorstellen. Beide Beispiele stammen aus dem Projekt *HeBe – Herkunft und Bildungserfolg*, das 2013 bis 2017 an den Universitäten Chemnitz und Hamburg durchgeführt wurde.⁵

3.1 Beispiel 1: Erklärungswert migrationsbedingter Faktoren

Das erste Beispiel betrifft die Frage, ob die sozio-ökonomische Herkunft und ihre Konfundiertheit mit migrationsbedingten Lebensbedingungen das Phänomen der Bildungsbenachteiligung von Migranten tatsächlich in dem Maße klären, wie dies verbreitet angenommen wird. Im Raum steht hier eine wissenschaftliche Kontroverse über die Frage, ob Faktoren, die sich der Migration verdanken, überhaupt einen eigenständigen Erklärungswert besitzen (Kalter, 2008; vgl. auch Nauck, 2009). Tatsächlich bringt die inzwischen zum Standard gewordene empirische Kontrolle von sozio-ökonomischen Faktoren oft migrationsbedingte Faktoren zum Verschwinden; Bildungsbenachteiligung ist demnach eine Folge der fehlenden materiellen Ressourcen der Herkunftsfamilie. Dass es dennoch nicht überflüssig ist, sich der Klärung des Problems migrationsbedingter Benachteiligung weiterhin und genauer anzunehmen, zeigt sich in Untersuchungen, in denen Migrantengruppen vergleichend in den Blick genommen werden (etwa Segeritz, Walter & Stanat, 2010).

Auf einen solchen Vergleich bezieht sich die HeBe-Untersuchung, über die ich hier abschließend berichte. Anlass für die Studie war die internationale, insbesondere in den USA wiederkehrend gewonnene Erkenntnis, dass Migrantenkinder mit südost-asiatischem Hintergrund trotz der sozio-ökonomisch benachteiligten Lebenssituation und anderer geringer Ressourcen ihrer Familien in hohem Maße bildungserfolgreich sind (Lee & Zhou, 2014). Wenngleich die Anzahl von Kindern und Jugendlichen entsprechender Herkunft in Deutschland zu gering ist, um sie in Large Scale-Untersuchungen zu berücksichtigen, liegen Hinweise auf den aus humankapitaltheoretischer Sicht ungewöhnlichen Bildungserfolg auch für Deutschland vor (Walter, 2011). Auch Daten der amtlichen Bildungsstatistik

5 Das Projekt wurde im Forschungsschwerpunkt „Chancengerechtigkeit und Teilhabe“ des Bundesministeriums für Bildung und Forschung gefördert (Az. 01JC1123A). Es stand unter der Leitung von Prof. Dr. Bernhard Nauck (Technische Universität Chemnitz) und Prof. Dr. Ingrid Gogolin (Universität Hamburg). Auf der Projektwebsite www.hebe.uni-hamburg.de ist das Design der Studie detailliert dargestellt.

deuten in diese Richtung. So besuchten im Schuljahr 2013/14 rund 47 Prozent der deutschen, 20 Prozent der türkischen, aber 64 Prozent der vietnamesischen Jugendlichen ein Gymnasium (Statistisches Bundesamt, 2014).

Studien zeigen, dass vietnamesische Eltern in Deutschland über ein geringeres Sozialkapital verfügen als türkische Eltern; beide Migrantengruppen wiederum besitzen geringeres Sozialkapital als Eltern aus altansässigen Familien (Nauck & Lotter, 2014). Auch die ökonomische Situation vietnamesischer Familien in Deutschland ist für Bildungserfolg nachteilig; sie verfügen über ein geringeres Einkommen als türkische Familien, und beide ethnischen Gruppen verfügen über ein durchschnittlich deutlich geringeres Einkommen als altansässige Familien (Mikrozensus; Statistisches Bundesamt, 2010). Von Vorteil für Jugendliche vietnamesischer Herkunft sollte sein, dass ihre Eltern ein höheres Bildungsniveau besitzen als Eltern türkischer Herkunft. Knapp 31 % der deutschen Mütter mit mindestens einem Kind im Schulalter besitzen eine (Fach-)Hochschulreife; bei vietnamesischen Müttern trifft dies auf knapp 32 % zu, aber bei Müttern türkischer Herkunft nur auf knapp 15 % (Statistisches Bundesamt, 2010).

Obwohl Erklärungsversuche für Bildungsbenachteiligung, die sich vornehmlich auf die unterschiedliche Ressourcenausstattung der Herkunftsfamilien von Migranten beziehen, durch eine Vielzahl von empirischen Studien eindrucksvoll belegt sind und die ihnen unterliegenden Mechanismen über innerfamiliäre Entscheidungsprozesse und die kumulativen Effekte an Übergängen in der Bildungskarriere nachgewiesen werden konnten, wird ihre Verallgemeinerungsfähigkeit durch den Bildungserfolg vietnamesischer Jugendlicher in Frage gestellt. Für die HeBe-Studie (vgl. zum folgenden Nauck & Schnoor, 2015) wurde ein Erklärungsmodell entwickelt, das Unterschiede in der kulturellen Transformation von Ressourcen in familiäre Bildungsinvestitionen einbezieht.⁶ Erfragt wurden familiäre Akkulturationsstrategien (z. B. Verwendung und Förderung der Sprache des Aufnahmelandes in der familialen Interaktion), familiäre Sozialisationspraktiken (z. B. elterliche Supervision und Kontrolle der Kinder bzw. Jugendlichen, praktizierter Erziehungsstil). Als Kriterium des Bildungserfolgs wurden Noten in Deutsch und Mathematik, gewichtet an den Unterschieden zwischen den Schulformen, einbezogen. Familiäre Ressourcen, sprachliche Akkulturation und Sozialisationspraktiken wurden in standardisierten, mehrsprachigen CAPI-Interviews mit Müttern deutscher, türkischer und vietnamesischer Herkunft erfragt, wobei bilinguale Interviewer eingesetzt werden. Zudem wurden produktive Fähigkeiten in emergierendem bildungssprachlichen Deutsch erhoben; dazu später mehr.

Zu den Ergebnissen der Untersuchung gehört, dass deutliche Unterschiede in der Ausstattung der deutschen, türkischen und vietnamesischen Herkunftsfamilien

6 Die Daten zum Projekt HeBe wurden in einem 2 x 3 x 4 Design zwischen Herbst 2012 und Herbst 2013 erhoben: in zwei Aufnahmekontexten (Hamburg und Sachsen), mit drei unterschiedlichen Herkunftsnationalitäten (deutsch, türkisch und vietnamesisch) der Mütter und vier Altersgruppen der Kinder (ca. 6-, ca. 12- und ca. 15-jährig). Insgesamt stehen Daten von 1523 Mutter-Kind-Dyaden zur Verfügung. Die hier nachfolgend berichteten Beobachtungen beziehen sich auf die Gruppe der ca. 15-jährigen Schüler(innen).

mit ökonomischem, sozialem und kulturellem Kapital zu verzeichnen sind. Ermittelt wurden zudem Unterschiede im Bildungserfolg, die den oben berichteten Tendenzen entsprechen. Die Analysen ergaben jedoch, dass diese Unterschiede nur sehr begrenzt auf die Ressourcenausstattung zurückgeführt werden können. „Innerhalb der ethnischen Gruppen tragen die Ressourcen kaum zur Differenzierung des Bildungserfolgs bei, wenn Akkulturation und Sozialisation in den Familien kontrolliert werden“ (Nauck & Schnoor, 2015, S. 649). Es scheint, dass der Einfluss der Ressourcen in den Herkunftsfamilien auf den Bildungserfolg ihrer Kinder deutlich geringer ist, als das in den autochthonen Familien der Fall ist. Dies deutet auf eine geringe Variabilität in den Lebenslagen von Migrantenfamilien im Vergleich zu altansässigen Familien hin. Zugleich zeigen die Analysen, dass die familialen Ressourcen zwischen den ethnischen Gruppen erheblich differenzieren. Die Autoren kommen zu dem Schluss, dass der geringere Bildungserfolg der Jugendlichen türkischer Herkunft weitgehend auf die geringe Verfügbarkeit von familialen Ressourcen zurückzuführen ist. Dies trifft aber auf die vietnamesische Herkunftsguppe nicht zu, denn trotz ihrer ungünstigen Ressourcenlage erzielen ihre Kinder ähnlich hohe Bildungserfolge wie Gleichaltrige aus nichtgewanderten Familien, und es ist davon auszugehen, dass sie bei günstigerer Ressourcenlage noch bildungserfolgreicher wären. Es lohnt sich demnach, der Frage weiter nachzugehen, welche familialen Orientierungen und Praktiken es eigentlich sind, denen Wirkungen auf den Bildungserfolg ihrer Kinder zukommt.

3.2 Beispiel 2: Nachteil durch Herkunftssprache

Zu den wissenschaftlich und öffentlich vieldiskutierten Vermutungen darüber, welche familialen Praktiken den Bildungserfolg der Kinder aus Migrantenfamilien negativ beeinflussen, gehört das Sprechen der Herkunftssprache. Wie bereits angesprochen, wird der Gebrauch der Herkunftssprache in der Familie zu den Risikofaktoren für Bildungserfolg gerechnet. Die darauf deutenden Resultate der Large Scale-Studien erscheinen robust, und es kann kein Zweifel daran bestehenden, dass adäquate Kenntnisse der Schul- und Unterrichtssprache eine wesentliche Voraussetzung für Bildungserfolg überhaupt sind. Die Annahme, dass eine dementsprechende Praxis in der Familie ein Mittel zur Förderung der Fähigkeiten in dieser Sprache ist, ist daher naheliegend. Vor dem Hintergrund, dass zugleich ein erheblicher Bestand an Forschung darauf deutet, dass Zwei- oder Mehrsprachigkeit an sich keinen Anlass gibt, von Benachteiligung auszugehen (Poarch & Bialystok, 2017), wurde die Frage nach der potentiellen Funktion des familialen Sprachgebrauchs für die Aneignung des Deutschen als Bildungssprache (*academic language*) und für Bildungserfolg im HeBe-Projekt eingehender untersucht. Grundannahme dabei ist, dass bildungssprachlichen Fähigkeiten eine Schlüssel-funktion bei der Erbringung schulischer Leistungen zukommt, vergleichbar mit anderen bildungsrelevanten Bereichen des Humankapitals wie mathematischen oder naturwissenschaftlichen Fähigkeiten (Gogolin & Duarte, 2016).

Die Analysen zu diesen Fragen (Schnoor, 2018)⁷ beruhen auf gemessenen produktiven Sprachdaten im Deutschen der Kinder aus der HeBe-Stichprobe. Bei den eingesetzten Instrumenten handelt es sich um Weiterentwicklungen von HAVAS5, FörMig-Tulpenbeet und FörMig-Bumerang (Reich, Roth & Neumann, 2007; Klinger et al., 2019). Mit diesen Instrumenten werden altersentsprechende Sprachprodukte elizitiert. Die Aufgabenstellungen sind schul- und unterrichtstypisch, aber nicht curriculumbasiert. Durch das identische Konstruktionsprinzip der Instrumente für die jeweilige Altersgruppe wird ein standardisierter Rahmen gesetzt, der die Grundlage dafür bildet, die gemessenen Sprachstände zwischen den Gruppen zu vergleichen. Die Auswertung erfolgt anhand von fünf Indikatoren (Aufgabenbewältigung; Anzahl Verben (types); Anzahl Satzverbindungen (types); Anzahl bildungssprachliche Elemente (tokens); Textlänge).⁸ Erhoben wurden hierneben differenzierte Daten zur familialen Sprachpraxis sowie zum Verfügen über und zum Umgang mit bildungsrelevanten Medien. Ferner wurden vermittels eines sehr differenzierten Instruments zur Selbsteinschätzung sprachlicher Fähigkeiten, das auf dem Europäischen Referenzrahmen für Sprachen beruht (Klinger, 2019), Informationen der Mütter zu ihren Verstehens- und Produktionsfähigkeiten im schriftlichen und mündlichen Deutsch erhoben.

Schnoors eingehende Analysen führen zu dem Ergebnis, dass die geringere Ausstattung mit strukturellen Ressourcen und der geringere Akkulturationsgrad, den Migrantenfamilien im Vergleich zu einheimischen aufweisen, die geringeren Fähigkeiten im Deutschen weitgehend klären, die die Kinder aus Migrantenfamilien im Vergleich denen deutscher Herkunft aufweisen. Insofern bestätigen sich hier vorliegende Ergebnisse zu maßgeblichen Einflussfaktoren auf die Aneignung bildungsrelevanter Fähigkeiten im Migrationskontext. Zugleich aber zeigen die Analysen, dass der erwartungswidrig hohe Bildungserfolg der Schülerinnen und Schüler vietnamesischer Herkunft nicht mit ihren bildungssprachlichen Deutschfähigkeiten zu begründen ist (Schnoor, 2018, S. 111). Signifikante Effekte auf die Fähigkeiten der Kinder im Deutschen ergeben sich in der vietnamesischen Herkunftsgruppe lediglich bei den Faktoren „Besuch einer vorschulischen Einrichtung“ und bei den selbsteingeschätzten Deutschfähigkeiten der Mütter. Für beide Migrantengruppen erwiesen sich ferner bildungsrelevante Kontakte als einflussreich auf die Fähigkeiten im Deutschen. Die „Menge“ des Deutschgebrauchs in der Familie zeigte hingegen keinen Effekt. Die Unterschiede im Bildungserfolg der Schüler(innen) türkischer und vietnamesischer Herkunft lassen sich nach diesen Analysen auf Effekte des sozialen Hintergrunds und des Migrationshintergrunds zurückführen (Schnoor, 2018, S. 113), wohinge-

7 Der folgende Bericht beruht auf den Analysen zur Inaugural-Dissertation von Birger Schnoor mit dem Titel *Soziale Herkunft und Bildungssprache. Intergenerationale Humankapitalinvestitionen in Familien deutscher, türkischer und vietnamesischer Herkunft in Deutschland*, deren Publikation in Vorbereitung ist (2019). Ich danke Birger Schnoor für die Möglichkeit, von den Ergebnissen seiner Untersuchung berichten zu dürfen.

8 Messinvarianz nach Altersgruppen und Herkunftsgruppen ist gegeben; vgl. Schnoor (2018), Kapitel 3.2.1.2 und 3.2.1.3.

gen Unterschiede zur Referenzgruppe ohne Migrationshintergrund die geringere Ausstattung der Migranten mit strukturellen Ressourcen und ihren geringeren Akkulturationsgrad widerspiegeln.

Mit Bezug auf den familialen Sprachgebrauch lässt sich aus diesen Analysen folgern, dass es für den potentiellen Bildungserfolg nicht auf das Quantum, sondern auf die Qualität des sprachlichen Inputs und auf die Bereitstellung von Lerngelegenheiten ankommt, die Kinder in ihren Familien erfahren. Die von Schnoor erzielten Ergebnisse sind im Einklang mit solchen aus internationalen Studien, die bereits darauf verwiesen, dass es für die (bildungs-)sprachliche Entwicklung von Kindern weniger entscheidend ist, welche Sprache(n) in der Familie gesprochen wird (oder werden). Relevant sei vielmehr die Bildungsnähe des sprachlichen Inputs, den Kinder erhalten – gleichgültig, in welcher Sprache dieser gegeben werde (Leseman, Scheele, Mayo & Messer, 2007).

Beide grob umrissenen Ergebnisse der HeBe-Studie sprechen dafür, dass die Auseinandersetzung mit Faktoren, die Bildungserfolg begünstigen oder beeinträchtigen, durchaus nicht mit der Beachtung des sozio-ökonomischen Feldes erledigt ist. Insbesondere dann, wenn es um die Frage geht, welche Interventionsmöglichkeiten dem Erziehungs- und Bildungssektor zur Verfügung stehen, um herkunftsbedingte Benachteiligung zu beeinflussen, wird es relevant, Faktoren einzubeziehen, die für Migration spezifisch sind. Es kann also das Fazit gezogen werden, dass auf die Frage nach den Folgen der Migration für Bildung und Erziehung ein guter Bestand an empirisch gestütztem Wissen vorhanden ist. Aber vieles harrt zugleich der Klärung, auch in dem breit beackerten Feld der Klärung von Gründen für soziale Ungleichheit.

Literatur

- Ammon, U. (1972). *Dialekt, soziale Ungleichheit und Schule*. Weinheim: Beltz.
- Auernheimer, G. (2003). *Einführung in die Interkulturelle Pädagogik*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Bade, K. J. (2007). Versäumte Integrationschancen und nachholende Integrationspolitik. In K. J. Bade & H.-G. Hiesserich (Hrsg.), *Nachholende Integrationspolitik und Gestaltungsperspektiven der Integrationspraxis*. Göttingen: V&R unipress.
- Baros, W. & Kempf, W. (Hrsg.) (2014). *Erkenntnisinteressen, Methodologie und Methoden interkultureller Bildungsforschung*. Berlin: Regener.
- Barth, F. (Hrsg.) (1969). *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*. Oslo: Universitetsforlaget.
- Baumert, J. (Hrsg.) (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Wiesbaden: Springer VS.
- Baumert, J., Bos, W. & Lehmann, R. (Hrsg.) (2000). *TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Pflichtschulzeit*. Opladen: Leske+Budrich.
- Baumert, J. & Schümer, G. (2002). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb im nationalen Vergleich. In Deutsches PISA-Konsortium (Hrsg.), *PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich* (S. 159–202). Opladen: Leske+Budrich.

- Becker, B. (2010). *Bildungsaspirationen von Migranten: Determinanten und Umsetzung in Bildungsergebnisse* (MZES Arbeitspapiere, 137). Mannheim: MZES.
- Becker-Mrotzek, M. & Roth, H.-J. (Hrsg.) (2017). *Sprachliche Bildung. Grundlagen und Handlungsfelder*. Münster: Waxmann.
- Bernstein, B. (1977). *Class, codes and control. Towards a theory of educational transmissions*. London: Routledge & Kegan.
- Boos-Nünning, U. & Gogolin, I. (1988). Sprachdiagnose bei ausländischen Schulanfängern. Resultate der empirischen Prüfung eines ‚Sprachtests‘. *Deutsch lernen*, 3/4, 3–71.
- Boos-Nünning, U., Hohmann, M., Reich, H. H. & Wittek, F. (1983). *Aufnahmeunterricht, Muttersprachlicher Unterricht, Interkultureller Unterricht*. München: Oldenbourg.
- Bos, W. (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Cortina, K. S., Baumert, J., Leschinsky, A. & Mayer, K. U. (2005). *Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick*. Reinbek: Rowohlt Taschenbuch.
- Diefenbach, H. & Nauck, B. (2000). Der Beitrag der Migrations- und Integrationsforschung zur Entwicklung der Sozialwissenschaften. In I. Gogolin & B. Nauck (Hrsg.), *Migration, gesellschaftliche Differenzierung und Bildung* (S. 37–52). Opladen: Leske+Budrich.
- Diehl, C., Hunkler, C. & Kristen, C. (Hrsg.) (2016). *Ethnische Ungleichheiten im Bildungsverlauf. Mechanismen, Befunde, Debatten*. Wiesbaden: Springer VS.
- Gogolin, I. (2008). *Der monolinguale Habitus der multilingualen Schule*. Münster: Waxmann.
- Gogolin, I. (2016). Interkulturelle Bildungsforschung. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (S. 263–279). Wiesbaden: Springer.
- Gogolin, I. (2017). Sprachliche Bildung als Feld von sprachdidaktischer und erziehungswissenschaftlicher Forschung. In M. Becker-Mrotzek & H.-J. Roth (Hrsg.), *Sprachliche Bildung. Grundlagen und Handlungsfelder* (S. 37–54). Münster: Waxmann.
- Gogolin, I. & Duarte, J. (2016). Bildungssprache. In J. Kilian, B. Brouèr & D. Lüttenberg (Hrsg.), *Handbuch Sprache in der Bildung*. (Handbücher Sprachwissenschaft, 21, S. 478–499). Berlin: Walter de Gruyter
- Gogolin, I. & Krüger-Potratz, M. (2019). *Einführung in die Interkulturelle Pädagogik*. Opladen: Barbara Budrich (UTB).
- Gogolin, I. & Nauck, B. (Hrsg.) (2000). *Migration, gesellschaftliche Differenzierung und Bildung*. Opladen: Leske+Budrich.
- Grabowski, J. (Hrsg.) (2014). *Fähigkeitskonzepte im Bereich von Sprache, Medien und Kultur. Sinn und Unsinn von Kompetenzen*. Opladen: Barbara Budrich.
- Haag, N., Kocaj, A., Jansen, M., & Kuhl, P. (2017). Soziale Disparitäten. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 213–236). Münster: Waxmann.
- Hadjar, A. & Hupka-Brunner, S. (Hrsg.) (2013). *Geschlecht, Migrationshintergrund und Bildungserfolg*. Weinheim: Beltz Juventa.
- Hamburger, F., Seus, L. & Wolter, O. (1981): Über die Unmöglichkeit, Politik durch Pädagogik zu ersetzen. *Unterrichtswissenschaft*, 9, 158–167.
- Heckmann, F. (1982). Minderheiten. Begriffsanalyse und Entwicklung einer historisch-systematischen Typologie. In J. Swift (Hrsg.), *Bilinguale und multikulturelle Erziehung* (S. 93–117). Würzburg: Königshausen und Neumann.
- Helbig, M. & Schneider, T. (2014). *Auf der Suche nach dem katholischen Arbeitermädchen vom Lande. Religion und Bildungserfolg im regionalen, historischen und internationalen Vergleich*. Wiesbaden: Springer VS.

- Herwartz-Emden, L. (1986). *Türkische Familien und Berliner Schule. Die deutsche Schule im Spiegel von Einstellungen, Erwartungen und Erfahrungen türkischer Eltern – eine empirische Untersuchung*. Berlin: Express Edition.
- Herwartz-Emden, L. (1991). Migrantinnen und ihre Familien in der Bundesrepublik Deutschland. Ein Bericht zum Forschungsstand. *Ethnizität und Migration* 2(7), 5–29.
- Herwartz-Emden, L. & Westphal, M. (2000). Methodische Fragen in interkulturellen Untersuchungen. In I. Gogolin & B. Nauck (Hrsg.), *Migration, gesellschaftliche Differenzierung und Bildung*. Opladen: Leske + Budrich.
- Hohmann, M. (Hrsg.) (1976). *Unterricht mit ausländischen Kindern*. Düsseldorf: Pädagogischer Verlag Schwann (Publikation alfa).
- Kalter, F. (2008). Ethnische Ungleichheiten auf dem Arbeitsmarkt. In M. Abraham & T. Hinz (Hrsg.), *Arbeitsmarktsoziologie. Probleme, Theorien, empirische Befunde* (S. 303–332). Wiesbaden: Springer VS.
- Kempert, S., Edele, A., Rauch, D., Wolf, K. M. Paetsch, J., Darsow, A., Maluch, J. & Stanat, S. (2016). Die Rolle der Sprache für zuwanderungsbezogene Ungleichheiten im Bildungserfolg. In C. Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf. Mechanismen, Befunde, Debatten* (S. 157–241). Wiesbaden: Springer VS.
- Kilian, J., Brouèr, B. & Lüttenberg, D. (Hrsg.) (2016). *Handbuch Sprache in der Bildung* (Handbücher Sprachwissenschaft, 21). Berlin: Walter de Gruyter.
- Klemm, Klaus (1979). Ausländerkinder in deutschen Schulen – Zahlen und Prognosen. In G. Hansen, K. Klemm & T. Rauschenbach, (Hrsg.), *Kinder ausländischer Arbeiter* (S. 31–44). Essen: Neue Deutsche Schule Verlag.
- Klinger, T. (2019). Selbsteinschätzung von Sprachfähigkeiten: Erste Schritte der Entwicklung eines Instruments. In T. Klinger, J. Duarte, I. Gogolin, B. Schnoor & M. Lagemann (Hrsg.), *Sprachentwicklung im Kontext von Mehrsprachigkeit. Hypothesen, Methoden, Forschungsperspektiven*. Wiesbaden: Springer VS.
- Klinger, T., Duarte, J., Gogolin, I., Schnoor, B. & Lagemann, M. (Hrsg.) (2019). *Sprachentwicklung im Kontext von Mehrsprachigkeit. Hypothesen, Methoden, Forschungsperspektiven*. Wiesbaden: Springer VS.
- Krüger-Potratz, M. (1994). „Dem Volke eine andere Muttersprache geben“. Zur pädagogischen Diskussion über Zwei- und Mehrsprachigkeit in der Geschichte der Volksschule. *Zeitschrift für Pädagogik* 40(1), 81–96.
- Krüger-Potratz, M. (1999). Stichwort. Erziehungswissenschaft und kulturelle Differenz. *Zeitschrift für Erziehungswissenschaft* 2(2), 149–165.
- Krüger-Potratz, M. (2000). Schulpolitik und ‚fremde Kinder‘. Eine Schule, ‚die die eigene Art schützt und die fremde achtet‘. In I. Gogolin & B. Nauck (Hrsg.), *Migration, gesellschaftliche Differenzierung und Bildung* (S. 365–384). Opladen: Leske + Budrich.
- Krüger-Potratz, M. (Hrsg.) (2005). *Interkulturelle Bildung. Eine Einführung*. Münster: Waxmann.
- Krüger-Potratz, M., Hansen, G. & Jasper, D. (1991). *Anderssein gab es nicht. Ausländer und Minderheiten in der DDR*. Münster: Waxmann.
- Lee, J. & Zhou, M. (2014). The success frame and achievement paradox: The costs and consequences for Asian Americans. *Race and Social Problems*, 6, 38–55.
- Lemke Muniz de Faria, Y.-C. (2002). *Zwischen Fürsorge und Ausgrenzung. Afro-deutsche „Besatzungskinder“ im Nachkriegsdeutschland* (Metropol Dokumente, Texte, Materialien, 43). Berlin: Metropol.
- Leseman, P. P. M., Scheele, A. F., Mayo, A. Y. & Messer, M. H. (2007). Home literacy as special language environment to prepare children for school. *Zeitschrift für Erziehungswissenschaft*, 10(3), 334–355.

- McElvany, N., Ohle, A., El-Khechem, W., Hardy, I. & Cinar, M. (2017). Förderung sprachlicher Kompetenzen. Das Potential der Familiensprache für den Wortschatzerwerb aus Texten. *Zeitschrift für Pädagogische Psychologie*, 31(1), 13–25.
- Mehringer, V. & Herwartz-Emden, L. (2018). Bildungsforschung. In I. Gogolin, V. Georgi, M. Krüger-Potratz, D. Lengyel & U. Sandfuchs (Hrsg.), *Handbuch Interkulturelle Pädagogik* (S. 194–198). Bad Heilbrunn: Klinkhardt.
- Müller, A. & Stanat, P. (2006). Schulischer Erfolg von Schülerinnen und Schülern mit Migrationshintergrund. Analysen zur Situation von Zuwanderern aus der ehemaligen Sowjetunion und aus der Türkei. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen. Vertiefte Analysen im Rahmen von PISA 2000* (S. 221–255). Wiesbaden: Springer VS.
- Müller, H. (Hrsg.) (1974). *Ausländerkinder in deutschen Schulen. Ein Handbuch*. Stuttgart: Klett.
- Nauck, B. (1985). *Arbeitsmigration und Familienstruktur. Eine Analyse der mikrosozialen Folgen von Migrationsprozessen*. Frankfurt a. M.: Campus.
- Nauck, B. (1988). Sozialstrukturelle und individualistische Migrationstheorien. Elemente eines Theorienvergleichs. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 40, 15–39.
- Nauck, B. (2009). Sozialtheorie und Gesellschaftstheorie: Ein problematischeres Verhältnis. In P. Hill, F. Kalter, J. Kopp, C. Kroneberg & R. Schnell (Hrsg.), *Hartmut Essers erklärende Soziologie. Kontroversen und Perspektiven* (S. 252–281). Frankfurt a. M.: Campus.
- Nauck, B. (2011). Kulturelles und soziales Kapital als Determinanten des Bildungserfolgs bei Migranten? In R. Becker (Hrsg.), *Integration durch Bildung. Bildungserwerb von jungen Migranten in Deutschland* (S. 71–93). Wiesbaden: VS.
- Nauck, B. & Lotter, V. (2014). Bildungsspezifisches Sozialkapital in einheimischen, türkischen und vietnamesischen Familien in Deutschland. In A. Steinbach, M. Hennig & O. Arránz Becker (Hrsg.), *Familie im Fokus der Wissenschaft* (S. 225–253). Wiesbaden: Springer.
- Nauck, B. & Özel, S. (1986). Erziehungsvorstellungen und Sozialisationspraktiken in türkischen Migrantenfamilien. Eine individualistische Erklärung interkulturell vergleichender empirischer Befunde. *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie*, 6, 285–312.
- Nauck, B. & Schnoor, B. (2015). Against all odds? Bildungserfolg in vietnamesischen und türkischen Familien in Deutschland. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 67(4), 633–657.
- Neumann, U. (1980). *Erziehung ausländischer Kinder*. Düsseldorf: Schwann.
- Oltmer, J. (2017). *Migration. Geschichte und Zukunft der Gegenwart*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Poarch, G. J. & Bialystok, E. (2017). Assessing the implications of migrant multilingualism for language education. *Zeitschrift für Erziehungswissenschaft*, 20(2), 175–191.
- Reich, H. H., Roth, H.-J. & Neumann, U. (Hrsg.) (2007). *Sprachdiagnose und Förderung* (Reihe FörMig edition, Band 3). Münster: Waxmann.
- Reich, H. H., Gogolin, I., Krüger-Potratz, M. & Neumann, U. (1990). Folgen der Arbeitsmigration für Bildung und Erziehung FABER. Kurzfassung des Antrags an die DFG. *Deutsch lernen*, 1, 70–88.
- Reich, H. H. & Merkens, H. (1993). Folgen der Arbeitsmigration für Bildung und Erziehung. *Unterrichtswissenschaft*, 21(2), 100–105.
- Reich, H. H. & Wittek, F. (Hrsg.) (1984). *Migration, Bildungspolitik und Pädagogik* (Berichte und Materialien der Forschungsgruppe ALFA Nr. 16). Essen: ALFA.
- Roth, H.-J. (2010). Vom Suchhorizont zur Querschnittsaufgabe. Überlegungen zur Positionierung Interkultureller Bildung im Übergang zur Diversity Education. In M. Krüger-Potratz, U. Neumann & H. H. Reich (Hrsg.), *Bei Vielfalt Chancen-*

- gleichheit. *Interkulturelle Pädagogik und durchgängige Sprachbildung* (S. 90–99). Münster: Waxmann.
- Ruhloff, J. (1983). Bildung und national-kulturelle Orientierung. *Rassegna di Pedagogia: Pädagogische Umschau*, XLI, 249–261.
- Schneider, F. (1961). *Vergleichende Pädagogik*. Heidelberg: Quelle & Meyer.
- Schnoor, B. (2018). *Soziale Herkunft und Bildungssprache. Intergenerationale Humankapitalinvestitionen in Familien deutscher, türkischer und vietnamesischer Herkunft in Deutschland*. Inaugural-Dissertation. Universität Hamburg, Hamburg. Fakultät für Erziehungswissenschaft. Wiesbaden: Springer VS.
- Segeritz, M., Walter, O. & Stanat, P. (2010). Muster des schulischen Erfolgs von jugendlichen Migranten in Deutschland: Evidenz für segmentierte Assimilation? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62(1), 113–138.
- Stanat, P. (2006). Disparitäten im schulischen Erfolg: Forschungsstand zur Rolle des Migrationshintergrunds. *Unterrichtswissenschaft*, 34(2), 98–124.
- Statistisches Bundesamt (2010). *Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund* (Reihe 2.2., Ergebnisse des Mikrozensus 2009, Fachserie 1). Wiesbaden: Destatis.
- Statistisches Bundesamt (2014). *Bildung und Kultur. Allgemeinbildende Schulen* (Fachserie 11, Reihe 1). Wiesbaden. Destatis.
- Tenorth, H.-E. (2016). Bildungstheorie und Bildungsforschung, Bildung und kulturelle Basiskompetenzen – ein Klärungsversuch, auch am Beispiel der PISA-Studien. In J. Baumert & K.-J. Tillmann (Hrsg.), *Empirische Bildungsforschung. Der kritische Blick und die Antwort auf die Kritiker. Zeitschrift für Erziehungswissenschaft, Sonderheft 31*, 45–71.
- Walter, O. (2011). Der Schulerfolg vietnamesischer und philippinischer Jugendlicher in Deutschland. Eine Analyse auf der Grundlage der Erweiterungsstichprobe von PISA 2003. *Zeitschrift für Erziehungswissenschaft*, 14(3), 397–419.
- Wessel, L. & Prediger, S. (2018). Brauchen mehrsprachige Jugendliche eine andere fach- und sprachintegrierte Förderung als einsprachige? Differentielle Analysen zur Wirksamkeit zweier Interventionen in Mathematik. *Zeitschrift für Erziehungswissenschaft*, 21(2), 361–382.
- Yakut, A., Reich, H. H., Neumann, U. & Boos-Nünning, U. (1986). *Zwischen Elternhaus und Arbeitsamt. Türkische Jugendliche suchen einen Beruf*. Berlin: Express Edition.

Knut Schwippert

Was wird aus den Büchern? Sozialer Hintergrund von Lernenden und Bildungsungleichheit aus Sicht der international vergleichenden Erziehungswissenschaft

Zusammenfassung

Internationale empirische Bildungsstudien sind vor allem dazu angelegt, auf Eigenschaften von Bildungssystemen aufmerksam zu werden, die sich erst im Vergleich mit anderen Bildungssystemen zeigen, da beispielsweise innerhalb eines Systems keine Varianz bezüglich eines Merkmals besteht oder aufgrund einer fehlenden Normvorstellung das Ausmaß eines Effekts innerhalb eines Systems nicht bewertet werden kann. Eine besondere Aufmerksamkeit hat bei diesen Betrachtungen die Frage der Abhängigkeit des Schulerfolgs vom sozialen Hintergrund erfahren. Zur Erfassung des sozialen Hintergrunds hat sich insbesondere in westeuropäisch geprägten Staaten eine Orientierung an der Kapitaltheorie von Bourdieu etabliert, in der eine Unterscheidung des Status anhand von ökonomischen, sozialen und kulturellen Merkmalen (Kapital) vorgenommen wird. In internationalen Bildungsstudien hat sich die Abfrage des heimischen Buchbesitzes als gut zu erhebender und robuster Indikator für kulturelles und ökonomisches Kapital bewährt. Vor dem Hintergrund der rasanten Entwicklung digitaler Medien und dessen Einzug in private Haushalte und Schulen stellt sich die Frage, ob der Buchbesitz überhaupt noch als geeigneter Indikator für den ökonomisch-kulturellen Hintergrund von Familien genutzt werden kann. Der Beitrag geht dieser Frage anhand einer Analyse der deutschen Teildatensätze der Internationalen Grundschul-Lese-Untersuchung (IGLU) von 2001 bis 2016 nach. Anhand von probabilistischen Testmodellen wird die Messeigenschaft der Büchervariablen im Verlauf von 15 Jahren nachgezeichnet. Hierbei stellt sich die Information über den heimischen Buchbesitz als robuster und stabiler Indikator zur Beschreibung des bildungsrelevanten familiären Hintergrunds im zeitlichen Verlauf dar.

Prof. Dr. Knut Schwippert, Fakultät für Erziehungswissenschaft, EW 1: Allgemeine und International Vergleichende Erziehungswissenschaft, Universität Hamburg, Von-Melle-Park 8, 20146 Hamburg, Deutschland
E-Mail: knut.schwippert@uni-hamburg.de

Schlagworte*Heimischer Buchbesitz; kultureller Hintergrund; digitale Medien***What's about the books? Social background of students and educational opportunities from the perspective of international large-scale surveys****Abstract**

International large-scale assessments are designed to identify characteristics of educational systems which only can be identified by comparing different systems. The comparison is necessary once within a country characteristic does not show any variance or if normative standards are not available to that a characteristic could be related. Special attention in this context is given to the relation of social background of students and their success in school. As theoretical framework – especially in European countries – the work from Pierre Bourdieu is often used in which social, cultural and economic capitals are distinguished. In international large-scale surveys the question to students and/or their parents about the books at home is tradition since some decades. This indicator is an easy to survey and robust indicator for cultural and economic capital. Taken into account the accelerating availability of digital media in the households the question rises, if the question on the books at home is still appropriate. This questions lies in the focus of the article. Answers are sought by analyzing background information and form the German data of the Progress in International Reading Literacy Study (PIRLS) in the four cycles from 2001 to 2016. By using models of probabilistic item response theory the characteristics of the indicators available – which address the cultural and economic background of the students – are investigated over a 15 year period. The results show that the information about the books at home is still a robust and stable indicator for the description of the educational background of the students.

Keywords*Books at home; Cultural background; Digital media***1. Entwicklung und Ziele von groß angelegten Schulvergleichsuntersuchungen**

Ende der 1950er Jahre wurde durch eine Gruppe von interessierten Personen die Frage aufgegriffen, ob Bildungssysteme anhand quantitativer empirischer Untersuchungen verglichen werden können. Im Rahmen einer 1959 bis 1961 durchgeführten Machbarkeitsstudie in 12 Staaten (Foshay, 1962) wurden sowohl Leistungsdaten als auch Hintergrundmerkmale erhoben. Aus dieser Initiative entwickelte sich die International Association for the Evaluation of Educational

Achievement (IEA), die bis heute namhafte internationale Bildungsstudien initiiert und organisiert (Husén & Postlethwaite, 1996). Da die IEA als Wegbereiter solcher Studien angesehen werden kann, liegt der Fokus dieses Beitrags auf der Initiierung und Entwicklung von IEA-Studien. Die damals wie heute im Fokus stehende Frage ist, anhand welcher Merkmale Staaten bzw. deren Bildungssysteme angemessen verglichen werden können. Eine immer wiederkehrende und in vielen der teilnehmenden Staaten formulierte Frage richtet sich darauf, ob bzw. in welchem Umfang der Lernerfolg von Schülerinnen und Schülern von Merkmalen des sozialen Hintergrunds abhängt. Zum einen geht es um die Frage, wie stark sich der Zusammenhang zwischen den Staaten gleicht bzw. unterscheidet und ob ggf. weitere Merkmale identifiziert werden können, die die Unterschiedlichkeit zwischen den Staaten erklären helfen. Dazu werden die Tests und Befragungsinstrumente in die jeweiligen Landessprachen übersetzt. Auch wenn dies mit hohem Aufwand (Übersetzungskontrolle und Rückübersetzungen zur Qualitätssicherung) betrieben wird (Ebbs & Wry, 2017), bleibt offen, ob die augenscheinlich gleich erhobenen Informationen auch in allen Staaten Gleiches erfassen – mit anderen Worten, ob sie gleich zu interpretieren sind und damit auch als valide Information zum Vergleich herangezogen werden können. Dies sind einerseits materielle Ausstattungsmerkmale und andererseits auch Merkmale von Prozessen bzw. Personen. Der Ansatz, anhand einiger empirisch gut zu erfassender Indikatoren bzw. Skalen im Rahmen eines Modells ein reduziertes empirisches Abbild der Wirklichkeit darzustellen, ist zum Standard quantitativer empirischer Bildungsforschung geworden (Rost, 2004). Im Sinne der Sparsamkeitsprämisse der Wissenschaft wird hierbei Wert daraufgelegt, eine reale Situation bzw. einen realen Prozess mit so wenigen Merkmalen wie möglich zu erfassen und zu analysieren. Bei diesem Ansatz geht es nicht – wie bei anderen empirischen Vorgehensweisen – um die umfangreiche und möglichst vollständige Beschreibung von Situationen oder Prozessen, sondern um die exemplarische und sparsame Beschreibung von Gesetzmäßigkeiten, die in Modellen als reduziertes Abbild der Wirklichkeit beschrieben und geprüft werden kann.

Neben der inhaltlichen und der am jeweiligen Erkenntnisinteresse ausgerichteten Auswahl der zu erfassenden Informationen spielt deren empirische Güte eine zentrale Rolle. So wird die Anwendung der (Haupt-)Testgütekriterien Objektivität, Reliabilität und Validität nicht mehr hinterfragt – was nicht heißt, dass diese Kriterien in ihrer Auslegung in den letzten Jahrzehnten gleichgeblieben sind (Kane, 2013). Die modellorientierte Entwicklung von groß angelegten internationalen Bildungsstudien unterliegen der Herausforderung, die für die Modellprüfungen notwendigen Informationen sowohl national als auch international so zu erfassen, dass Vergleiche sowohl innerhalb als auch zwischen Staaten möglich werden und darüber hinaus – bei Trenduntersuchungen – auch Analysen über die Zeit hinweg realisiert werden können.

1.1 Vergleich – von was?

Die Intention der an der Initiierung internationaler empirischer Bildungsstudien beteiligten Wissenschaftlerinnen und Wissenschaftler sind nicht die oft aus diesen Untersuchungen (i. d. R. von Dritten) abgeleiteten *Liga-Tabellen*. Ziel und Zweck dieser Untersuchungen ist es, den Horizont zu erweitern, von dem auf das eigene (nationale) Bildungssystem geschaut werden kann. Auch wenn kein normativer Diskurs oder gesellschaftlicher Konsens bestimmter Phänomene vorliegt, kann im Vergleich mit anderen Staaten beurteilt werde, ob zum Beispiel die Abhängigkeit des Lernerfolgs der Schülerinnen und Schüler vom sozialen Hintergrund als gering, mittel oder stark zu charakterisieren ist. Es gibt Merkmale, die national innerhalb von Bildungssystemen kaum variieren, wodurch eine empirische Prüfung der Wirkung bzw. Wirksamkeit innerhalb des Systems nicht möglich ist. Erst wenn diese Merkmale (unter sonst gleichen Bedingungen) in ihrer Variation über Staatengrenzen hinweg international vergleichend betrachtet werden, lassen sich Aussagen über mögliche Ursachen- und Wirkungszusammenhänge formulieren.

Je nach politischem, historischem und kulturellem Hintergrund hat sich in jedem Staat ein individuelles Bildungssystem entwickelt und etabliert. Ein entsprechender Vergleich von intentional und historisch bedingt variierenden Systemen ist eine Herausforderung. Jeder geplante Vergleich unterliegt der Frage „kann man Äpfel mit Birnen“ vergleichen? Wilfried Bos antwortete einmal auf diese Frage: „Ja – wenn mich der Anteil von Vitamin C in 100 Gramm Fruchtfleisch interessiert, kann ich Äpfel mit Birnen vergleichen“. Die Kunst des internationalen Vergleichs von Bildungssystemen liegt somit darin, relevante und auch tatsächlich vergleichbare Merkmale in den Blick zu nehmen und unter Berücksichtigung verschiedener (bekannter) Rahmenbedingungen in Relation zu setzen. In den international groß angelegten empirischen Bildungsstudien werden neben Leistungsdaten auch Rahmenbedingungen anhand von Fragebögen erfasst, die sich – je nach Fragestellung und Fokus der Untersuchung – an verschiedene Akteure im Bildungssystem richten und aus verschiedenen Wissenschaftsdisziplinen speisen. Im Fokus von IEA-Studien stehen oft Schülerinnen und Schüler, in der Regel ihre Lehrkräfte und die Schulleitungen, aber auch ihre Eltern oder Personen aus der Bildungsadministration und/oder -politik. Die durch die Befragung dieser Akteure geschaffene Datenbasis kann einerseits generelle Überblicke und eine Gegenüberstellung von Merkmalen der Bildungssysteme erlauben und andererseits auch Vergleiche von Bildungssystemen unter Dokumentation des Vergleichsrahmens ermöglichen. Die IEA ist hierbei traditionell an dem gemeinsamen Erkenntnisinteresse der nationalen Bildungsexpertinnen und Bildungsexperten der beteiligten Staaten orientiert. Somit gelingt es (gerechtfertigte) Vergleiche zu beschreiben und gleichzeitig die Grenzen der Vergleiche zu dokumentieren.

1.2 Tests

Ein Hauptaugenmerk im Rahmen von groß angelegten internationalen Schulvergleichsuntersuchungen liegt auf der Betrachtung der durch Schülerinnen und Schüler erreichten Kompetenzen. Kompetenzen werden hierbei als die Fähigkeit beschrieben, in variablen Situationen neuen Herausforderungen begegnen zu können (Klieme & Hartig, 2007; Weinert, 1999). Seit den ersten Studien der IEA liegt hierbei der Fokus auf den in Regelschulsystemen als zentral identifizierten Domänen: Mathematik, Naturwissenschaft (in Deutschland insbesondere: Physik, Chemie, Geographie, Biologie), Lesen, Schreiben und erste Fremdsprache (wenn diese Englisch ist). In späteren Untersuchungen sind – in Reflexion internationaler gesellschaftlicher Veränderungen – unter anderem auch Untersuchungen zur politischen Bildung, zur Klassenraumführung und (zunächst) Verfügbarkeit und (später) Nutzung von Computern hinzugekommen (Goy, van Ackeren & Schwippert, 2008).

Unter Berücksichtigung nationaler Curricula wird ein gemeinsamer internationaler Rahmen zur Erfassung der Kompetenzen von Schülerinnen und Schülern gesteckt. In diesem Rahmen werden Erwartungen formuliert, was Schülerinnen und Schüler sicher auf mittlerem oder auf hohem Niveau bearbeiten können sollten. Anhand von fachlich variierenden Testaufgaben kann so ein Leistungsspektrum empirisch abgebildet werden. Aktuelle Definitionen der in den Tests erfassten Kompetenzen lauten exemplarisch für die Progress in International Reading Literacy Study (PIRLS) und für die International Computer and Information Literacy Study 2013 (ICILS 2013) wie folgt:

- „The PIRLS definition of reading literacy is grounded in IEA’s 1991 study, in which reading literacy was defined as ‘the ability to understand and use those written language forms required by society and/or valued by the individual’” (Mullis, Martin & Sainsbury, 2015, S. 11).
- Die ICILS 2013-Studie untersucht, welche Kompetenzen Jugendliche im Umgang mit Computern und Informationstechnologien haben, die notwendig sind, sich aktiv in einer technologischen Welt bewegen zu können. Computer and information literacy (ICT) ist definiert als: „an individual’s ability to use computers to investigate, create and communicate in order to participate effectively at home, at school, in the workplace and in society“ (Fraillon, Schulz & Ainley, 2013, S. 17).

Beiden Definitionen ist gemein, dass sie den Lebensweltbezug der betrachteten Kompetenzen betonen.¹ Somit lassen sich die betrachteten Kompetenzen als notwendige – wenn auch nicht hinreichende – Voraussetzungen für Bildung verstehen. Nach Klafkis (2007) Auffassung ist Bildung die Befähigung zur vernünftigen Selbstbestimmung, Mitbestimmung und Solidaritätsfähigkeit, die zur Gewinnung von Individualität und Gemeinschaftlichkeit führt. Solch ein Verständnis von

1 Vgl. auch die entsprechende Definitionen zur Naturwissenschaft (Jones, Wheeler & Centurino, 2013, S. 26) und zur Mathematik (Grønmo, Lindquist, Arora & Mullis, 2013, S. 11).

Bildung setzt in einer modernen Gesellschaft die Beherrschung verschiedener Kulturtechniken voraus. Um in dieser Gesellschaft selbständig, selbstbestimmt aber auch verantwortungsvoll agieren zu können, rücken national wie auch international Kulturtechniken in den Fokus, „die als grundlegende Prämissen für die Teilhabe an gesellschaftlicher Kommunikation durch schulische Arbeit universalisiert werden müssen – als historische Gestalt dessen, was heute allgemeine Bildung heißen kann“ (Tenorth, 2005, S. 26): Erst lesen und schreiben zu können, eröffnet den Individuen die Möglichkeit, an einer modernen Informationsgesellschaft teil zu haben. Auch Kenntnisse in Mathematik und den Naturwissenschaften stellen wichtige Voraussetzungen dar, sich selbstbestimmt und verantwortungsvoll in der Gesellschaft zu bewegen, sich in und mit ihr zu entwickeln (vgl. Schwippert, 2015).

Bei der Erfassung von Kompetenzen in groß angelegten internationalen empirischen Bildungsstudien geht es nicht um eine detaillierte Individualdiagnostik, sondern um die Erfassung von Kompetenzständen in den betrachteten Bildungssystemen. Entsprechend wird das Studiendesign so optimiert, dass valide Aussagen über Gruppen von Schülerinnen und Schülern getroffen werden und nicht differenzierte Aussagen in Bezug auf einzelne Personen. Dies hat den positiven Nebeneffekt, dass – abstrakt betrachtet – die individuelle Tagesform der Schülerinnen und Schüler keine solche Rolle spielt wie beispielsweise im Rahmen von (zentralen) Abschlussprüfungen. Aufgrund der umfangreichen Befragung von Schülerinnen und Schülern kann davon ausgegangen werden, dass sich Tagesform abhängige Unterschiede zwischen „wahrer“ Kompetenz und gezeigter Kompetenz (diese wird auch als Performanz bezeichnet) im Mittel über alle Beobachtungen hinweg ausgleichen. Lediglich systematische Beeinflussungen (z. B. durch *Priming* erzeugte bewusste negativ-Manipulation der Ergebnisse oder in Gruppen induzierte Verweigerungen) würden das Gesamtbild verfälschen. Aber selbst wenn solche Abweichungen einen Einfluss auf die Gesamtergebnisse haben würden, wäre die dahinterliegende Haltung bzw. Einstellung auch als ein Merkmal des Bildungssystems zu interpretieren. Jedoch wäre in diesem Fall wichtig, dass solche Merkmale identifiziert und dokumentiert werden.

1.3 Hintergrundmerkmale

Wie bereits erwähnt ist das Ziel von groß angelegten Schulvergleichsuntersuchungen nicht das Ranking von Staaten. Vielmehr werden Merkmale der Bildungssysteme miteinander in Beziehung gesetzt, um so – aus der Interpretation von Ähnlichkeiten und Unterschieden zwischen Bildungssystemen – Schlussfolgerungen für die Reform oder Weiterentwicklung von Bildungs- bzw. Schulsystemen, Unterricht und Rahmenbedingungen des Lehrens und Lernens ziehen zu können. Je nach Wahrnehmung von be- bzw. anstehenden Herausforderungen für das Bildungssystem werden hierzu bewährte Modelle zur Erklärung von Unterschieden innerhalb und zwischen Bildungssystemen – je nach Ebene des Erkenntnisinteresses – herangezogen.

Schon seit den ersten Untersuchungen der IEA wurden Vergleiche und Relationen vorgestellt, ohne Rangreihen der beteiligten Staaten vorzunehmen (Foshay, 1962; Husén, 1967). Im Fokus stand vielmehr die Identifikation von Mustern im Sinne von Gemeinsamkeiten bzw. Unterschieden, aufgrund deren Reformen bzw. Innovationen in den jeweiligen Bildungssystemen initiiert werden konnten. Eine besondere Rolle hat hierbei seit der ersten internationalen empirischen Bildungsstudie die Frage gespielt, inwieweit die individuellen Voraussetzungen von Schülerinnen und Schülern im Schulsystem eine Rolle spielen. Hierbei steht die Frage im Raum, ob familiär bedingte (soziale) Unterschiede sich durch Schule und Unterricht verringern, gleichbleiben oder sogar vergrößern.

Im Zuge der Weiterentwicklung der IEA-Studien wurde das Repertoire an Hintergrundmerkmalen der Schülerinnen und Schüler – insbesondere orientiert an Befunden aus der Soziologie – weiter ausdifferenziert. Hierbei spielt auch in den aktuellen Studien eine differenzielle Betrachtung des sozialen, kulturellen und ökonomischen Kapitals (vgl. insb. Bourdieu, 1977; Bourdieu, 1983; Coleman, 1988) der Schülerinnen und Schüler international eine zentrale Rolle. Diese soll in diesem Beitrag im Fokus stehen. Eine nähere Betrachtung von Untersuchungen, die sich der Bedeutung und Wirkung von Faktoren auf Ebene der (nationalen) Bildungspolitik bzw. -administration, der Schulen bzw. des Unterrichts und den damit im Zusammenhang stehenden Tätigkeiten von Lehrkräften im internationalen Kontext widmen, muss anderen Beiträgen vorbehalten bleiben, da der Rahmen dieses Beitrags sonst überschritten würde.

Treibende Kraft bei der Analyse von Bildungssystemen ist die stetige Auseinandersetzung mit Fragen zur sozialen Ungleichheit, welche sich kontinuierlich durch das sich wandelnde Verständnis von Unterschiedlichkeit – also Heterogenität – in Schule und Unterricht aktualisiert (vgl. Schwippert, 2011). Das sich ändernde Verständnis ist nicht nur auf sich gesellschaftlich wandelnde Anforderungen (z. B. Gleichheit von Jungen und Mädchen; Inklusion von Schülerinnen und Schülern mit Beeinträchtigungen im Regelschulsystem) zurückführen, sondern auch auf veränderte Aufgaben, die als wandelnde Rahmenbedingungen im zeitlichen Verlauf unterschiedliche Herausforderungen induzieren (z. B. Berücksichtigung der Kinder von angeworbenen ausländischen Arbeitnehmern; fluchtbedingter Migration; Konzentration der Bevölkerung in Ballungsräumen).

Als übliches Repertoire zur Erfassung individueller Heterogenität bei Schülerinnen und Schülern werden neben dem Geschlecht und Migrationserfahrungen in der Familie insbesondere auch bildungsrelevante Hintergrundmerkmale erhoben. Hierzu zählen der Bildungshintergrund der Eltern, die Ausstattung mit bildungsrelevanten Gütern und die Ausstattung mit Gütern, die von dem ausgeübten Beruf der Eltern und der hierdurch induzierten ökonomischen Situation abhängen.

Bereits in der internationalen Reading Literacy Study der IEA, deren Haupterhebung ab 1990 stattfand, wurde die Frage nach dem heimischen Buchbesitz als eine der zentralen und erklärungskräftigsten Indikatoren für das in der Familie vorhandene kulturelle und (auch) ökonomische Kapital erfragt (Elley, 1994; Lehmann, Peek, Pieper & Stritzky, 1991). Der besondere Charme dieses

Indikators liegt darin begründet, dass er von Schülerinnen und Schülern ganz unterschiedlicher Klassenstufen erfragt werden kann, im Gegensatz zu Variablen wie Bildungsabschlüsse und Berufe der Eltern, die von diesen erfragt werden müssen. Im internationalen Vergleich kann damit sowohl in Grundschulen als auch im Sekundarbereich gezeigt werden, wie stark in den verschiedenen beteiligten Staaten die Testergebnisse vom kulturellen bzw. ökonomischen Status der Familien abhängen – und auch in Deutschland erweist sich diese Variable als bedeutsam bei der Untersuchung von Unterschieden zwischen Schülerinnen und Schülern (Stubbe & Goy, 2013). So zeigt sich in Deutschland in der seit 2001 alle fünf Jahre durchgeführten IGLU-Studie und der seit 2007 im vierjährigen Turnus durchgeführten Trends in International Mathematics and Science Study (TIMSS), dass die durch den Buchindikator abgebildete Bildungsungleichheit in Grundschulen, wenn überhaupt nur in wenigen anderen Staaten größer ausfällt als in Deutschland (Hußmann, Stubbe & Kasper, 2017; Stubbe, Schwippert & Wendt, 2016).

1.4 Herausforderung internationaler Vergleich und Dokumentation von Entwicklungen

He und van de Vijver (2013b) weisen darauf hin, dass eine zu stark datenbezogene technische Interpretation von Skalen insbesondere im Kontext international vergleichender oder auch längsschnittlich angelegter empirischer Bildungsstudien zu kurz greifen kann. Sie mahnen eine umsichtige Konstruktion von Fragen und Skalen an, die in Studien eingesetzt werden, wenn Staaten mit unterschiedlichen Kulturen im Bildungssystem gemeinsam in den Blick genommen werden sollen (so wie es in internationalen Vergleichsstudien üblich ist). He und van de Vijver (2013a) verweisen darüber hinaus darauf, dass neben der Notwendigkeit einer genauen Analyse der Items und Skalen (z.B. durch DIF-Modelle (differential item functioning), konfirmatorischen Faktoranalysen oder anhand von Mehrebenenanalysen) auch auf eine standardisierte und international einheitliche Testdurchführung zu achten ist – wie dies etwa bei den IEA-Studien wie TIMSS und PIRLS üblich ist. Darüber hinaus empfehlen sie eine Dekomposition möglicher Quellen für beobachtete Unterschiede (Varianzen), die möglicherweise auf die Konstruktion, die Auswertungsmethode oder auf Einzelitems zurückgeführt werden können. Sie weisen darauf hin, dass die Identifikation unterschiedlicher Varianzquellen notwendig ist, um so Effekte, die auf Individuen, Gruppen oder Staaten zurückzuführen sind, unterscheiden zu können. Einige dieser Unterschiede könnten dann durch kulturvergleichende Analysen erklärt werden, andere wiederum könnten auf tatsächliche (objektive) Unterschiede zwischen den Vergleichsgruppen zurückführbar sein. He und van de Vijver (2013a) haben in ihren Untersuchungen gezeigt, dass bei empirischen Vergleichsuntersuchungen kulturbedingte Hauptfaktoren bei Likert-Skalen identifizierbar waren, welche (unter Berücksichtigung der Überprüfung von Messinvarianz) für kulturvergleichende Analysen genutzt werden konnten. Bei vertiefenden Analysen von internatio-

nal eingesetzten Skalen erscheint es daher notwendig, bei der Interpretation der Befunde auch auf mögliche weitere – bisher nicht explizit modellierte oder beschriebene latente Konstrukte – zu achten und diese auch bei der Interpretation zu berücksichtigen.

Neben der Herausforderung bei der Identifizierung von Indikatoren bzw. Skalen zum internationalen Vergleich von Staaten in groß angelegten Vergleichsuntersuchungen ergibt sich aufgrund des Designs der Studien als Trend-Untersuchungen eine weitere Herausforderung. Seit 2001 wird PIRLS (mit Deutscher Beteiligung seit 2001)² und seit 1995 TIMSS (mit deutscher Beteiligung in der Grundschule seit 2007)³ in wiederkehrenden Zyklen durchgeführt. Wenn Vergleiche einerseits zwischen den Staaten aber auch im zeitlichen Verlauf (insb. innerhalb der Staaten) betrachtet werden sollen, sind diese Indikatoren und Skalen gleichzuhalten. Um mit den Studien jedoch sich wandelnde Bedingungen in dynamischen Schulsystemen angemessen zu erfassen, muss hierbei der Gestaltung und Weiterentwicklung der Hintergrundfragebögen eine besondere Aufmerksamkeit zukommen. In diesem Zusammenhang stellt sich gerade heute die Frage, welche Bedeutung Büchern – in diesem Kontext als einem der zentralen Indikatoren für die Dokumentation von Bildungsungleichheiten – in einer sich medial rasant wandelnden Welt für das Lehren und Lernen von Schülerinnen und Schülern zukommt und welche Möglichkeiten bei der Modellierung bestehen, wenn Veränderungen der Indikatoren festgestellt werden.

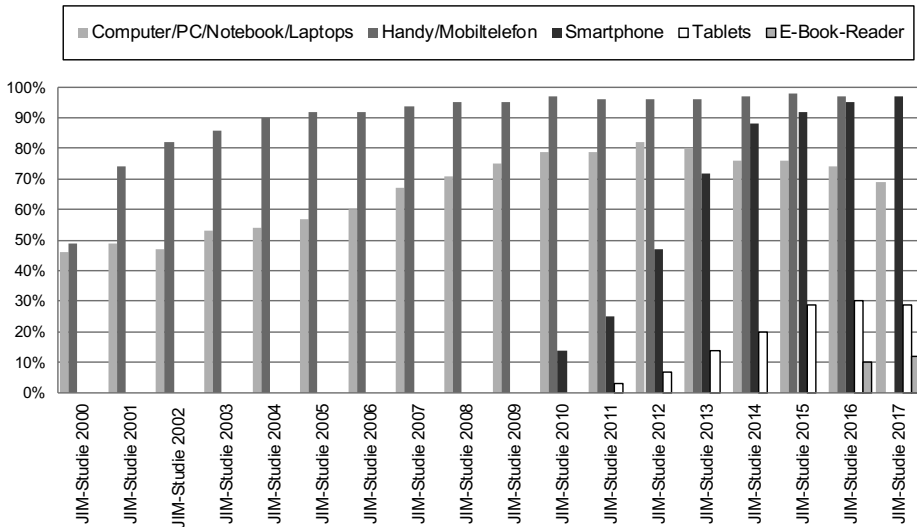
2. Ableitung der Fragestellung und empirische Grundlage

Wie zuvor skizziert nimmt seit Jahren die standardmäßig abgefragte Zahl des heimischen Buchbesitzes eine zentrale Rolle in internationalen Vergleichsuntersuchungen ein. Bei der Interpretation multivariater Zusammenhänge offenbaren sich vielschichtige Bezüge. Neben der numerischen Information über den Umfang des Buchbestandes umfasst die Information über den Buchbesitz auch Merkmale des kulturellen, des sozialen und auch des ökonomischen Kapitals einer Familie (Schulte, Hartig & Pietsch, 2014). Sie steht darüber hinaus auch als Indikator für die Anwendung von Kulturtechniken. Somit steht der Indikator für den heimischen Buchbesitz auch für eine bildungsorientierte Haltung.

2 PIRLS/IGLU (Bos et al., 2007; Bos et al., 2003; Bos, Tarelli, Bremerich-Vos & Schwippert, 2012; Hußmann, Wendt, Bos et al., 2017).

3 TIMSS (Bos et al., 2008; Bos, Wendt, Köller & Selter, 2012; Wendt et al., 2016).

Abbildung 1: Aufstellung zum Wandel der Verfügbarkeit digitaler Medien (eigene Darstellung auf Grundlage der JIM-Ergebnisse von 2000 bis 2017)



Anmerkung. Die Information zum E-Book-Reader wurde erst ab 2016 bei der Befragung berücksichtigt.

Vielfach ist belegt, dass der Indikator des heimischen Buchbesitzes einen überraschend hohen Anteil an Leistungsvarianz auf Individualebene erklärt (z. B. Schwippert, Wendt & Tarelli, 2012) – betrachtet man die Entwicklung der Verbreitung digitaler Medien, erscheint es aus heutiger Sicht nicht mehr reversibel, dass die bisher durch den Buchbesitz erfasste Information über die (zeitgemäße) Anwendung von Basiskompetenzen durch die Betrachtung der Nutzung von digitalen Medien zunächst ergänzt und langfristig möglicherweise auch abgelöst wird. Abbildung 1 aus der Zusammenfassung der von 2000 bis 2017 durchgeführten „Jugend, Information, (Multi-)Media“ (JIM) Studien (mpfs, n.d.) ist zu entnehmen, dass bis 2012 ein kontinuierlicher Zuwachs an Computern zu verzeichnen ist, der ab 2010 durch den Aufwuchs von Smartphones bzw. ab 2011 durch Tablet-Computer ergänzt wird. Diese Medien haben – neben Büchern – eine wichtige Rolle bei der Informationsbeschaffung in Lernkontexten übernommen. Explizit ist ab 2016 auch die Nutzung von E-Book-Readern aufgeführt, deren Verbreitung und Nutzung sicherlich mit einer Abnahme der Anzahl von Printmedien in Haushalten einhergehen wird. Damit ist zu vermuten, dass der Nutzung von digitalen Medien zukünftig eine entsprechende Erklärungskraft bei der Analyse von (Bildungs-) Disparitäten zukommen wird, wie heute noch dem Buchbesitz.

Um zu untersuchen, inwieweit sich die Bedeutung des heimischen Buchbesitzes im Kontext der Erfassung des kulturellen und ökonomischen Kapitals von Familien in Deutschland in den letzten 15 Jahren (zw. 2001 und 2016) verändert hat, wird dessen Eigenschaft im Kontext der Modellierung des latenten Konstrukts des kulturellen und ökonomischen Kapitals anhand der vier Erhebungszyklen von PIRLS

in Deutschland untersucht. Somit wird zunächst *ex post* der Frage nachgegangen, wie sich der Status des heimischen Buchbesitzes empirisch abbildet, um anschließend *ex ante* interpretierend der Frage „Was wird aus den Büchern?“ nachzugehen.

Wenn Skalen in verschiedenen Zyklen einer Untersuchung in den Hintergrundfragebögen wiederholt aufgenommen wurden, bieten probabilistische Testmodelle (item response theory – IRT) gute Möglichkeiten, diese einerseits über Staaten hinweg und andererseits auch im Laufe der Zyklen zu modellieren und diese damit sowohl international als auch im Längsschnitt als vergleichbare Kennwerte verfügbar zu machen. Wie bei Mullis, Martin und Hooper (2016) beschrieben, wurde auch im Kontext der TIMSS- und PIRLS-Untersuchungen IRT-Modelle zur Konstruktion von Skalen eingesetzt.

Geeignete IRT-Modelle bieten die Möglichkeit, auch bei Variation von einzelnen Items zwischen den Staaten oder über die Studienzyklen hinweg, diese bei der Berechnung der Gesamtskalen berücksichtigen zu können (*missing by design*). Anhand der Analysen kann dann festgestellt werden, ob die in den Skalen berücksichtigten Items in der erwarteten Relation zu den anderen erfassten Variablen stehen und damit in der Zusammenfassung das theoretisch abgeleitete latente Konstrukt empirisch angemessen erfasst wird. Diese Analysen der Skalen bietet die Möglichkeit, Abweichungen festzustellen, die in den Testeigenschaften der Items liegen. Werden solche Abweichungen (*bias*) identifiziert, können diese Items entweder aus den Modellen entfernt werden oder – wenn möglich – durch angepasste IRT-Modelle weiterhin berücksichtigt werden. Je nach Ausprägung und Eigenschaft der Abweichung kann dies auch auf eine notwendige Modifizierung der Befragungsinstrumente für zukünftige Erhebungen hinweisen. Um in jedem der realisierten Studienzyklen angemessene und reliable Antworten zu erhalten, bedeuten diese Veränderungen in den Messmodellen bzw. den Skalen jedoch auch, dass ein direkter Vergleich über die Studienzyklen hinweg nicht mehr zulässig wäre. Wird auf die Veränderung der Skaleneigenschaften bei der Interpretation der Befunde keine Rücksicht genommen, kann dies zu Fehlinterpretationen bei der Bewertung von Unterschieden führen. Darüber hinaus stellen Skalen, die nur auf einer geringen Anzahl von Items basieren, eine besondere Herausforderung dar. Gustafsson und Rosén (2014) weisen darauf hin, dass dies insbesondere für Skalen gilt, die im Längsschnitt einen Wandel erfahren haben und nur wenige Ankeritems zur Verbindung der verschiedenen Studienzyklen aufweisen. In ihrer Schlussfolgerung weisen Gustafsson und Rosén (2014) daher darauf hin, dass insbesondere der Auswahl der Ankeritems besondere Beachtung zukommen muss. Die Büchervariable ist bei den nachfolgend vorgestellten Analysen eine dieser Ankeraufgaben, sodass die Betrachtung ihrer psychometrischen Eigenschaften im Fokus liegen wird.

Für die Analysen werden die deutschen Datensätze der IEA PIRLS-Untersuchung (in Deutschland als Internationale Grundschul-Lese-Untersuchung – IGLU bekannt) von 2001 (n = 7633), 2006 (n = 7899), 2011 (n = 4000) und 2016

(n = 3959) genutzt.⁴ Die Schülerinnen und Schüler wurden im Rahmen einer systematischen Stichprobenziehung so bestimmt, dass die Stichprobe als repräsentativ für Deutschland gelten kann (für technische Details zur Stichprobenziehung siehe u.a. LaRoche, Joncas & Foy, 2017). Im Rahmen der IGLU-Studie wurden in Deutschland Viertklässlerinnen und Viertklässler am Ende des Schuljahrs (jeweils im Frühjahr des angegebenen Jahrs der Haupterhebung) befragt. In allen Zyklen wurden den Kindern insgesamt zwei Texte im Umfang von ca. 4 Seiten (abhängig von der Bebilderung der Texte) vorgelegt, zu denen im Anschluss 12 bis 17 Fragen gestellt wurden (Hußmann, Wendt, Kasper, Bos & Goy, 2017). Im Anschluss wurde den Kindern nach der Testung ein Fragebogen zur Erfassung verschiedener Hintergrundinformationen vorgelegt. Testungen und Befragungen wurden durch kindgerechte Pausen unterbrochen. In PIRLS/IGLU wurden zusätzlich auch die Eltern der beteiligten Viertklässlerinnen und Viertklässler befragt, sodass diese Informationen als zusätzliche Hintergrundinformationen über die Kinder zur Verfügung standen.

In allen Zyklen liegen Informationen vor über (1) den heimischen Buchbesitz (5-stufig), (2) den höchsten Bildungsabschluss der Eltern (3-stufig; aus Elternbefragung) und die Information (dichotom), ob (3) zu Hause immer Deutsch gesprochen wird, (4) es einen Computer gibt, (5) das Kind einen eigenen Schreibtisch hat, (6) es eine Tageszeitung und (7) es einen Rasenmäher im Haushalt gibt und ob (8) das Kind ein eigenes Zimmer hat. Damit liegen über alle vier Studienzyklen acht Ankeraufgaben vor. Weitere Indikatoren zur ökonomischen und kulturellen Situation sind z.T. einmalig in einzelnen Zyklen, aber auch wiederholt erfragt worden. Tabelle 1 gibt Auskunft über die Verfügbarkeit der Merkmale für die vier Studienzyklen.

Aus den Hintergrundfragebögen der Schülerinnen und Schüler und ihrer Eltern werden die Informationen für die in diesem Beitrag im Fokus stehenden Analysen (unter Berücksichtigung des *sampling-weights* HOUWGT (Student House Weight)) genutzt. Hierzu werden die durch die Befragungen ermittelten Einzelinformationen zum ökonomischen und kulturellen Hintergrund der Kinder bzw. deren Familien herangezogen, um hieraus einen kombinierten Kennwert zu bilden. Als bewährtes und robustes IRT-Modell wird hierzu das ein-parametrische logistische Modell (Rasch-Modell) in drei Variationen herangezogen. Als Kriterium zur Berücksichtigung der Hintergrundmerkmale (Items) in den Modellen wurde der gewichtete Fit (MNSQ) herangezogen. Bei allen nachfolgend beschriebenen Modellen weisen die berücksichtigten Items⁵ einen sehr guten Fit auf ($0.9 < \text{MNSQ} < 1.1$). Zur Berechnung ist das Programm ConQuest genutzt worden (Wu, Adams, Wilson & Haldane, 2007).

4 Die Daten sind am 1.4.2018 vom PIRLS-Server am Boston College herunter geladen worden [www.pirls.org].

5 Lediglich das Item zur Verfügbarkeit des Mobiltelefons (2006 und 2011) wurde aufgrund schlechter Fit-Werte von den Analysen ausgeschlossen.

Tabelle 1: Erfragte Merkmale der Haushalte in Deutschland in PIRLS 2001, 2006, 2011 und 2016 im Überblick

	2001	2006	2011	2016
Bücher im Haushalt	*	*	*	*
Höchster Bildungsabschluss Eltern	*	*	*	*
Sprachgebrauch	*	*	*	*
Computer	*	*	*	*
Eigener Schreibtisch	*	*	*	*
Tageszeitung	*	*	*	*
Rasenmäher	*	*	*	*
Persönliche Bücher	*	*	*	*
Geschirrspülmaschine	*			
Eigenes Zimmer	*	*	*	*
Computer-Lernprogramme	*	*	*	
Internetzugang	*		*	*
Wörterbuch	*			
Kunstgegenstände (Bilder)	*			
Zweites Auto	*	*	*	
Klavier		*	*	*
Haushaltshilfe				*

Anmerkung. 2006 und 2011 wurde auch nach einem eigenen Mobiltelefon gefragt. Dieses Item wurde aufgrund schlechter FIT-Werte nicht berücksichtigt.

- I) Gemeinsame Modellierung der Hintergrundvariablen mit einheitlichen Itemparametern in allen vier Erhebungszyklen. Berücksichtigung der Itemschwierigkeit bei dichotomen Items und Modellierung von polytomen Items durch Bestimmung der Übergangsschwellen zwischen den Kategorienfunktionen (vgl. Rost, 2004).
(ConQuest Modell: $\text{item} + \text{item}^*\text{step}$).
- II) Gemeinsame Modellierung aller erfragten Hintergrundvariablen mit Ermittlung der generellen Itemschwierigkeiten und der Unterschiede zwischen den Erhebungszyklen. Zusätzlich Bestimmung zyklenspezifischer Schwellenparameter.
(ConQuest Modell: $\text{item} + \text{ncycle} + \text{item}^*\text{step} + \text{item}^*\text{step}^*\text{ncycle}$).

III) Gemeinsame Modellierung ausschließlich der Hintergrundvariablen, die in allen Erhebungszyklen gefragt wurden. Die polytomen Items (Bücher und höchster Bildungsabschluss der Eltern) wurden zur Verbesserung der Fit-Werte dichotomisiert;⁶ anschließend wurden die Itemschwierigkeiten, die Zykleneffekte und die zyklenspezifischen Itemparameter modelliert. (ConQuest Modell: $\text{item} + \text{ncycle} + \text{item} * \text{ncycle}$).

3. Ergebnisse

Model I: Prüfung allgemeiner Itemeigenschaften unter Berücksichtigung der Daten aus vier Zyklen: Im ersten Analyselauf wurden die Itemparameter aller berücksichtigten Variablen bestimmt. Tabelle 2 sind die Itemschwierigkeiten zu entnehmen. Für die polytomen Items zum heimischen Buchbesitz (ASBGBOOK) und des höchsten Bildungsabschlusses der Eltern (H_ISCED_R) wurden zusätzlich die Schwellenparameter berechnet. In Abbildung 2 sind für alle dichotomen Items die Itemschwierigkeiten und für die polytomen Items die Schwellenparameter gemäß Thurstone (Wu et al., 2007) abgetragen. Dieser ist zu entnehmen, dass sich die vier Schwellen der Büchervariablen (1.1 bis 1.4) gleichmäßig in der Verteilung der Itemschwierigkeiten verteilen. Insgesamt zeigt sich eine gute Streuung der Items im Vergleich zur Verteilung der Personenparameter, auch wenn eine etwas dichtere Indikatorisierung im mittleren Bereich wünschenswert wäre.

Model II: Differenzielle Schwellenparameter der Büchervariablen nach Zyklen: In einem weiteren Modell wird der generelle Unterschied der Itemschwierigkeiten für die vier Zyklen überprüft. Die Analysen (vgl. Tabelle 3) zeigen, leichte Verschiebungen der Itemparameter für die Erhebungen von 2001, 2006, 2011 und 2016 – diese sind jedoch numerisch als sehr gering anzusehen (auch wenn sie aufgrund der sehr großen Stichprobenumfänge als signifikant gekennzeichnet sind). Diese Unterschiede könnten auch auf Unterschiede in den Schülerinnen- und Schülerstichproben in den verschiedenen Zyklen zurückzuführen sein – eine Gegenüberstellung der mittleren Personenparameter für die Zyklen (2001 = 1,068; 2006 = 1,053; 2011 = 1,054; 2016 = 1,066) zeigt jedoch nur marginale Unterschiede, die deutlich unter den festgestellten Unterschieden in den mittleren Itemschwierigkeiten der vier Zyklen (vgl. Tabelle 4) liegen.

6 Dichotomisierung: Buchbesitz 0 = bis 100 Bücher vs. 1 = mehr als 100 Bücher; höchster Bildungsabschluss 0 = bis Universitätsbesuch ohne BA Abschluss vs. 1 = Universitätsabschluss BA und höher.

Tabelle 2: Itemschwierigkeiten und Fit-Werte der Skalierung der Hintergrundmerkmale für die Deutschen PIRLS-Daten von 2001, 2006, 2011 und 2016

Item	Item-Schwierigkeit	S.E.	weighted MNSQ
Bücher im Haushalt	0,785	0,007	1,08
Sprachgebrach	-0,685	0,011	1,02
Computer	-1,172	0,011	0,96
Eigener Schreibtisch	-1,918	0,012	0,96
Persönliche Bücher	-2,282	0,012	0,95
Tageszeitung	0,586	0,010	0,96
Geschirrspülmaschine	0,072	0,012	0,94
Eigenes Zimmer	-0,731	0,011	0,96
Computer-Lernprogramme	0,097	0,010	1,04
Internetzugang	0,276	0,011	1,06
Wörterbuch	-1,522	0,012	1,00
Kunstgegenstände (Bilder)	-0,487	0,012	1,03
Rasenmäher	0,384	0,010	0,91
Zweites Auto	1,079	0,010	0,98
Klavier	2,487	0,011	0,95
Haushaltshilfe	1,871	0,012	1,07
Höchster Bildungsabschluss Eltern	1,160	0,043	1,02

Abbildung 2: Stem-and-leaf Darstellung von Personenparametern und Itemschwierigkeiten für die Skalierung der Hintergrundmerkmale für die Deutschen PIRLS-Daten von 2001, 2006, 2011 und 2016

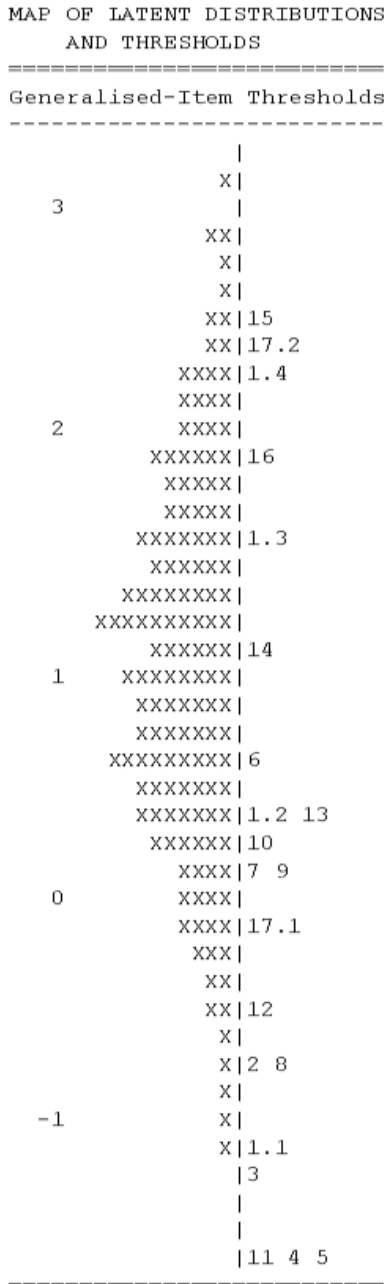
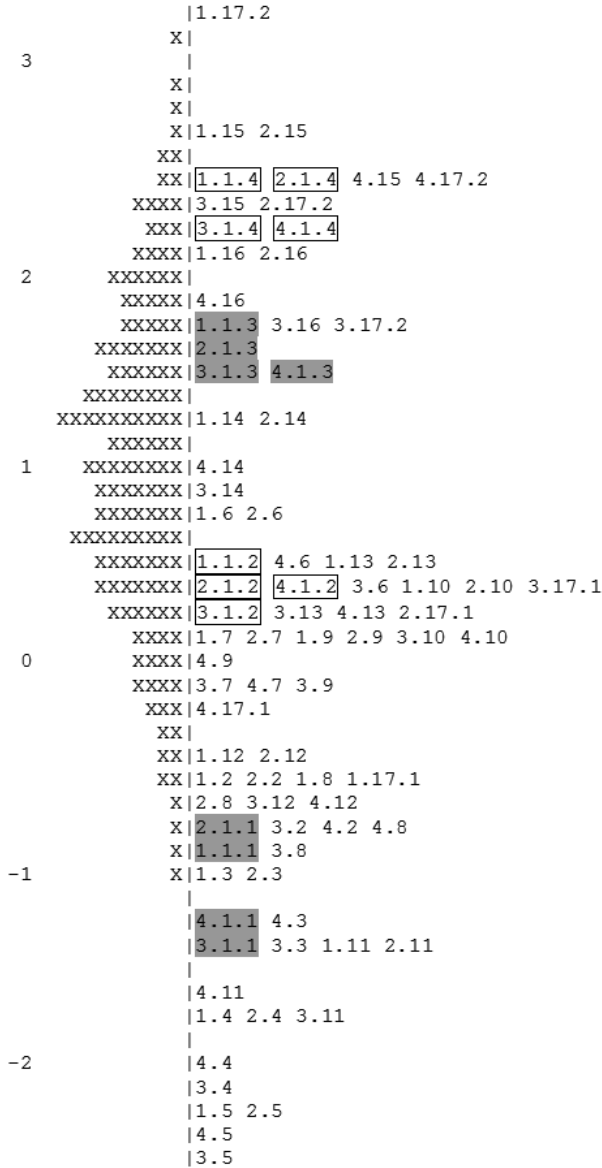


Tabelle 3: Genereller Zyklen-Effekt bei der Modellierung der Itemschwierigkeiten für die Skalierung der Hintergrundmerkmale für die Deutschen PIRLS-Daten von 2001, 2006, 2011 und 2016

Zyklus	Item-Schwierigkeit	S.E.	weighted MNSQ
1 - PIRLS 2001	0,125	0,006	1,02
2 - PIRLS 2006	0,119	0,006	0,97
3 - PIRLS 2011	-0,156	0,007	0,98
4 - PIRLS 2016	-0,088	0,011	0,95

In Abbildung 3 sind die zyklenspezifischen Itemschwierigkeiten der Büchervariable (Thurstonian thresholds) herausgehoben, die sich als nur wenig variierend erweisen. Lediglich die Parameter für die erste Schwelle (x.1.1) zeigen sich etwas stärker gespreizt als für die Schwellen 2, 3 und 4. Da ConQuest die zyklenspezifischen Itemschwierigkeiten (Item*Zyklus) aufgrund systematisch fehlender Werte für bestimmte Variablen (siehe Tabelle 1 – Items, die in verschiedenen Zyklen nicht erfragt wurden) nicht ausgibt, ist abschließend noch ein Modell vorgestellt, welches ausschließlich auf den Variablen basiert, die in allen vier Zyklen erhoben wurde.

Abbildung 3: Verortung der Schwellenparameter (stem-and-leaf plot) für den heimischen Buchbesitz in den Deutschen PIRLS-Daten von 2001, 2006, 2011 und 2016



Model III: Vollständiges DIF-Modell für in allen Zyklen erfragte Variablen: In allen vier PIRLS-Zyklen wurden in Deutschland acht Variablen einheitlich erfasst. Basierend auf diesen Variablen werden im Model III die generelle Itemschwierigkeit, die Unterschiede in den Itemschwierigkeiten in den vier Zyklen und die darüber hinausgehenden zyklenspezifischen Item-Variationen (DIF) berechnet (ConQuest Modell: $\text{item} + \text{ncycle} + \text{item} * \text{ncycle}$). Als Ergebnis der Berechnung bietet ConQuest eine Dekomposition der drei Itemparameter-Komponenten. Durch die Zusammenführung der drei Komponenten lassen sich die Itemschwierigkeitswerte je Item und Zyklus berechnen. Diese item- und zyklusspezifische Zusammenführung der von ConQuest ausgegebenen Parameter ist in Abbildung 4 visualisiert.

Abbildung 4: Zyklenspezifische Itemschwierigkeiten für die Merkmale, die in den Erhebungszyklen 2001, 2006, 2011 und 2016 erfragt wurden

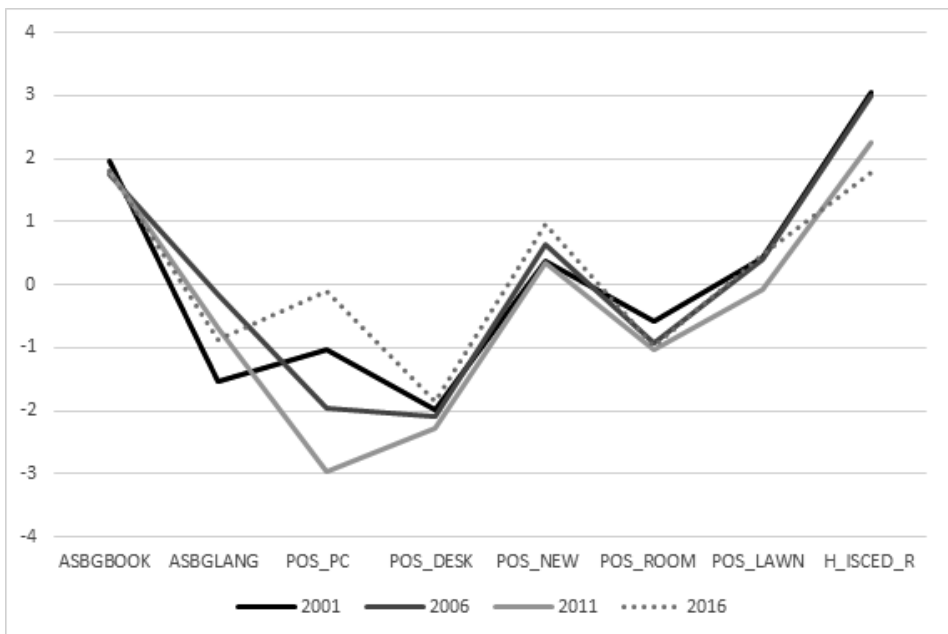


Abbildung 4 ist zu entnehmen, dass die zyklenspezifische Variation der Itemschwierigkeit bei der Büchervariable (ASBGBOOK) besonders gering ausfällt. Etwas größere Schwankungen zeigen sich bei der Information nach dem eigenen Schreibtisch (_DESK), der Verfügbarkeit einer Tageszeitung (_NEW), des eigenen Zimmers (_ROOM) und auch beim Rasenmäher (_LAWN). Etwas stärkere Unterschiede zeigen sich beim heimischen Sprachgebrauch (ASBGLANG) und dem höchsten Bildungsabschluss der Eltern (H_ISCED_R). Eine auffallend hohe Spreizung weist in dem Modell der Parameter für die Information über die Verfügbarkeit des Computers (_PC) in den verschiedenen Studienzyklen auf. Mit diesen Befunden kann gezeigt werden, dass die insbesondere in der Diskussion ste-

hende Büchervariable bei den vorliegenden Studienzyklen als stabiler Indikator und Ankeritem bei der Modellierung von kulturellem und ökonomischem Hintergrund genutzt werden kann. Allerdings verweist die starke Schwankung und damit die variierende Bedeutung des Computerbesitzes auf die Sensitivität des Konstrukts, wenn die inhaltliche Bedeutung der einzelnen Items über die Zeit variiert.⁷ Die erwarteten Wechselwirkungen zwischen Buchbesitz und Computer (inkl. Tablet, Smartphone und ggf. E-Book-Reader) verweisen auf die Notwendigkeit, die erwarteten Bedeutungsverschiebungen empirisch im Blick zu behalten.

Vergleich der Modelle: Zum Vergleich der Modelle, die auf der gleichen Personen- und Itemauswahl basieren, wurden Modellanpassungstests (goodness of fit) gerechnet. Tabelle 4 enthält die entsprechenden Werte für die drei Modelle.

Tabelle 4: Modellanpassung für die Bestimmung der zyklusunspezifischen und zyklusspezifischen Skalierungsläufe

	Modell I	Modell II	Modell III
AIC	326145,73	324821,99	166787,11
BIC	326323,14	325152,63	167053,23
CAIC	326345,14	325193,63	167086,23

Die durchweg niedrigeren Werte der drei Indices im Modell III (vgl. Tabelle 4) verweisen auf die beste Anpassung des sparsamsten Modells – während im direkten Vergleich der Modelle I und II, in denen alle verfügbaren Ressourcen als Items aufgenommen waren, die differenzierte Modellierung der Itemschwierigkeiten nach den Erhebungszyklen eine bessere Passung für das Modell II aufweist.

4. Diskussion

Bei der Bestimmung der Itemschwierigkeitsparameter für die als polytom kodierete Büchervariable zeigen sich zwar numerische (und aufgrund der sehr großen Stichproben auch signifikante) Unterschiede zwischen den Erhebungszyklen, die jedoch inhaltlich als wenig bedeutsam interpretiert werden können. Somit konnte gezeigt werden, dass die verschiedenen Abstufungen der Büchervariablen im Verlauf der letzten 15 Jahre praktisch keine Veränderung erfahren haben. Lediglich für die unterste Schwelle – am Übergang von 10 und weniger zu mehr als 10 Büchern – zeigt sich eine gewisse Spreizung über die Jahre. Hierbei ist jedoch zu beachten, dass die unterste Kategorie von unter 10 % der Schülerinnen und Schüler angegeben wird und entsprechend die für die verschiedenen Studienzyklen ermittelten Schwellen nicht nur eine Folge der Messeigenschaften des Items sind, sondern

⁷ Während sich 2001, 2006 und 2011 generell nach der Verfügbarkeit eines Computers erkundigt wurde, ist 2016 nach dem Besitz eines eigenen Computers gefragt worden.

auch von einer sich wandelnden Zusammensetzung der Schülerschaft abhängt (vgl. hierzu Hußmann, Stubbe et al., 2017).

Betrachtet man die Skalen insgesamt und damit auch die Komposition der berücksichtigten Indikatoren zum ökonomischen und kulturellen Status, erweisen sich diese als gut geeignet, um Unterschiede zwischen den Schülerinnen und Schülern in den jeweiligen Erhebungszyklen aber auch über diese Zyklen im zeitlichen Verlauf hinweg zu beschreiben. Der Modellvergleich zeigte eine besonders gute Passung des sparsamsten Modells an. Eine Entscheidung hierfür wäre aber eine empiristische Entscheidung, die außer Acht lässt, dass sich durch das Auslassen von Items auch die Interpretation des modellierten latenten Konstrukts ökonomischen und kulturellen Kapitals inhaltlich verändert und in diesem Fall auch einengt. Gerade wenn von einer sich über die Zeit verändernden Gesellschaft mit wandelnden Entscheidungsmöglichkeiten betreffend der individuellen Ausstattung mit Bildungsgütern ausgegangen wird, erscheint eine Fokussierung nur auf die gemeinsamen (Anker-)Items inhaltlich nicht angemessen. Vielmehr sollte die im Rahmen der Skalierung zulässige Variation von Items über die Zeit genutzt werden und gleichzeitig die mögliche zeitlich bedingte Interpretationsvalidität der Skalen in den Fokus rücken (Kane, 2013). Betreffend der Frage „Was wird aus den Büchern?“ kann zusammenfassend festgestellt werden, dass für den beobachteten Zeitraum von 2001 bis 2016 eine Betrachtung der Bildungsungleichheit auch mit der Information über den heimischen Buchbestand angemessen und gerechtfertigt erscheint. Somit kann die häufig geäußerte Frage, ob die Büchervariable zur Beschreibung des sozialen Hintergrunds sich schon gewandelt hätte, verneint werden. Allerdings zeigt die Berücksichtigung von Items, wie zum Beispiel die Frage nach dem heimischen Computer (dessen Bedeutsamkeit im zeitlichen Verlauf eine Veränderung erfahren hat), dass eine besondere Berücksichtigung bei der Modellierung angezeigt ist. Um diese Variable, die zu den jeweiligen Messzeitpunkten gute Messeigenschaften aufweist, nicht aus dem Modell entfernen zu müssen, wäre eine Alternative, diese Variable nicht übergreifend als Ankeritem zu verwenden, sondern jeweils als zyklusspezifisches Item (messzeitpunktspezifisch) zu berücksichtigen. So würde eine sich nachzeichenbare Veränderung des Items insbesondere als ökonomisches Merkmal im Modell Berücksichtigung finden. Gerade vor dem Hintergrund der im Beitrag skizzierten Veränderung in der Verfügbarkeit von digitalen Medien, die – wird die bisherige Entwicklung weiter in die Zukunft projiziert – zunehmend Printmedien substituieren könnten, ist abzusehen, dass auch mit der Information über den heimischen Buchbesitz zukünftig entsprechend umgegangen werden muss. Daher wird bei den Trendstudien PIRLS 2021 und TIMSS 2019 diese sich wandelnde Funktion der Büchervariablen zur Beschreibung des sozialen Hintergrunds zu überprüfen und in der Befunddarstellung zu reflektieren sein. Gleichzeitig ist zu eruieren, welche alternativen bildungsrelevanten Güter zukünftig in den Blick genommen werden müssen, um eine angemessene Darstellung von Unterschieden im sozialen Hintergrund realisieren zu können. Bis dato kann jedoch festgestellt werden, dass

die Information über den heimischen Buchbesitz als Ankeritem in der Betrachtung des sozialen Hintergrunds im Zeitraum von 2001 bis 2016 genutzt werden kann.

4.1 Ausblick

Sicherlich wird der heimische Buchbesitz weiterhin eine erklärungskräftige Variable von Bildungsdisparitäten sein –, aber die inhaltliche Interpretation des Zusammenhangs und damit der Validität wird sich nachhaltig durch die Verfügbarkeit und Nutzung digitaler Medien wandeln. Aus theoretischer Perspektive (vgl. Bourdieu, 1983) verrät die Büchervariable sowohl etwas über das kulturelle als auch das ökonomische Kapital von Familien. Empirisch konnten diese vielschichtigen Zusammenhänge nachgewiesen werden (vgl. u. a. Ehmke & Jude, 2010; Stubbe et al., 2016). Der heimische Buchbesitz stellt somit bisher eine wichtige und vielschichtige Ressource dar, die jedoch zur Steuerung bzw. in diesem Zusammenhang zur Kompensation von Bildungsbenachteiligungen – wenn überhaupt – zunächst nur in sehr geringem Umfang genutzt werden kann. So wäre bei dem Versuch kompensatorisch zu wirken selbstverständlich, dass die isolierte Steigerung der Bücher in einem Haushalt (durch Geschenke bzw. Subventionen) nicht mit einer tatsächlichen Steigerung des kulturellen bzw. ökonomischen Kapitals (im Sinne Bourdieus) und in der Folge somit auch nicht mit einer Steigerung der in groß angelegten Schulvergleichsuntersuchung im Fokus stehenden Kompetenzen einhergeht. Die wirksame Größe ist und bleibt der bildungsrelevante Umgang mit Büchern bzw. zukünftig auch mit digitalen Medien, die zumindest vergleichbare Inhalte bieten. Die Verfügbarkeit von digitalen Medien und die Kompetenz, diese nutzen zu können, stellen bereits heute einen zu berücksichtigenden und in absehbarer Zukunft (von 5 bis 10 Jahren) sicherlich zunehmend wichtigeren Faktor im Bildungskontext dar. Jedoch ist wie bisher beim heimischen Buchbestand auch bei digitalen Medien nicht allein der Besitz der Faktor, der die Bildungsunterschiede bei den Schülerinnen und Schülern erklären kann (Lorenz & Gerick, 2014). Erst bei der Betrachtung des Zusammenspiels von Verfügbarkeit und Nutzung gewinnen wir Erkenntnisse darüber, wie die selbständige Gestaltung und Entwicklung des eigenen Handelns in der (informierten Wissens-)Gesellschaft wirken.

4.2 Grenzen der vorliegenden Betrachtungen

Die in dem Beitrag vorgelegten Betrachtungen fokussieren auf die praktische Umsetzung von Messmodellen in einem Bildungssystem. In Erweiterung internationaler Analysen zur Messinvarianz von Hintergrundmerkmalen (Wendt, Kasper & Trendtel, 2017), bei dem die Einheitlichkeit der Modellierung im internationalen Vergleich im Fokus stand, sind – vor dem Hintergrund der hier vorgestellten Befunde – in einem nächsten Schritt internationale Vergleiche vorzunehmen, bei denen soziale, ökonomische und kulturelle Merkmale zwischen den betrach-

teten Staaten variieren und somit auch als weitere Varianzkomponenten zu berücksichtigen sind. Bei einem solchen globalen Vergleich wäre zu berücksichtigen, dass die – auch in diesem Beitrag – herangezogenen empirisch vielfach replizierten Erklärungsmodelle ihren Ursprung in einer insbesondere nord-westlich geprägten Tradition haben. Es wäre somit zu prüfen, ob und wie entsprechenden, Modellierungen für andere Regionen der Welt anzupassen wären.

Literatur

- Bos, W., Bonsen, M., Baumert, J., Prenzel, M., Selter, C. & Walther, G. (Hrsg.) (2008). *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.) (2007). *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G. & Valtin, R. (Hrsg.) (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (Hrsg.) (2012). *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Wendt, H., Köller, O. & Selter, C. (Hrsg.) (2012). *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bourdieu, P. (1977). Cultural reproduction and social reproduction. In J. Karabel & A. H. Halsey (Hrsg.), *Power and ideology in education* (S. 487–511). New York, NY: Oxford University Press.
- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (Soziale Welt, Sonderbd. 2, S. 183–198). Göttingen: Schwartz.
- Coleman, J. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 95–120.
- Ebbs, D. & Wry, E. (2017). Translation and layout verification for PIRLS 2016. In M. O. Martin, I. V. S. Mullis & M. Hooper (Hrsg.), *Methods and procedures in PIRLS 2016*. Zugriff am 04.06.2018 unter https://pirls.bc.edu/publications/pirls/2016-methods/P16_MP_Chap7_Translation_Layout_Verification.pdf
- Ehmke, T. & Jude, N. (2010). Soziale Herkunft und Kompetenzerwerb. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 231–253). Münster: Waxmann.
- Elley, W. B. (1994). *The IEA study of reading literacy: Achievement and instruction in thirty-two school systems*. Exeter: Pergamon.
- Foshay, A. W. (1962). The background and the procedures of the twelve-country study. In A. W. Foshay, R. L. Thorndike, F. Hotyat, D. A. Pidgeon & D. A. Walker (Hrsg.), *Educational achievements of thirteen-year-olds in twelve countries. Results of an international research project, 1959–61* (S. 7–19). Hamburg: Unesco Institute for Education.
- Fraillon, J., Schulz, W. & Ainley, J. (2013). *International Computer and Information Literacy Study. Assessment framework*. Zugriff am 04.06.2018 unter https://research.acer.edu.au/ict_literacy/9/

- Goy, M., van Ackeren, I. & Schwippert, K. (2008). Ein halbes Jahrhundert internationale Schulleistungsstudien. Eine systematische Übersicht. *Tertium Comparationis*, 14(1), 77–107.
- Grønmo, L. S., Lindquist, M., Arora, A. & Mullis, I. V. S. (2013). TIMSS 2015 mathematics framework. In I. V. S. Mullis & M. O. Martin (Hrsg.), *TIMSS 2015 assessment frameworks* (S. 11–27). Zugriff am 03.02.2017 unter http://timssandpirls.bc.edu/timss2015/downloads/T15_Frameworks_Full_Book.pdf
- Gustafsson, J.-E. & Rosén, M. (2014). Quality and credibility of international studies. In R. Strietholt, W. Bos, J.-E. Gustafsson & M. Rosén (Hrsg.), *Educational policy evaluation through international comparative assessments* (S. 19–32). Münster: Waxmann.
- He, J. & van de Vijver, F. J. R. (2013a). A general response style factor: Evidence from a multi-ethnic study in the Netherlands. *Personality and Individual Differences*, 55, 794–800.
- He, J. & van de Vijver, F. J. R. (2013b). Methodological issues in cross-cultural studies in educational psychology. In G. A. D. Liem & A. B. I. Bernado (Hrsg.), *Advancing cross-cultural perspectives on educational psychology* (S. 39–55). Charlotte, NC: Information Age Publishing Inc.
- Husén, T. (Hrsg.) (1967). *International study of achievement in mathematics. A comparison of twelve countries/international project for the evaluation of educational achievement (IEA)*. Stockholm: Almqvist & Wiksell.
- Husén, T. & Postlethwaite, T. N. (1996). A brief history of the International Association for the Evaluation of Educational Achievement (IEA). *Assessment in Education*, 3(2), 129–141.
- Hußmann, A., Stubbe, T. C. & Kasper, D. (2017). Soziale Herkunft und Lesekompetenzen von Schülerinnen und Schülern. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 195–217). Münster: Waxmann.
- Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E.-M., McElvany, N., Stubbe, T. C. & Valtin, R. (2017). *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Hußmann, A., Wendt, H., Kasper, D., Bos, W. & Goy, M. (2017). Ziele, Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU 2016). In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 29–77). Münster: Waxmann.
- Jones, L. R., Wheeler, G. & Centurino, V. A. S. (2013). TIMSS 2015 science framework. In I. V. S. Mullis & M. O. Martin (Hrsg.), *TIMSS 2015 assessment frameworks* (S. 29–58). Zugriff am 03.02.2017 unter http://timssandpirls.bc.edu/timss2015/downloads/T15_Frameworks_Full_Book.pdf
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 55, 1–73. Zugriff am 03.02.2017 unter onlinelibrary.wiley.com/doi/10.1111/jedm.11200/epdf
- Klafki, W. (2007). *Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik*. Weinheim: Beltz.
- Klieme, E. & Hartig, J. (2007). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. *Zeitschrift für Erziehungswissenschaft, Sonderheft 8*, 11–29.
- LaRoche, S., Joncas, M. & Foy, P. (2017). Sample design in PIRLS 2016. In M. O. Martin, I. V. S. Mullis & M. Hooper (Hrsg.), *Methods and procedures in PIRLS*

2016. Zugriff am 04.06.2018 unter https://pirls.bc.edu/publications/pirls/2016-methods/P16_Methods_and_Procedures.pdf
- Lehmann, R. H., Peek, R., Pieper, I. & Stritzky, R. v. (1991). Lesefähigkeit und Lesegewohnheiten von Achtkläßlern. *Deutschunterricht*, 44(6), 410–417.
- Lorenz, R. & Gerick, J. (2014). Neue Technologien und die Leseleistung von Grundschulkindern. In B. Eickelmann, R. Lorenz, M. Vennemann, J. Gerick & W. Bos (Hrsg.), *Grundschule in der digitalen Gesellschaft. Befunde aus den Schulleistungstudien IGLU und TIMSS 2011* (S. 59–71). Münster: Waxmann.
- mpfs – Medienpädagogischer Forschungsverbund Südwest (n.d.). *Basisuntersuchungen zur Mediennutzung von Kindern und Jugendlichen. JIM-Studie*. Zugriff am 04.06.2018 unter <https://www.mpfs.de/studien/?tab=tab-18-1>
- Mullis, I. V. S., Martin, M. O. & Hooper, M. (2016). Measuring changing educational contexts in a changing world: Evolution of the TIMSS and PIRLS questionnaires. In M. Rosén, K. Y. Hansen & U. Wolff (Hrsg.), *Cognitive abilities and educational outcomes* (S. 207–222). Cham: Springer International Publishing.
- Mullis, I. V. S., Martin, M. O. & Sainsbury, M. (2015). PIRLS 2016 reading framework. In I. V. S. Mullis & M. O. Martin (Hrsg.), *PIRLS 2016 assessment framework* (S. 11–29). Zugriff am 03.02.2017 unter http://timss.bc.edu/pirls2016/downloads/P16_FW_Chap1.pdf
- Rost, J. (2004). *Testtheorie. Testkonstruktion*. Bern: Huber.
- Schulte, K., Hartig, J. & Pietsch, M. (2014). Der Sozialindex für Hamburger Schulen. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine daten- und theoriegestützte Schulentwicklung. Konzeption und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)* (HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Bd. 13; S. 67–80). Münster: Waxmann.
- Schwippert, K. (2011). Von der Beschreibung zur Interpretation. Konzepte der Betrachtung von Chancengleichheit. In C. Tarnai (Hrsg.), *Sozialwissenschaftliche Forschung in Diskurs und Empirie* (S. 35–48). Münster: Waxmann.
- Schwippert, K. (2015). Wie viel Bildung steckt in groß angelegten Schulvergleichsuntersuchungen? In H.-J. Fischer, H. Giest & K. Michalik (Hrsg.), *Bildung im und durch Sachunterricht* (Probleme und Perspektiven des Sachunterrichts, Bd. 25; S. 39–50). Bad Heilbrunn: Klinkhardt.
- Schwippert, K., Wendt, H. & Tarelli, I. (2012). Lesekompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 191–207). Münster: Waxmann.
- Stubbe, T. C. & Goy, M. (2013). Besitzen wollen und sich leisten können. Analysen zur dimensional getrennten Erfassbarkeit von ökonomischem und kulturellem Kapital in Familien. In K. Schwippert, M. Bensen & N. Berkemeyer (Hrsg.), *Schul- und Bildungsforschung. Diskussionen, Befunde und Perspektiven. Festschrift für Wilfried Bos* (S. 203–222). Münster: Waxmann.
- Stubbe, T. C., Schwippert, K. & Wendt, H. (2016). Soziale Disparitäten der Schülerleistungen in Mathematik und Naturwissenschaften. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland und im internationalen Vergleich* (S. 229–316). Münster: Waxmann.
- Tenorth, H.-E. (2005). Grundbildung – Institutionelle Restriktion oder legitimes Programm? In M. Götz & K. Müller (Hrsg.), *Grundschule zwischen den Ansprüchen der Individualisierung und Standardisierung* (S. 17–30). Wiesbaden: Springer.
- Weinert, F. E. (1999). *Concepts of competence. Definition and selection of competencies*. Zugriff am 04.06.2018 unter <https://pdfs.semanticscholar.org/8b88/efa9d5e0a4b605aea6e5e3b9ec640bebo89.pdf>

- Wendt, H., Bos, W., Selter, C., Köller, O., Schwippert, K. & Kasper, D. (Hrsg.) (2016). *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Wendt, H., Kasper, D. & Trendtel, M. (2017). Assuming measurement invariance of background indicators in international comparative educational achievement studies: A challenge for the interpretation of achievement differences. *Large-scale Assessments in Education*, 5(10).
- Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, S. A. (2007). *ACER ConQuest version 2.0. Generalised item response modelling software*. Camberwell: Australian Council for Educational Research.

Birgit Eickelmann, Julia Gerick & Mario Vennemann

Unerwartet erfolgreiche Schulen im digitalen Zeitalter – Eine Analyse von Schulmerkmalen resilienter Schultypen auf Grundlage der IEA-Studie ICILS 2013

Zusammenfassung

Die fortschreitende Digitalisierung stellt Schulen bei der Vermittlung von computer- und informationsbezogenen Kompetenzen vor neue Herausforderungen. In diesem Zusammenhang konnte die International Computer and Information Literacy Study (ICILS) 2013 zeigen, dass die Achtklässlerinnen und Achtklässler in Deutschland im internationalen Vergleich lediglich über durchschnittliche computer- und informationsbezogene Kompetenzen verfügen. Zudem konnten substantielle soziale Disparitäten hinsichtlich eines kompetenten Umgangs mit neuen Technologien festgestellt werden. Ausgehend von der Beobachtung, dass sich einerseits zahlreiche Schulen in besonders herausfordernden Lagen im besonderen Maße für die Unterstützung des Erwerbs digitaler Kompetenzen ihrer Schülerinnen und Schüler engagieren und sich andererseits für andere Domänen wie Lesen und Mathematik auf der Basis empirischer Studien Schulen identifizieren lassen, die trotz herausfordernder sozialer Schülerkomposition ein hohes Kompetenzniveau erreichen, betrachtet der Beitrag unerwartet erfolgreiche Schulen im Bereich der computer- und informationsbezogenen Kompetenzen. Im Fokus steht daher, inwiefern sich diese auch als organisational resilient bezeichneten Schulen aufgrund von Schulmerkmalen zu einer Typologie zusammenfassen lassen und wie sich relevante schulische Merkmale auch unter Berücksichtigung des Schulleitungshandeln innerhalb der vier identifizierbaren Schulgruppen beschreiben lassen.

Prof. Dr. Birgit Eickelmann (corresponding author) · Dr. Mario Vennemann, Institut für Erziehungswissenschaft, Fakultät für Kulturwissenschaften, Universität Paderborn, Warburger Straße 100, 33098 Paderborn, Deutschland
E-Mail: birgit.eickelmann@uni-paderborn.de
mario.vennemann@upb.de

Prof. Dr. Julia Gerick, Fachbereich Schulpädagogik, Sozialpädagogik, Behindertenpädagogik und Psychologie in Erziehung und Unterricht (EW 2), Fakultät für Erziehungswissenschaft, Universität Hamburg, Von-Melle-Park 8, 20146 Hamburg, Deutschland
E-Mail: julia.gerick@uni-hamburg.de

Schlagworte

ICILS 2013; Schulische Resilienz; Computer- und informationsbezogene Kompetenzen; Schuleffektivitätsforschung; Schulentwicklung im digitalen Zeitalter

Unexpectedly successful schools in a digital age – An analysis of school characteristics of resilient school types based on the IEA study ICILS 2013

Abstract

The ongoing digitalization poses new challenges to schools concerning computer and information literacy. In this context, the International Computer and Information Literacy Study (ICILS) 2013 has shown that grade 8 students in Germany are scoring near the international average in computer and information literacy. In addition, substantial social disparities concerning computer and information literacy of grade 8 students have been observed. Proceeding from the observation, that on the one hand many schools especially located in challenging areas are particularly engaged in supporting student's digital literacy and innovatively designing learning with digital media and on the other hand for other domains such as reading literacy and mathematics on the basis of empirical studies it has been identified that some schools in spite of challenging social student bodies have higher average achievement in literacy, this article addresses schools which were unexpectedly successful in the competence area Computer and Information Literacy and aims to investigate how these organizational resilient schools can be classified as a typology based on similar school characteristics and how relevant school characteristics in consideration of school leadership within the four identified different school groups are.

Keywords

ICILS 2013; Educational resilience; Computer and information literacy; School effectiveness; School development in the digital age

1. Einleitung: Die Relevanz des schulischen Kontextes für die erfolgreiche Implementierung digitaler Medien in Schulen

Der digitale Wandel stellt Schulen in Deutschland und weltweit derzeit vor große Herausforderungen (Eickelmann, 2018; Voogt, Knezek, Christensen & Lai, 2018). Neben neuen Anforderungen, die sich auf die Gestaltung von Lehr- und Lernprozessen hinsichtlich einer lernförderlichen Nutzung der Potenziale digitaler Medien in Schulen beziehen (vgl. u.a. Scheiter, 2017) gewinnt vor allem die Förderung computer- und informationsbezogener Kompetenzen (computer and information literacy, CIL) an Relevanz, damit Schülerinnen und Schüler künfti-

gen beruflichen, privaten und gesellschaftlichen Anforderungen gewachsen sind (Fraillon, Ainley, Schulz, Friedman & Gebhardt, 2014; KMK, 2016). Die IEA-Studie International Computer and Information Literacy Study (ICILS) 2013 konnte in diesem Zusammenhang zeigen, dass die Achtklässlerinnen und Achtklässler in Deutschland im internationalen Vergleich lediglich über ein durchschnittliches Niveau computer- und informationsbezogener Kompetenzen verfügen und fast ein Drittel der Jugendlichen nur die untersten beiden von fünf Kompetenzstufen erreicht (Bos, Eickelmann & Gerick, 2014).

Weiterhin konnte in ICILS 2013 erstmalig über eine computerbasierte Erhebung der computer- und informationsbezogenen Kompetenzen der Schülerinnen und Schüler empirisch gezeigt werden, dass auch für diese fächerübergreifende Querschnittskompetenz (u. a. Eickelmann, Bos, Gerick & Kahnert, 2014) substantielle soziale Disparitäten sichtbar werden. So erzielen Schülerinnen und Schüler aus weniger privilegierten Familien – je nach Indikatorisierung ihrer sozialen Lage – computer- und informationsbezogene Kompetenzen, die im Mittel um etwa 50 Punkte unter dem Niveau der Kompetenzen von Schülerinnen und Schülern aus privilegierten sozialen Lagen anzusiedeln sind (Wendt, Vennemann, Schwippert & Drossel, 2014).

Zu berücksichtigen ist jedoch, dass im Kontext der festgestellten Disparitäten durch die Fokussierung auf Durchschnittswerte weder national noch international berücksichtigt wurde, dass es in den an ICILS 2013 teilnehmenden Bildungssystemen auch Schulen gibt, die sich unabhängig von der sozialen Komposition der Schülerschaft bereits vor Jahren auf den Weg gemacht haben, den kompetenten und reflektierten Umgang aller Schülerinnen und Schüler mit digitalen Medien zu fördern. Im Zuge des Lernens und Lehrens mit digitalen Medien kann – wie auch für andere fächerspezifische Domänen (vgl. Bensen, Bos & Rolff, 2008; Fend, 2008; Klieme, Baumert & Schwippert, 2000) – angenommen werden, dass der Einzelschule als pädagogische Handlungseinheit im Bereich der Förderung von computer- und informationsbezogenen Kompetenzen eine besondere Bedeutung zukommt. Dies konnte bereits seit Ende der 1990er Jahre durch zahlreiche nationale und internationale Studien gezeigt werden, die die Relevanz der Einzelschule für die gelingende Implementierung digitaler Medien in Schulentwicklungsprozesse und in schulische Lehr- und Lernprozesse untersuchen und relevante schulische Merkmale identifizieren konnten (Eickelmann, 2011; Kozma, 2003; Petko, Egger & Wespi, 2015; Schaumburg, Prasse, Tschackert & Blömeke, 2007; Schuhmacher, 2004; Schulz-Zander, 2001, 2004), aber nicht die computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern erfasst haben. Hier schließt die ICILS-2013-Studie eine Forschungslücke und bildet daher die Grundlage für den vorliegenden Beitrag. Aus Sicht der Schulentwicklungsforschung und vor dem Hintergrund der skizzierten ausgewählten Forschungsbefunde zu sozialen Disparitäten erscheinen im Zusammenhang mit der Vermittlung von computer- und informationsbezogenen Kompetenzen dabei besonders diejenigen Schulen interessant, die trotz einer herausfordernden Schülerkomposition überdurchschnittlich hohe mittlere computer- und informationsbezogene Kompetenzstände

hervorbringen, und damit für den Bereich des Lehrens und Lernens mit digitalen Medien als „organisational resilient“ (Racherbäumer, Funke, Van Ackeren & Clausen, 2013, S. 250) bezeichnet werden können. Eine Schule wäre demnach für den Bereich der computer- und informationsbezogenen Kompetenzen als organisational resilient zu bezeichnen, wenn es ihr trotz eines vergleichsweise hohen Anteils von Schülerinnen und Schülern aus sozioökonomisch wenig privilegierten Familien gelingt, in dem betrachteten Kompetenzbereich überdurchschnittlich hohe mittlere Kompetenzen hervorzubringen. Solche Schulen sind demnach im Zuge der Prozesse der Digitalisierung unerwartet erfolgreich, da sich für Deutschland insgesamt die oben beschriebenen sozialbedingten Bildungsdisparitäten zeigen.

Anknüpfend an dieses in der Schul- und insbesondere in der Schulentwicklungsforschung thematisierte Verständnis der besonderen Relevanz der Schulebene weist die empirische Forschung bereits für andere Kompetenzbereiche auf das Phänomen der organisationalen Resilienz hin und zeigt, dass erwartungswidrig erfolgreiche Schulen über bestimmte schulische Merkmale sowie über ein ausgeprägtes Schulleitungshandeln (Leadership) verfügen (Muijs, Harris, Chapman, Stoll & Russ, 2004). Für den Bereich der computer- und informationsbezogenen Kompetenzen fehlen bisher entsprechende Befunde. Diese Forschungslücke bearbeitet der vorliegende Beitrag.

In der Zusammenschau bisheriger Ergebnisse liegen damit insgesamt, insbesondere auch für Deutschland, bisher vergleichsweise wenige Forschungsergebnisse vor, die relevante gemeinsame Merkmale von erfolgreichen Schulen im Hinblick auf die Vermittlung von computer- und informationsbezogenen Kompetenzen herausarbeiten. Aus methodischer Sicht ist diesbezüglich besonders hervorzuheben, dass die vorliegenden Befunde vielfach auf Mehrebenenmodellierungen zurückgreifen und dabei die Gemeinsamkeiten von Schulen im Sinne einer Typologie nicht berücksichtigen (vgl. u. a. Eickelmann, Gerick & Koop, 2017; Gerick, Eickelmann & Bos, 2017). Daher wird im Rahmen dieses Beitrags in methodischer Hinsicht ein strukturentdeckender Ansatz gewählt.

Mit dem vorliegenden Beitrag wird vor dem Hintergrund des identifizierten Forschungsdesiderates und seiner Relevanz für die Schulforschung und die schulische Praxis im internationalen Vergleich zunächst die Forschungsfrage bearbeitet, ob das Phänomen der organisationalen Resilienz von Schulen auch im Bereich der computer- und informationsbezogenen Kompetenzen zu identifizieren ist. Weiterhin wird untersucht, ob sich die identifizierten Schulen anhand ihrer Merkmale zu gemeinsamen Profilen gruppieren lassen und ob sich insbesondere auch das Schulleitungshandeln – wie von Muijs et al. (2004) dargestellt – an diesen Schulen systematisch unterscheidet. In die Sekundäranalysen der ICILS-2013-Daten gehen neben Deutschland alle Bildungssysteme ein, in denen die Achtklässlerinnen und Achtklässler im Rahmen der Studie ein signifikant höheres Kompetenzniveau erreicht haben als die Achtklässlerinnen und Achtklässler in Deutschland (Bos et al., 2014; Fraillon et al., 2014; siehe auch Abschnitt 4). Damit soll, im Verständnis der internationalen Schulforschung, die Möglichkeit genutzt werden, durch internationale Vergleiche von anderen Bildungssystemen zu lernen

und die Situation des eigenen Bildungssystems im betrachteten Untersuchungsbereich einzuordnen.

Ausgehend von diesem Forschungsinteresse wird im Folgenden zunächst auf die für den Beitrag relevante theoretische Verortung der Studie ICILS 2013 abgehoben (Abschnitt 2). Im Anschluss daran wird der Forschungsstand zu denjenigen schulischen Merkmalen dargestellt, die in diesem Beitrag theoriegeleitet genutzt werden, um die organisational resilienten Schulen in ICILS 2013 zu typisieren (Abschnitt 3). Dazu zählen Merkmale, die in anderen Studien als bedeutsam für die Schulentwicklung mit digitalen Medien herausgearbeitet wurden und die neben der Verfügbarkeit und Nutzungshäufigkeit von digitalen Medien beispielsweise auch die Kompetenzeinschätzungen der Lehrpersonen und die Einstellung des Kollegiums gegenüber den Potenzialen digitaler Medien für das Lehren und Lernen an der Schule umfassen. Zur Gewährleistung der Nachvollziehbarkeit der hier durchgeführten Analysen wird anknüpfend an den Forschungsstand das methodische Vorgehen erläutert (Abschnitt 4). Nach der Darstellung der Analyseergebnisse (Abschnitt 5) erfolgt eine inhaltliche und methodische Diskussion der Befunde (Abschnitt 6).

2. Theoretische Verortung: Das Prozessmodell schulischer Wirksamkeit im Rahmen von ICILS 2013

Das in ICILS 2013 unter dem Begriff der computer- und informationsbezogenen Kompetenzen zusammengeführte Konstrukt wird in Anlehnung an die verschiedenen sinnverwandten und oftmals synonym verwendeten Begrifflichkeiten wie digital literacy oder ICT-Literacy beschrieben (Eickelmann, Bos, et al., 2014). Im Sinne eines Literacy-Ansatzes definieren Eickelmann, Bos et al. (2014, S. 45) computer- und informationsbezogenen Kompetenzen gemäß Fraillon et al. (2014) im Rahmen von ICILS 2013 als:

individuelle Fähigkeiten einer Person [...], die es ihr erlauben, Computer und neue Technologien zum Recherchieren, Gestalten und Kommunizieren von Informationen zu nutzen und diese zu bewerten, um am Leben im häuslichen Umfeld, in der Schule, am Arbeitsplatz und in der Gesellschaft erfolgreich teilzuhaben.

Die computer- und informationsbezogenen Kompetenzen bestehen in diesem Ansatz aus zwei inhaltlichen Teilbereichen, die sich jeweils weiter in verschiedene Aspekte differenzieren lassen (vgl. auch Senkbeil et al., 2014). Teilbereich I *Informationen sammeln und organisieren* umfasst drei Aspekte, unter anderem ‚Auf Informationen zugreifen und Informationen bewerten‘. Teilbereich II *Informationen erzeugen und austauschen* deckt vier Aspekte ab, die sich zum

Beispiel darauf beziehen, *Informationen kommunizieren und austauschen* sowie *Informationen sicher nutzen* zu können.

Das ICILS 2013 zugrundeliegende theoretische Rahmenmodell zum Erwerb von computer- und informationsbezogenen Kompetenzen greift auf ein Prozessmodell schulischen Lernens zurück (vgl. Fraillon, Schulz & Ainley, 2013). Es kategorisiert relevante Faktoren, die der Mehrebenenstruktur des Erwerbs computer- und informationsbezogener Kompetenzen Rechnung tragen (Gerick, 2018) und baut in systematischer Art und Weise auf Erkenntnissen zu und Annahmen aus früheren Schulwirksamkeitsmodellen auf, die darauf abzielen, das Zusammenspiel von bildungsbezogenen Eingangs- (Input) und Prozessvariablen sowie den Outcomes zu konzeptualisieren (Bos et al., 2003; Ditton, 2000; Scheerens & Bosker, 1997; Vennemann, 2018).

Das Rahmenmodell von ICILS unterscheidet Voraussetzungen (antecedents) und Prozesse (processes) und differenziert dabei vier verschiedene Ebenen: (1) die Ebene des gesellschaftlichen Kontextes, (2) die Schul- und Klassenebene, (3) die Ebene des häuslichen Umfeldes der Schülerinnen und Schüler sowie (4) die Ebene der Schülerinnen und Schüler.

Um die computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern (Outcome) zu erklären, werden in diesem Ansatz die Voraussetzungen und die Prozesse im Zusammenhang betrachtet. In diesem Modell werden die Voraussetzungen als relevante Faktoren betrachtet, die aber keinen direkten Einfluss auf die Lernergebnisse von Schülerinnen und Schülern haben. Im Kontrast dazu wird angenommen, dass die Prozessfaktoren in einem direkten Zusammenhang mit den computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern stehen. Im Folgenden werden die vier Ebenen in gegebener Kürze dargestellt (vgl. dazu auch im Detail Eickelmann, Bos, et al., 2014).

Die gesellschaftliche Kontextebene (wider community) bezieht sich auf die Voraussetzungen für den Erwerb computer- und informationsbezogener Kompetenzen u.a. auf Aspekte der Struktur des Bildungssystems. Hinsichtlich der Prozesse stellen vor allem technologiebezogene Strategien und Curricula sowie Ziele auf Ebene des Bildungssystems relevante Faktoren dar.

Die *Schul- und Klassenebene* bezieht sich zum einen auf schulbezogene oder klassenbezogene Voraussetzungen, wie beispielsweise die Ausstattung mit digitalen Medien. Zum anderen lassen sich auf dieser Ebene im Hinblick auf Prozesse zum Beispiel das Handeln der Schulleitung oder die Nutzung digitaler Medien durch Lehrpersonen verorten.

Auf *Ebene des häuslichen Umfeldes* wird primär der familiäre Hintergrund einbezogen. Sie bezieht sich auf Voraussetzungen wie den sozioökonomischen Status (SES) der Schülerfamilien oder den Zuwanderungshintergrund. Auf der Prozessebene werden die häusliche Nutzung von und der Austausch mit Familienmitgliedern über digitale Medien als relevant für den Erwerb von computer- und informationsbezogenen Kompetenzen erachtet (vgl. auch Eickelmann, Bos, et al., 2014).

Auf *Ebene der Schülerinnen und Schüler* lassen sich die für den Kompetenzerwerb als relevant erachteten individuellen Merkmale wie Alter, Geschlecht und Bildungsaspiration als Voraussetzungen verorten. Im Sinne der unterschiedlichen Prozessmerkmale wird auf dieser Ebene angenommen, dass sich Einstellungs- (z. B. computerbezogene Selbstwirksamkeitserwartung) und Verhaltensaspekte (z. B. Nutzung von digitalen Medien) direkt auf die computer- und informationsbezogenen Kompetenzen auswirken (Eickelmann, Bos, et al., 2014).

Über die im internationalen Rahmenkonzept berücksichtigten Prädiktoren hinaus konnten insbesondere für Deutschland weitere Faktoren als relevant identifiziert werden. Hierzu gehören u. a. Aspekte der Lehrerverkooperation (vgl. Drossel & Eickelmann, 2017; Drossel, Eickelmann, van Ophuysen & Bos, 2018) und das Schulleitungshandeln (Gerick, 2018; Gerick, Eickelmann, Drossel & Lorenz, 2016).

An dieses Rahmenmodell anknüpfend stehen in diesem Beitrag die Faktoren auf der Schulebene im Vordergrund. Weiterhin wird durch die Berücksichtigung der sozialen Herkunft der Schülerinnen und Schüler implizit ebenfalls auf die Ebene des häuslichen Umfeldes sowie durch den anvisierten internationalen Vergleich auch auf die gesellschaftliche Kontextebene abgehoben.

3. Forschungsstand und forschungsleitende Fragestellungen

Vor dem Hintergrund noch fehlender Befunde zu organisational resilienten Schulen für den überfachlichen Kompetenzbereich der computer- und informationsbezogenen Kompetenzen, werden im folgenden Abschnitt aktuelle Forschungsergebnisse zu solchen Faktoren auf der Schulebene zusammengetragen, die sich im Hinblick auf die erfolgreiche Implementierung digitaler Medien in Schulen sowie zur Förderung computer- und informationsbezogener Kompetenzen als relevant herausgestellt haben (u. a. Eickelmann, 2011; Gerick et al., 2017; Lorenz, Eickelmann & Gerick, 2015). Die vorliegenden Befunde werden nach schulischen Voraussetzungen (vgl. Abschnitt 3.1) und schulischen Prozessmerkmalen (vgl. Abschnitt 3.2) differenziert dargestellt.

3.1 Schulische Bedingungsfaktoren auf Ebene der Voraussetzungen

Auf der Ebene der Voraussetzungen erweist sich vor dem Hintergrund des nationalen und internationalen Forschungsstandes (a) die Verfügbarkeit von schulischen IT-Ressourcen, (b) die Kompetenzeinschätzung der Lehrpersonen hinsichtlich des Unterrichtens mit digitalen Medien sowie (c) die Einstellungen der Lehrpersonen bzw. des Lehrerkollegiums als Ganzes zum Lernen mit digitalen Medien als besonders relevant.

3.1.1 Verfügbarkeit schulischer IT-Ressourcen

Hinsichtlich der Verfügbarkeit von Informations- und Kommunikationstechnologien (IKT) zeigen sowohl nationale als auch internationale Untersuchungen, dass deren Vorhandensein und eine für unterrichtliche Zwecke passende Ausstattung dieser Medien – nachvollziehbarerweise – eine notwendige Vorbedingung für die Nutzung digitaler Medien darstellt (Eickelmann et al., 2017; European Commission, 2013). Dabei besteht weitestgehend Konsens darüber, dass die Ausstattung von Schulen mit digitalen Medien per se nicht vermag, das Lernen und Lehren mit digitalen Medien zu verbessern (Lai, 2008; Law, Pelgrum & Plomp, 2008)

3.1.2 Die auf digitale Medien bezogenen Kompetenzeinschätzungen der Lehrpersonen

Ein weiteres relevantes Merkmal auf Schulebene stellt die Kompetenzeinschätzung der Lehrpersonen dar. Nationale und internationale Studien weisen in diesem Zusammenhang wiederholt auf die Bedeutung der selbsteingeschätzten Kompetenzen im Umgang mit digitalen Medien hin (vgl. z.B. Eickelmann, 2011; Hatlevik, 2016; Saudelli & Ciampa, 2016)

Für Deutschland konnte im Rahmen von ICILS 2013 auf Grundlage einer repräsentativen Stichprobe diesbezüglich zudem gezeigt werden, dass Lehrpersonen, die in der achten Jahrgangsstufe unterrichten, zwar mehrheitlich berichten, dass sie im Stande sind, Unterricht vorzubereiten, der den Einsatz von digitalen Medien beinhaltet, diese Einschätzung im internationalen Vergleich jedoch eher gering ausfällt (Gerick, Schaumburg, Kahnert & Eickelmann, 2014). Dies ist für den vorliegenden Beitrag insofern relevant, als dass Endberg und Lorenz (2017) im Rahmen des Länderindikators 2016 zeigen konnten, dass die eigene Einschätzung des Könnens und des Wissens zum Einsatz digitaler Medien in bestimmten Lehr- und Lernsituationen direkt mit der Häufigkeit der Computernutzung im Unterricht in Verbindung steht (Bos et al., 2016).

3.1.3 Einstellungen des Kollegiums zu Potenzialen digitaler Medien für das Lehren und Lernen

Dass die Einstellungen von Lehrpersonen eine förderliche bzw. auch hinderliche Wirkung auf die Nutzung von digitalen Medien im Unterricht haben, ist ebenfalls durch internationale empirische Forschung breit belegt (für einen Überblick vgl. Buabend-Andoh, 2012, siehe auch Petko, 2012; Schweizer & Horn, 2014; Tomczyk et al., 2015; Yang & Leung, 2015).

Auch für die Situation in Deutschland liegen Erkenntnisse zur Relevanz der Einstellungen der Lehrkräfte vor. Im Rahmen von ICILS 2013 konnte gezeigt wer-

den, dass die Lehrkräfte, die in der achten Jahrgangsstufe unterrichten, „mehrerheitlich eine positive Sichtweise auf den Einsatz neuer Technologien im Unterricht haben“, in allen anderen an der Studie teilnehmenden Bildungssystemen jedoch größere Anteile der Lehrpersonen den positiven Sichtweisen auf den Einsatz von digitalen Medien in der Schule zustimmen (Gerick et al., 2014, S. 176). Weiterhin stellen Eickelmann und Vennemann (2017) für eine Auswahl europäischer Länder dar, dass die Einstellungen der Lehrpersonen – direkt oder indirekt – mit der Nutzung digitaler Medien in Verbindung stehen. Ähnliche Ergebnisse legen Endberg, Lorenz und Senkbeil (2015) im Kontext des Länderindikators 2015 (Bos et al., 2015) vor: Hier zeigt sich, dass Lehrpersonen, die positiv gegenüber der Nutzung digitaler Medien eingestellt sind, auch unter Kontrolle von individuellen Hintergrundmerkmalen im Durchschnitt häufiger digitale Medien im Unterricht nutzen als Lehrpersonen, die dem Einsatz digitaler Medien im Unterricht eher skeptisch gegenüberstehen.

3.2 Prozessfaktoren auf der Schulebene

Im Folgenden werden relevante Prozessmerkmale auf der Schulebene näher in den Fokus der Betrachtung gerückt. In den nächsten Abschnitten werden dafür (a) die Nutzung digitaler Medien, (b) die Rolle der Förderung von computer- und informationsbezogenen Kompetenzen sowie (c) das Ausmaß der Lehrerkooperation näher fokussiert.

3.2.1 Nutzung digitaler Medien

Empirische Befunde zur Häufigkeit der Nutzung digitaler Medien im Unterricht liegen im internationalen Vergleich und auf Grundlage von repräsentativen Stichproben mit der ICILS 2013-Studie vor. So ist ein zentrales Ergebnis aus ICILS 2013, dass die Lehrpersonen in Deutschland im internationalen Vergleich digitale Medien selten nutzen. Lediglich etwa ein Drittel der Lehrkräfte gibt an, digitale Medien in ihrem Unterricht regelmäßig, d.h. mindestens einmal in der Woche einzusetzen (vgl. Eickelmann, Schaumburg, Drossel & Lorenz, 2014). Damit positionieren sich die Lehrpersonen in Deutschland auf dem letzten Platz im Vergleich der 21 Bildungssysteme. Zwar zeigen aktuelle Befunde für Deutschland ein etwas positiveres Bild (vgl. Lorenz, Endberg & Eickelmann, 2017), dennoch muss davon ausgegangen werden, dass die Häufigkeit der unterrichtlichen Nutzung von digitalen Medien durch Lehrpersonen in Deutschland verhältnismäßig gering ausgeprägt ist

3.2.2 Die Rolle der Förderung von computer- und informationsbezogenen Kompetenzen in der Schule

Der Erwerb von computer- und informationsbezogenen Kompetenzen kann auf der Schulebene durch eine Vielzahl von Maßnahmen gefördert werden. In diesem Zusammenhang machen einige Studien darauf aufmerksam, dass eine der wichtigsten Schulmerkmale die Rolle und das Engagement der Schulleitung zu sein scheint (Dexter, 2008; Eickelmann, 2011). Diese sind grundlegend an der Gestaltung der Schulumwelt beteiligt und können die Implementierung von digitalen Medien in die Schule und den Unterricht beispielsweise dadurch fördern, dass sie eine gemeinsame Schulvision und gemeinsame Ziele lancieren (Ottestad, 2013), bestimmte schulische Strategien (Eickelmann et al., 2017) oder ganz allgemein schulische Richtlinien und Curricula zur Nutzung digitaler Medien im Unterricht implementieren (Biagi & Loi, 2013).

3.2.3 Lehrerkooperation

Wie auch für andere Fächerdomänen (z.B. Lesen oder Mathematik; vgl. Bottia, Valentina, Moller, Mickelson & Stearns, 2016; Ronfeldt & Owens, 2015) hat sich auch für die Förderung von computerbezogenen Kompetenzen bzw. für die Nutzung digitaler Medien das Ausmaß der Lehrerkooperation als relevant herausgestellt (Kiru, 2018). So zeigen vorliegende Studien sowohl, dass Lehrerkooperation an Schulen durch digitale Medien ganz allgemein intensiviert werden kann (Eickelmann & Schulz-Zander, 2008) als auch dass intensivere Formen der Lehrerkooperation sich positiv auf den Einsatz digitaler Medien im Unterricht auswirken (vgl. Dexter, Seashore & Anderson, 2002; Drossel & Eickelmann, 2017; Prasse, 2012).

Trotz der positiven Wirkungen, die der Lehrerkooperation im Kontext des Lernens und Lehrens zugeschrieben werden, konnte ICILS 2013 zeigen, dass die Lehrpersonen in Deutschland im Vergleich zu den 20 übrigen an ICILS 2013 teilnehmenden Bildungssystemen am wenigsten häufig mit anderen Lehrpersonen kooperieren (Drossel, Schulz-Zander, Lorenz & Eickelmann, 2016; Gerick et al., 2014).

3.3 Forschungsdesiderat und Forschungsfragen: Merkmale unerwartet erfolgreicher Schulen im internationalen Vergleich

Ausgehend von der Relevanz schulischer Merkmale, für erwartungswidrig erfolgreiche – oder auch organisational resiliente – Schulen im digitalen Zeitalter, wird im vorliegenden Beitrag den folgenden forschungsleitenden Fragestellungen nachgegangen:

1. Lassen sich in Deutschland und in weiteren Bildungssystemen Schulen identifizieren, die trotz herausfordernder sozioökonomischer Schülerkomposition überdurchschnittlich hohe mittlere Kompetenzstände im Bereich der computer- und informationsbezogenen Kompetenzen aufweisen?
2. Lassen sich diese Schulen anhand von schulischen Merkmalen zu Schultypen zusammenfassen?
3. Wie lässt sich hinsichtlich der gefundenen Schultypen insbesondere das Schulleitungshandeln als zentraler Faktor im Hinblick auf die Integration von digitalen Medien in Schule und Unterricht beschreiben?

4. Daten und Methoden

In den folgenden Abschnitten wird das methodische Vorgehen der vorliegenden Untersuchung ausgeführt. Zunächst wird die verwendete Stichprobe beschrieben (vgl. Abschnitt 4.1), bevor auf die verwendete Datengrundlage sowie auf die verwendeten Analysemethoden eingegangen wird (vgl. Abschnitt 4.2).

4.1 Stichprobe

Zur Beantwortung der oben dargestellten Fragestellungen werden Sekundäranalysen auf der Grundlage der IEA-Studie ICILS 2013 (Bos, Eickelmann, Gerick, et al., 2014) durchgeführt. Mit ICILS 2013 wurden erstmalig im internationalen Vergleich computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der achten Jahrgangsstufe computerbasiert erfasst und mittels umfangreicher Hintergrundfragebögen auch auf die Rahmenbedingungen des Kompetenzerwerbs fokussiert (vgl. u. a. Fraillon et al., 2014). Bei der Stichprobe in ICILS 2013 handelt es sich um mehrfachstratifizierte Klumpenstichproben, die in jedem der 21 an ICILS 2013 teilnehmenden Bildungssysteme eine repräsentative Stichprobe von Achtklässlerinnen und Achtklässlern sowie der Lehrpersonen, die in der achten Jahrgangsstufe unterrichten, bereitstellen. Für die Analysen in diesem Beitrag wurden neben Deutschland weiterhin alle Bildungssysteme von ICILS 2013 einbezogen, in denen die Schülerinnen und Schüler im internationalen Vergleich im Durchschnitt ein signifikant höheres Niveau der computer- und informationsbezogenen Kompetenzen erzielen als die Achtklässlerinnen und Achtklässler in Deutschland (vgl. Tabelle 1).

Tabelle 1: Übersicht über die Stichprobengrößen (*N*) differenziert nach Personen-
gruppe und Bildungssystem

Bildungssysteme ^A	Stichprobenumfang		
	Schülerinnen und Schüler	Lehrpersonen	Schulen
Australien	5 326	3 495	311
Dänemark	1 767	728	103
Deutschland	2 225	1 386	136
Kanada (Ontario)	3 777	443	193
Norwegen	2 436	1 158	138
Polen	2 870	2 228	156
Republik Korea	2 888	2 189	150
Tschechische Republik	3 066	2 126	170
Gesamt	24 355	13 753	1 357

Anmerkungen. ^A In den Niederlanden wurden keine Daten zum sozioökonomischen Status der Schülerschaft erhoben. Daher können die Niederlande in diesem Beitrag nicht berücksichtigt werden, da die Information zum mittleren SES der Schule zur Identifizierung organisatorisch resilienter Schulen benötigt werden (vgl. auch Wendt et al., 2014).

4.2 Datengrundlage und verwendete Analysemethoden

4.2.1 Verfahren zur Identifikation organisational resilienter Schulen

Zur Beantwortung der ersten Forschungsfrage werden sowohl das mittlere Niveau der computer- und informationsbezogenen Kompetenzen der Achtklässlerinnen und Achtklässler im computerbasierten Test (Plausible Values) als auch Angaben zum mittleren SES aus dem Schülerfragebogen herangezogen (HISEI, vgl. Ganzeboom, De Graaf & Treiman, 1992).

Zur Identifikation resilienter Schulen werden dazu in einem ersten Schritt Schülerindividualdaten (HISEI und CIL) auf Schulebene aggregiert. Die Daten jedes in die Analysen einbezogenen Bildungssystems werden dann sowohl für den HISEI als auch für die computer- und informationsbezogenen Kompetenzen (CIL) in drei gleich große Bereiche aufgeteilt. In einem zweiten Schritt werden dann die Schulen als resiliente Schulen identifiziert, die dem unteren HISEI-Drittel und gleichzeitig dem oberen Drittel des Leistungsspektrums zuzuordnen sind.

4.2.2 Strukturentdeckende Verfahren und Bestimmung einer Typologie organisatorisch resilienter Schulen in ICILS 2013

Die zweite Forschungsfrage verfolgt das Ziel, eine Typologie *resilienter* Schulen anhand schulischer Merkmalen zu entwickeln. Dafür werden Latente-Profil-Analysen

(LPA; vgl. Gibson, 1959) basierend auf Daten des Lehrerfragebogens durchgeführt. Zur Vorbereitung der Analysen wurden die Lehrerangaben zu sechs ausgewählten Schul- und Lehrermerkmalen auf Schulebene aggregiert und zwar als mittlerer Anteil der Lehrpersonen, der die jeweils höchste Antwortkategorie angegeben hat.

Mittels LPA mit Mplus 7 (Muthén & Muthén, 2012) werden die betrachteten Schulen dann anhand der in Tabelle 2 dargestellten Items zu ähnlichen Gruppen zusammengefasst (vgl. Rost, 2004). Die Ermittlung einer passenden Gruppenlösung erfolgt mittels so genannter informationstheoretischer Maße für den Modellvergleich (Akaike, 1974; Schwarz, 1978).

Tabelle 2: Für die latente Profilanalyse (LPA) verwendete schulische Merkmale

Schul- und Lehrermerkmale	Itemformulierung und relevante Antwortkategorie
schulische Nutzungshäufigkeit von digitalen Medien	Wie oft benutzen Sie einen Computer an den folgenden Orten? – In der Schule beim Unterrichten (Antwortkategorie: Jeden Tag)
Verfügbarkeit schulischer IT-Ressourcen	In welchem Ausmaß stimmen Sie den folgenden Aussagen über den Einsatz von IT im Unterricht an Ihrer Schule zu? – In meiner Schule ist der Internetzugang eingeschränkt (z. B. langsame oder instabile Verbindung). (Antwortkategorie: stimme gar nicht zu)
Kompetenzeinschätzung der Lehrpersonen	Wie gut können Sie diese Tätigkeiten selbstständig an einem Computer ausführen? – Den Lernstand von Schülerinnen und Schülern überprüfen (Antwortkategorie: Ich weiß, wie das geht.)
Rolle der Förderung von computer- und informationsbezogenen Kompetenzen in der Schule	Mit wie viel Nachdruck haben Sie in Ihrem Unterricht mit der Referenzklasse in diesem Schuljahr die folgenden IT-bezogenen Fähigkeiten gefördert? – Reflexion des Vorgehens bei der Informationssuche (Antwortkategorie: Mit starkem Nachdruck)
Einstellungen des Kollegiums zu Potenzialen digitaler Medien für das Lehren und Lernen	In welchem Ausmaß stimmen Sie den folgenden Aussagen über den Einsatz von IT im Unterricht und zum Lernen an ihrer Schule zu? – Hilft den Schülerinnen und Schülern, Informationen wirksamer zu vertiefen und zu verarbeiten (Antwortkategorie: stimme voll zu)
Ausmaß Lehrerkooperation zur Verbesserung des Lernens mit digitalen Medien	In welchem Ausmaß stimmen Sie den folgenden Methoden und Prinzipien in Bezug auf die IT-Nutzung im Unterricht und für das Lernen zu? – Ich arbeite mit anderen Lehrkräften daran, die IT-Nutzung für den Unterricht im Klassenzimmer zu verbessern. (Antwortkategorie: stimme voll zu)

4.2.3 Die Relevanz des Schulleitungshandeln: deskriptive Statistiken

Zur Beantwortung der dritten Forschungsfrage werden differenziert nach den identifizierten Schultypen deskriptive Statistiken unter Bezugnahme von internationalen und nationalen Daten aus dem Schulleitungsfragebogen berichtet. Auf diese Weise wird die bisher eingenommene Lehrerperspektive um die Schulleitungsperspektive erweitert. Die Schulleitungsangaben werden auf die Schüler-ebene gewichtet, sodass die deskriptiven Statistiken auf die passenden Anteile der Achtklässlerinnen und Achtklässler betreffen, die die jeweilige Schule besuchen. Unterschiede in den Anteilen zwischen den Schultypen werden via t-Test zu-

fallskritisch abgesichert. Tabelle 3 zeigt die einbezogenen Schulleitungselemente im Überblick.

Tabelle 3: Variablen zur vertiefenden Analyse des Schulleitungshandelns

Merkmale des Schulleitungshandelns	Itemformulierung und relevante Antwortkategorie
Priorität der Schulleitung: Schaffung von Anreizen für die Lehrkräfte, um die Integration der IT-Nutzung in den Unterricht zu fördern	Welche Prioritäten gibt es an Ihrer Schule bezüglich folgender Möglichkeiten zur Unterstützung des Einsatzes von IT im Unterricht und zum Lernen? Schaffung von Anreizen für die Lehrkräfte, um die Integration der IT-Nutzung in den Unterricht zu fördern (Antwortkategorie: Hohe Priorität)
Priorität der Schulleitung: Entwicklung einer gemeinsamen pädagogischen Vision im Kollegium	Welche Priorität messen Sie den nachfolgenden Bereichen bei? Entwicklung einer gemeinsamen pädagogischen Vision im Kollegium (Antwortkategorie: Hohe Priorität)
Priorität der Schulleitung: Förderung des Computereinsatzes im Fachunterricht	Welche Priorität messen Sie den nachfolgenden Bereichen bei? Förderung des Computereinsatzes im Fachunterricht (Antwortkategorie: Hohe Priorität)
Priorität der Schulleitung: Entwicklung eines Schulkonzeptes zur systematischen Planung des Computereinsatzes in Lehr-/Lernkontexten	Welche Priorität messen Sie den nachfolgenden Bereichen bei? Entwicklung eines Schulkonzeptes zur systematischen Planung des Computereinsatzes in Lehr-/Lernkontexten (Antwortkategorie: Hohe Priorität)
Maßnahmen der Schulleitung in den letzten 3 Jahren: Bereitstellung von Entlastungsstunden zur Planung von Unterrichtsinnovationen mit digitalen Medien	Welche der folgenden Maßnahmen hat die Schulleitung in den letzten drei Jahren eingeleitet? Bereitstellung von Entlastungsstunden zur Planung von Unterrichtsinnovationen mit digitalen Medien (Antwortkategorie: Ja)
Maßnahmen der Schulleitung in den letzten 3 Jahren: Einführung von Lehrerteams/Arbeitsgruppen zur Koordination pädagogischer Innovationen	Welche der folgenden Maßnahmen hat die Schulleitung in den letzten drei Jahren eingeleitet? Einführung von Lehrerteams/Arbeitsgruppen zur Koordination pädagogischer Innovationen (Antwortkategorie: Ja)
Maßnahmen der Schulleitung in den letzten 3 Jahren: Durchführung von Unterrichtsbesuchen	Welche der folgenden Maßnahmen hat die Schulleitung in den letzten drei Jahren eingeleitet? Durchführung von Unterrichtsbesuchen (Antwortkategorie: Ja)
Thematisierung des Erwerbs computer- und informationsbezogener Kompetenzen im Schulprogramm	Wird in Ihrem Schulprogramm der Erwerb computer- und informationsbezogener Kompetenzen thematisiert? (Antwortkategorie: Ja)

5. Ergebnisse zur Identifikation und Beschreibung organisational resilienter Schulen im digitalen Zeitalter auf der Basis von ICILS 2013

Im Folgenden werden die Analysen entlang der in Abschnitt 3.3 formulierten Forschungsfragen berichtet. Dazu werden zunächst die Analysen zur Identifikation von organisatorisch resilienten Schulen im Bereich der computer- und informationsbezogenen Kompetenzen berichtet (vgl. Abschnitt 5.1), bevor die Typisierung dieser Schulen in Abschnitt 6.2 dargestellt wird. In einem abschließenden Analyseschritt werden die Angaben der Schulleitung zu einer näheren

Beschreibung der resilienten Schulen anhand des Schulleitungshandelns genutzt (vgl. Abschnitt 5.3).

5.1 Ergebnisse zur ersten Forschungsfrage: Identifikation von resilienten Schulen in ICILS 2013

Wie in Abschnitt 4.2.1 beschrieben, zeichnen sich die hier betrachteten organisatorisch resilienten Schulen dadurch aus, dass ihre Schülerschaft im Mittel dem unteren HISEI-Drittel und gleichzeitig dem oberen Drittel des Leistungsspektrums in den computer- und informationsbezogenen Kompetenzen zuzuordnen ist. Tabelle 4 zeigt die Verteilung differenziert nach teilnehmendem Bildungssystem.

Tabelle 4: Anteil organisational resilienter und nicht resilienter Schulen in den teilnehmenden Bildungssystemen (absteigend sortiert nach dem Anteil resilienter Schulen)

Bildungssystem	Gesamt		Unterer HISEI	Mittlerer HISEI	Hoher HISEI
	(N)	Resilient	Vergleichsschulen		
Polen	156	19.9	53.8	25.0	1.3
Deutschland	136	11.8	27.9	44.1	16.2
Australien	311	10.3	19.3	58.5	11.9
Republik Korea	150	7.3	21.3	55.3	16.0
Kanada (Ontario)	193	4.1	11.9	63.2	5.2
Norwegen	138	2.2	20.3	53.6	23.9
Dänemark	103	1.9	18.4	60.2	19.4
Tschechische Republik	170	1.2	42.9	35.3	20.6
Gesamt	1357	7.7	26.3	50.3	13.5

Die Ergebnisse zeigen, dass über alle acht Bildungssysteme hinweg ein Anteil von 7.7 % der Schulen als resilient identifiziert werden konnten. Die Anteile der einzelnen Bildungssysteme variieren zwischen 1.2 % in der Tschechischen Republik und 19.9 % in Polen. Der entsprechende Anteil an unerwartet erfolgreichen Schulen liegt in Deutschland bei 11.8 %. Die Tabelle zeigt darüber hinaus, dass resiliente Schulen im Vergleich zu den anderen Schulen aus dem unteren HISEI-Spektrum (Vergleichsschulen) in allen betrachteten Bildungssystemen unterrepräsentiert sind. Aus der Analyse zur Identifizierung der resilienten Schulen gehen insgesamt 105 Schulen in die weiteren Analysen ein.

5.2 Ergebnisse zur zweiten Forschungsfrage: Eine Typologie resilienter Schulen in ICILS 2013

Zur Vorbereitung der LPA werden zunächst deskriptive Statistiken der zugrundeliegenden Schul- und Lehrermerkmale ermittelt (vgl. Tabelle 5).

Tabelle 5: Deskriptive Statistiken der Schul- und Lehrermerkmale (aggregierte Angaben der Lehrpersonen)

Merkmal	Mittelwert	Median	SD	MIN	MAX
Schulische Nutzungshäufigkeit digitaler Medien	36.7	33.3	25.9	0.0	100.0
Verfügbarkeit IT-Ressourcen (Schulebene)	17.4	12.5	18.2	0.0	77.8
Kompetenzeinschätzung der Lehrpersonen	74.4	78.9	19.3	25.0	100.0
Rolle der Kompetenzförderung im Bereich CIL	64.8	66.7	19.2	0.0	100.0
Einschätzung der Potenziale für das Lernen	22.7	21.4	17.5	0.0	100.0
Ausmaß der Lehrkooperation	13.6	10.0	15.9	0.0	100.0

Die Ergebnisse verdeutlichen, dass alle Variablen insgesamt genug Varianz für die Analysen aufweisen. Um die zweite Forschungsfrage zu beantworten und um eine Typologie resilienter Schulen in ICILS 2013 zu identifizieren, werden mit Mplus LPAs mit unterschiedlicher Gruppenanzahl berechnet. Um aus diesen unterschiedlichen Lösungen die Gruppenlösung zu wählen, die am besten auf die Daten passt, werden sogenannte informationstheoretische Maße für den Modellvergleich genutzt (BIC, AIC, SABIC; vgl. Abschnitt 4.2.2). Tabelle 6 zeigt die Kennwerte der einzelnen Modelle im Überblick.

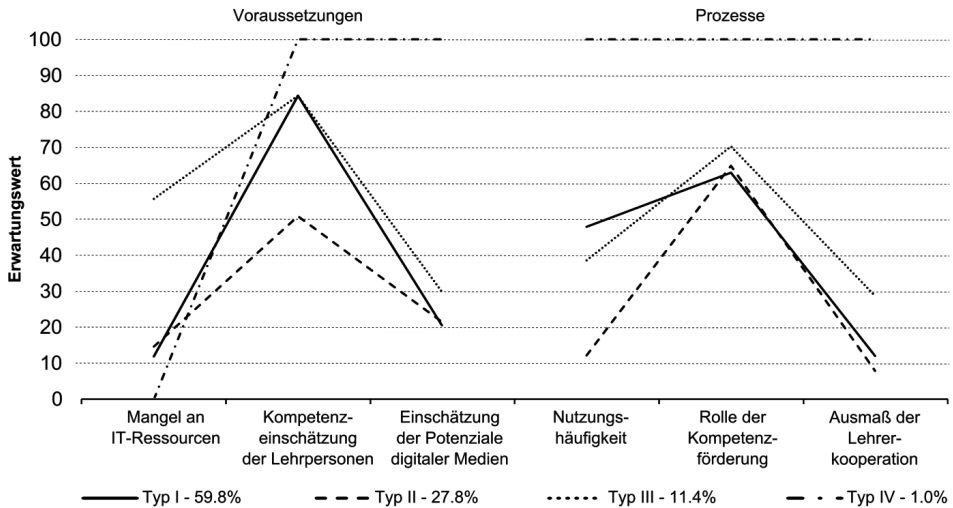
Tabelle 6: Übersicht über verschiedene Modellgütemaße der verschiedenen Gruppenlösungen (gewählte Gruppenlösung kursiv hervorgehoben)

Gruppenlösung	Parameter	AIC ¹	BIC ²	BIC _{Sample} ³
2-Gruppen-Lösung	19	5057.420	5106.339	5046.344
3-Gruppen-Lösung	26	5013.195	5080.138	4998.039
<i>4-Gruppen-Lösung</i>	<i>33</i>	<i>4979.743</i>	<i>5064.708</i>	<i>4960.506</i>
5-Gruppen-Lösung	40	4964.419	5067.408	4941.102
6-Gruppen-Lösung	47	4959.290	5080.301	4931.892
7-Gruppen-Lösung	54	4954.993	5094.028	4923.515
8-Gruppen-Lösung	61	4949.797	5106.854	4914.238
9-Gruppen-Lösung	68	4945.715	5120.795	4906.075

Anmerkungen: ¹ Akaike-Information-Criterion, ² Bayes-Information-Criterion, ³ Sample Size Adjusted BIC

Die Kennwerte der LPA weisen darauf hin, dass sich die identifizierten Schulen in Bezug auf schulische Merkmale am besten durch eine 4-Gruppen-Lösung beschreiben lassen (BIC_{\min} ; vgl. Rost, 2004). Abbildung 1 veranschaulicht die Ergebnisse der 4-Gruppen-Lösung, die im Folgenden näher beschrieben wird.

Abbildung 1: Erwartungswerte der sechs Items für die vier Typen der LPA-Lösung mit Anteil des Typs an der untersuchten Gesamtstichprobe (N = 105 Schulen)



Schultyp I: „Kompetente Schulen“ (59.8 % aller resilienten Schulen und 21.4 % der resilienten Schulen in Deutschland)

Schulgruppe I umfasst nahezu drei Fünftel aller Schulen über die acht einbezogenen Bildungssysteme hinweg. Charakteristisch für diese Schulgruppe ist eine vergleichsweise hohe Nutzungshäufigkeit digitaler Medien im Unterricht bei gleichzeitig wenig eingeschränkten IT-Ressourcen. Die (selbsteingeschätzten) IT-Kompetenzen der Lehrkräfte sind vergleichsweise hoch ausgeprägt. Die Rolle der Kompetenzförderung an der Schule ist mittelmäßig ausgeprägt. Resümierend lässt sich dieser Schultyp durch kompetente Lehrpersonen und gute Rahmenbedingungen beschreiben.

Schultyp II: „Zurückhaltende Schulen“ (27.8 % aller resilienten Schulen und 71.4 % der resilienten Schulen in Deutschland)

Mehr als ein Viertel (27.8 %) aller Schulen kann dem Schultyp II zugeordnet werden, welcher sich durch die geringste Nutzungshäufigkeit digitaler Medien für das Lehren und Lernen auszeichnet. Die schulischen IT-Ressourcen sind dabei allerdings kaum beeinträchtigt. Die selbsteingeschätzten Kompetenzen der Lehrpersonen liegen im mittleren Bereich. Potenziale digitaler Medien werden an Schulen dieses Schultyps wenig gesehen. Das Ausmaß an Lehrerkooperation ist gering ausgeprägt.

Schultyp III: „Pragmatische Schulen“ (11.4 % aller resilienten Schulen und 7.1 % der resilienten Schulen in Deutschland)

In Schulen der drittgrößten Gruppe (11.4 %) ist die Nutzungshäufigkeit digitaler Medien im Unterricht durchschnittlich ausgeprägt. Schulen, die dieser Schulgruppe zugeordnet werden können, zeichnen sich auch durch Mängel in den IT-Ressourcen aus. Die Kompetenzen der Lehrpersonen werden dagegen hoch eingeschätzt. Potenziale digitaler Medien werden wenig gesehen. Das Ausmaß der IT-bezogenen Lehrerkoooperation ist etwas höher ausgeprägt als in den bisherigen beiden Schulgruppen.

Schultyp IV: „Vorreiterschulen“ (1.0 % aller resilienten Schulen und 0.0 % der resilienten Schulen in Deutschland)

Schulen, die dem Schultyp IV zugeordnet werden können, zeichnen sich durch eine hohe Nutzungshäufigkeit digitaler Medien im Unterricht aus. Weiterhin werden von den Lehrpersonen keine Mängel in der IT-Ausstattung gesehen. Die Lehrpersonen schätzen ihre eigenen computerbezogenen Kompetenzen, die Wichtigkeit der Förderung der computer- und informationsbezogenen Kompetenzen ihrer Schülerinnen und Schüler und die Potentiale von digitalen Medien für das Lernen und Lehren als hoch ein. Weiterhin ist die IT-bezogene Lehrerkoooperation stark ausgeprägt. Schulen in dieser Schulgruppe können im Vergleich zu den anderen drei Gruppen als besonders ambitioniert bezeichnet werden.

6.3 Ergebnisse zur dritten Forschungsfrage

Gegenstand der dritten Fragestellung ist im Hinblick auf organisational resiliente Schulen in Deutschland, wie sich das Schulleitungshandeln im Hinblick auf die Integration von digitalen Medien in Schule und Unterricht innerhalb der identifizierten Schultypen in Deutschland beschreiben lässt. Wie in Abschnitt 4 erläutert, werden zur Berücksichtigung der Schulleitungsperspektive bei der Beschreibung der Schultypen acht Schulleitungssitems herangezogen. Tabelle 7 zeigt die Ergebnisse der deskriptiven Analysen.

Tabelle 7: Deskriptive Analysen zur Ausprägung verschiedener Merkmale des Schulleitungshandelns in den Gruppen I bis III

	Schultyp I: Kompetente Schulen		Schultyp II: Zurückhaltende Schulen		Schultyp III: Pragmatische Schulen	
	%	(SE)	%	(SE)	%	(SE)
16h: Priorität der Schulleitung: Schaffung von Anreizen für die Lehrkräfte, um die Integration der IT-Nutzung in den Unterricht zu fördern ¹	46.1	(23.9)	0.0	(0.0)	0.0	(0.0)
21a Priorität der Schulleitung: Entwicklung einer gemeinsamen pädagogischen Vision im Kollegium ¹	100.0	(0.0)	81.6	(13.3)	0.0	(0.0)
21g: Priorität der Schulleitung: Förderung des Computereinsatzes im Fachunterricht ¹	72.8	(13.2)	21.4	(14.5)	n	n
21h: Priorität der Schulleitung: Entwicklung eines Schulkonzeptes zur systematischen Planung des Computereinsatzes in Lehr-/Lernkontexten ¹	72.8	(13.2)	21.6	(14.7)	0.0	(0.0)
20a: Maßnahmen der Schulleitung in den letzten 3 Jahren: Bereitstellung von Entlastungsstunden zur Planung von Unterrichtsinnovationen mit digitalen Medien ²	72.8	(13.2)	33.7	(17.6)	0.0	(0.0)
20b: Maßnahmen der Schulleitung in den letzten 3 Jahren: Einführung von Lehrerteams/Arbeitsgruppen zur Koordination pädagogischer Innovationen ²	72.8	(13.2)	50.3	(18.9)	n	n
20g: Maßnahmen der Schulleitung in den letzten 3 Jahren: Durchführung von Unterrichtsbesuchen ²	100.0	(0.0)	39.7	(19.4)	0.0	(0.0)
26a: Thematisierung des Erwerbs computer- und informationsbezogener Kompetenzen im Schulprogramm ²	100.0	(0.0)	63.9	(25.3)	n	n

Anmerkungen: ¹ Kategorie „Hohe Priorität“

² Kategorie „Ja“

ⁿ Zu hoher Standardfehler aufgrund der kleinen Fallzahlen für Schultyp 3, sodass Ergebnis nicht belastbar ist

Zunächst wird deutlich, dass sich das Schulleitungshandeln an kompetenten Schulen (Schultyp I) durch klare Prioritäten der Schulleitung auszeichnet. Alle Achtklässlerinnen und Achtklässler an einer als „kompetent“ charakterisierten Schule (Schultyp I), besuchen eine Schule, an der die Schulleitung eine hohe Priorität auf die Entwicklung einer gemeinsamen pädagogischen Vision legt. Nahezu drei Viertel der Achtklässlerinnen und Achtklässler an Schulen dieses Typs besuchen eine Schule, an der die Schulleitung darüber hinaus eine hohe Priorität auf die Förderung des Computereinsatzes im Fachunterricht sowie auf die Entwicklung eines Schulkonzeptes zur systematischen Planung des Computereinsatzes in Lehr-/Lernkontexten legt. Weiterhin besuchen nahezu drei Viertel der Achtklässlerinnen und Achtklässler an einer Schule die-

ses Schultyps eine Schule, an der die Schulleitung in den letzten drei Jahren vor Erhebungszeitpunkt Maßnahmen zur Bereitstellung von Entlastungsstunden zur Planung von Unterrichtsinnovationen mit digitalen Medien sowie zur Einführung von Lehrerteams/Arbeitsgruppen zur Koordination pädagogischer Innovationen initiiert hat. An allen Schulen dieses Schultyps führt die Schulleitung Unterrichtsbesuche durch. Ebenso wird an allen Schulen dieses Schultyps der Erwerb computer- und informationsbezogener Kompetenzen im Schulprogramm thematisiert.

Das Schulleitungshandeln an den *zurückhaltenden Schulen* (Schultyp II) und damit an dem Schultyp, der in Deutschland in Bezug auf die Resilienz im Kontext computer- und informationsbezogener Kompetenzen am häufigsten zu finden ist, ist erkennbar weniger stark ausgeprägt. So besucht keine Achtklässlerin bzw. kein Achtklässler an einer Schule dieses Typs II eine Schule, an der die Schulleitung der Schaffung von Anreizen für die Lehrkräfte, um die Integration der IT-Nutzung in den Unterricht zu fördern, eine hohe Priorität zuweist. Dagegen besuchen mehr als vier Fünftel der Achtklässlerinnen und Achtklässler an einer Schule dieses Typs eine Schule, an der die Schulleitung eine hohe Priorität auf die Entwicklung einer gemeinsamen pädagogischen Vision im Kollegium legt. Weiterhin besucht weniger als ein Viertel der Achtklässlerinnen und Achtklässler an einer Schule dieses Typs eine Schule, an der die Schulleitung die Entwicklung eines Schulkonzeptes zur systematischen Planung des Computereinsatzes in Lehr-/Lernkontexten hoch priorisiert. Mittlere bis niedrige Anteile finden sich für die Maßnahmen der Schulleitung in den letzten drei Jahren vor Erhebungszeitpunkt. So besucht beispielsweise etwa ein Drittel der Schülerinnen und Schüler an einer Schule dieses Schultyps II eine Schule, an der die Schulleitung Entlastungsstunden zur Planung von Unterrichtsinnovationen mit digitalen Medien bereitstellt.

Das Schulleitungshandeln an den *pragmatischen Schulen* ist insgesamt eher schwach ausgeprägt. Im Handeln der Schulleitung liegen keine hohen Prioritäten hinsichtlich der Integration digitaler Medien in schulische Strukturen und Konzepte. Auch im Hinblick auf die Umsetzung von schulischen Maßnahmen in dieser Hinsicht lässt sich kein nennenswertes Engagement identifizieren. Aufgrund der kleinen Fallzahlen in dieser Kategorie und die damit verbundenen hohen Standardfehler können allerdings nicht für alle Bereiche Befunde vorgelegt werden (vgl. Tabelle 7).

Da es in Deutschland keine Schulen des Typs *Vorreiterschulen* gibt, entfällt an dieser Stelle die Beschreibung des Schulleitungshandelns.

6. Diskussion der Ergebnisse

Vor dem Hintergrund der zunehmenden Bedeutung der Vermittlung von Kompetenzen im Umgang mit neuen Technologien und digitalen Informationen für eine erfolgreiche Teilhabe an der Gesellschaft, stehen erwartungswidrig erfolg-

reiche Schulen im Fokus dieses Beitrags, also Schulen, die sich trotz der Befunde zu sozialen Disparitäten der ICILS-2013-Studie und trotz einer herausfordernden Schülerkomposition durch ein hohes Niveau computer- und informationsbezogener Schülerkompetenzen charakterisieren lassen. Mittels Sekundäranalysen der ICILS-2013-Daten wurde den Fragen nachgegangen, ob sich in Deutschland und weiteren Bildungssystemen solche Schulen identifizieren lassen, wie sich diese anhand von schulischen Merkmalen zu einer Typologie zusammenfassen lassen (Schultypen) und wie sich hinsichtlich der gefundenen Schultypen das Schulleitungshandeln beschreiben lässt.

Die Ergebnisse zur ersten Fragestellung zeigen, dass das Phänomen organisationaler Resilienz auch in dem Bereich der computer- und informationsbezogenen Kompetenzen relevant erscheint und über verschiedene Bildungssysteme hinweg im unterschiedlichen Maße auftritt. Insgesamt variiert die Anzahl *resilienter* Schulen zwischen den betrachteten Bildungssystemen. In Deutschland liegt der Anteil mit nahezu 12 % verhältnismäßig hoch. Dies ist insofern erfreulich, als dass es in Deutschland offensichtlich eine substantielle Anzahl von Schulen gibt, die es trotz benachteiligter Lage ihrer Schülerinnen und Schüler schaffen, diese kompetent im Umgang mit digitalen Medien und damit zukunftsfähig auszubilden.

Die Befunde zur zweiten Fragestellung verdeutlichen, dass es möglich ist, resiliente Schulen anhand von Schulmerkmalen zu gruppieren. So zeigen die Ergebnisse der LPA, dass sich über die verschiedenen Bildungssysteme hinweg vier Gruppen resilienter Schulen bilden lassen. Für Deutschland zeigt sich das Ergebnis, dass für einen Großteil der resilienten Schulen die klassischen Merkmale von Schulen, die im Bereich der Integration von digitalen Medien in das Lehren und Lernen (siehe Abschnitt 2) relevant sind, eine eher geringe Rolle zu spielen scheinen. Für rund ein Fünftel der Schulen in Deutschland finden sich jedoch in den resilienten Schulen Lehrpersonen, die ihre eigenen Kompetenzen im Hinblick auf den Umgang mit digitalen Medien hoch einschätzen und indem zudem kein Mangel an IT-Ressourcen berichtet wird. Dieses Ergebnis unterstreicht die besondere Rolle der Lehreraus- und -fortbildung sowie die Notwendigkeit der Schaffung einer geeigneten schulischen IT-Infrastruktur. Werden insbesondere diese Faktoren berücksichtigt, kann es möglicherweise gelingen, soziale Disparitäten, die sich im Kontext der ICILS-2013-Studie in Deutschland als besonders besorgniserregend herausgestellt haben, auf Schulebene zu reduzieren.

In Bezug auf die dritte Forschungsfrage wird deutlich, dass sich das Schulleitungshandeln zum Teil erheblich zwischen den identifizierten Schultypen unterscheidet. Besonders ausgeprägt ist das Schulleitungshandeln, und damit u. a. Prioritätensetzung der Schulleitung im Bereich der Integration digitaler Medien in die Schule, an den Schulen, die als „kompetente Schulen“ (Schultyp I) charakterisiert wurden. Dies erscheint erwartungskonform, da die Schulleitung als ein zentraler Akteur im Prozess der Implementation digitaler Medien in schulische Prozesse gilt und in dieser Hinsicht sowohl als Fach- als auch Prozess- und Machtpromotor wirken kann (vgl. u. a. Eickelmann & Gerick, 2018; Hunneshagen, 2005). Dass das Schulleitungshandeln an den betrachteten Schulen mit benachteiligter Schüler-

schaft besonders wichtig ist, ist jedoch ein Befund, der sich erstmals mit der vorliegenden Untersuchung in dieser Deutlichkeit empirisch zeigt. Hinsichtlich der schulischen Praxis wird daher die Relevanz des Schulleitungshandelns und damit auch der Schulleitungsprofessionalisierung im Kontext des digitalen Wandels deutlich, der bisher oft noch – auch bildungspolitisch – zu wenig mitgedacht wird.

Die Grenzen der vorliegenden Untersuchung liegen jedoch in dem Querschnittsdesign der Anlage der Studie. Im Ergebnis kann nicht abschließend geklärt werden, ob es tatsächlich auf die Schule und ihre Lehr- und Lernaktivitäten zurückzuführen ist, dass die Jugendlichen trotz wenig privilegierter sozialer Lage so ein hohes Niveau an computer- und informationsbezogenen Kompetenzen aufweisen oder ob dies auch auf andere Gründe zurückgeführt werden kann. Weiterhin könnte der Befund für Deutschland, dass nur so wenige Schulen dem Schultyp „kompetente Schulen (Schultyp I)“ zugeordnet werden konnten in methodischer Hinsicht darauf zurückzuführen sein, dass als Kriterium bei der LPA eine tägliche Nutzungshäufigkeit gewählt wurde, was – im Vergleich zu anderen Ländern in Deutschland nur an ganz wenigen Schulen zum Zeitpunkt der Erhebung der Daten überhaupt realisiert ist (siehe dazu auch Abschnitt 2). Hier werden weitere Untersuchungen, u. a. im Kontext des zweiten Zyklus‘ der Studie ICILS (ICILS 2018) Aufschluss über den Bestand der Ergebnisse geben können. Zu diskutieren ist zudem, wie das Kriterium der Nutzungshäufigkeit im Vergleich zu Maßen der Nutzungsqualität digitaler Medien empirisch bedeutsamer ist. Diesbezüglich weisen insbesondere qualitative Untersuchungen mit Videostudien- und Interviewanteilen besondere Potenziale auf. Weiterhin ergibt sich für Fragen nach Entwicklungen und kausalen Zusammenhängen die Perspektive der Analysen im Trendvergleich, die sich schließlich mit der Veröffentlichung der Ergebnisse und Daten von ICILS 2018 im Jahr 2019 und 2020 ergeben.

Literatur

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Biagi, F. & Loi, M. (2013). Measuring ICT use and learning outcomes. Evidence from recent econometric studies. *European Journal of Education*, 48(1), 28–42.
- Bonsen, M., Bos, W. & Rolff, H.-G. (2008). Zur Fusion von Schuleffektivitäts- und Schulentwicklungsforschung. In W. Bos, H. G. Holtappels, H. Pfeiffer, H.-G. Rolff & R. Schulz-Zander (Hrsg.), *Jahrbuch Schulentwicklung* (Band 15, S. 11–40). Weinheim: Juventa.
- Bos, W., Eickelmann, B. & Gerick, J. (2014). Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich. In W. Bos, B. Eickelmann, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil, R. Schulz-Zander & H. Wendt (Hrsg.), *ICILS 2013. Computer und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich* (S. 113–145). Münster: Waxmann.
- Bos, W., Eickelmann, B., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., Senkbeil, M., Schulz-Zander, R. & Wendt, H. (Hrsg.) (2014). *ICILS 2013*.

- Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich.* Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G. & Valtin, R. (Hrsg.) (2003). *Erste Ergebnisse aus IGLU – Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich.* Münster: Waxmann.
- Bos, W., Lorenz, R., Endberg, M., Eickelmann, B., Kammerl, R. & Welling, S. (Hrsg.) (2016). *Schule digital – der Länderindikator 2016. Kompetenzen von Lehrpersonen der Sekundarstufe I im Umgang mit digitalen Medien im Bundesländervergleich.* Münster: Waxmann.
- Bos, W., Lorenz, R., Endberg, M., Schaumburg, H., Schulz-Zander, R. & Senkbeil, M. (Hrsg.) (2015). *Schule digital – der Länderindikator 2015. Vertiefende Analysen zur schulischen Nutzung digitaler Medien im Bundesländervergleich.* Münster: Waxmann.
- Bottia, M. C., Valentina, L., Moller, S., Mickelson, R. A. & Stearns, E. (2016). Teacher collaboration and Latinos/as' mathematics achievement trajectories. *American Journal of Education*, 122, 505–535.
- Buabend-Andoh, C. (2012). Factors influencing teachers' adoption and integration of information and communication technology into teaching: A review of the literature. *International Journal of Education and Development using Information and Communication Technology*, 8(1), 136–155.
- Dexter, S. (2008). Leadership for IT in schools. In J. Voogt & G. Knezek (Hrsg.), *International handbook of information technology in primary and secondary education* (S. 543–554). New York, NY: Springer.
- Dexter, S., Seashore, K. R. & Anderson, R. E. (2002). Contributions of professional learning community to exemplary use of ICT. *Journal of Computer Assisted Learning*, 18(4), 489–497.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich. Schule, Sozialpädagogik, Hochschule* (S. 73–92). Weinheim: Beltz.
- Drossel, K. & Eickelmann, B. (2017). Teachers' participation in professional development concerning the implementation of new technologies in class – Different types of teachers and their relationship with the use of computers, ICT self-efficacy and emphasis on teaching ICT. *Large-scale Assessments in Education*, 5(19), 1–13.
- Drossel, K., Eickelmann, B., van Ophuysen, S. & Bos, W. (2018). Why teachers cooperate – An expectancy-value model of teacher cooperation. *European Journal of Psychology of Education*, 33(2), 1–22.
- Drossel, K., Schulz-Zander, R., Lorenz, R. & Eickelmann, B. (2016). Gelingensbedingungen IT-bezogener Lehrerverkooperation als Merkmal von Schulqualität. In B. Eickelmann, J. Gerick, K. Drossel & W. Bos (Hrsg.), *ICILS 2013. Vertiefende Analysen zu computer- und informationsbezogenen Kompetenzen von Jugendlichen* (S. 143–167). Münster: Waxmann.
- Eickelmann, B. (2011). Supportive and hindering factors to a sustainable implementation in ICT in schools. *Journal for Educational Research Online*, 3(1), 75–103.
- Eickelmann, B. (2018). Digitalisierung in der schulischen Bildung – Entwicklungen, Befunde und Perspektiven für die Schulentwicklung und die Bildungsforschung. In N. McElvany, F. Schwabe, W. Bos & H. G. Holtappels (Hrsg.), *Digitalisierung in der schulischen Bildung. Chancen und Herausforderungen* (S. 11–25). Münster: Waxmann.
- Eickelmann, B., Bos, W., Gerick, J. & Kahnert, J. (2014). Anlage, Durchführung und Instrumentierung von ICILS 2013. In W. Bos, B. Eickelmann, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil, R. Schulz-Zander & H. Wendt (Hrsg.), *ICILS 2013. Computer- und informationsbezogene Kompetenzen*

- von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich (S. 43–81). Münster: Waxmann.
- Eickelmann, B. & Gerick, J. (2018). Herausforderungen und Zielsetzungen im Kontext der Digitalisierung von Schule und Unterricht. Teil 3: Neue Aufgaben für die Schulleitung. *Schulverwaltung Hessen/Rheinland-Pfalz*, 23(7–8), 200–202.
- Eickelmann, B., Gerick, J. & Koop, C. (2017). ICT use in mathematics lessons and the mathematics achievement of secondary school students by international comparison: Which role do school level factors play? *Journal for Education and Information Technologies*, 22(4), 1527–1551.
- Eickelmann, B., Schaumburg, H., Drossel, K. & Lorenz, R. (2014). Schulische Nutzung von neuen Technologien in Deutschland im internationalen Vergleich. In W. Bos, B. Eickelmann, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil, R. Schulz-Zander & H. Wendt (Hrsg.), *ICILS 2013. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich* (S. 197–229). Münster: Waxmann.
- Eickelmann, B. & Schulz-Zander, R. (2008). Schuleffektivität, Schulentwicklung und digitale Medien. In W. Bos, H. G. Holtappels, H. Pfeiffer & H.-G. Rolff (Hrsg.), *Jahrbuch der Schulentwicklung Band 15* (S. 157–193). Weinheim: Juventa.
- Eickelmann, B. & Vennemann, M. (2017). Teachers' attitudes and beliefs regarding ICT in teaching and learning in European countries. *European Educational Research Journal*, 16(6), 733–761.
- Endberg, M. & Lorenz, R. (2017). Selbsteinschätzung medienbezogener Kompetenzen von Lehrpersonen der Sekundarstufe I im Bundesländervergleich und im Trend von 2016–2017. In R. Lorenz, W. Bos, M. Endberg, B. Eickelmann, S. Grafe & J. Vahrenhold (Hrsg.), *Schule digital – der Länderindikator 2017. Schulische Medienbildung in der Sekundarstufe I mit besonderem Fokus auf MINT-Fächer im Bundesländervergleich und Trends von 2015 bis 2017* (S. 151–176). Münster: Waxmann.
- Endberg, M., Lorenz, R. & Senkbeil, M. (2015). Einstellungen von Lehrpersonen der Sekundarstufe I zum Einsatz digitaler Medien im Unterricht. In W. Bos, R. Lorenz, M. Endberg, H. Schaumburg, R. Schulz-Zander & M. Senkbeil (Hrsg.), *Schule digital – der Länderindikator 2015. Vertiefende Analysen zur schulischen Nutzung digitaler Medien im Bundesländervergleich* (S. 95–140). Münster: Waxmann.
- European Commission (2013). *Survey of schools: ICT in education. Benchmarking access, use and attitudes to technology in Europe's schools*. Luxemburg: Publications Office of the European Union.
- Fend, H. (2008). *Schule gestalten. Systemsteuerung, Schulentwicklung und Unterrichtsqualität*. Wiesbaden: VS.
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T. & Gebhardt, E. (2014). *Preparing for life in a digital age. The IEA International Computer and Information Literacy Study International Report*. Melbourne: Springer.
- Fraillon, J., Schulz, W. & Ainley, J. (2013). *International Computer and Information Literacy Study: Assessment framework*. Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).
- Ganzeboom, H. B. G., De Graaf, P. M. & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.
- Gerick, J. (2018). School level characteristics and students' CIL in Europe – A latent class analysis approach. *Computers & Education*, 120, 160–171.
- Gerick, J., Eickelmann, B. & Bos, W. (2017). School-level predictors for the use of ICT in schools and students' CIL in international comparison. *Large-scale Assessments in Education*, 5(1), 1–13.
- Gerick, J., Eickelmann, B., Drossel, K. & Lorenz, R. (2016). Perspektiven von Schulleitungen auf neue Technologien in Schule und Unterricht. In B. Eickelmann, J. Gerick, K. Drossel & W. Bos (Hrsg.), *ICILS 2013. Vertiefende Analysen zu com-*

- puter- und informationsbezogenen Kompetenzen von Jugendlichen*. Münster: Waxmann.
- Gerick, J., Schaumburg, H., Kahnert, J. & Eickelmann, B. (2014). Lehr- und Lernbedingungen des Erwerbs computer- und informationsbezogener Kompetenzen in den ICILS-2013-Teilnehmerländern. In W. Bos, B. Eickelmann, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil, R. Schulz-Zander & H. Wendt (Hrsg.), *ICILS 2013. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich* (S. 147–196). Münster: Waxmann.
- Gibson, W. A. (1959). Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika*, 24(3), 229–252.
- Hatlevik, O. E. (2016). Examining the relationship between teacher's self-efficacy, their digital competence, strategies to evaluate information, and use of ICT at school. *Scandinavian Journal of Educational Research*, 61(5), 555–567.
- Hunneshagen, H. (2005). *Innovationen in Schulen*. Münster: Waxmann.
- Kiru, E. W. (2018). Mathematics teachers' use of information and communication technologies: An international comparison. *International Journal of Research in Education and Science*, 4(1), 165–177.
- Klieme, E., Baumert, J. & Schwippert, K. (2000). Schulbezogene Evaluation und Schulleistungsvergleiche – Eine Studie im Anschluss an TIMSS. In H.-G. Rolff, W. Bos, K. Klemm, H. Pfeiffer & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung. Band 11. Daten, Beispiele und Perspektiven* (S. 387–419). Weinheim: Juventa.
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik (2016). *Strategie der Kultusministerkonferenz „Bildung in der digitalen Welt“*. Berlin.
- Kozma, R. B. (Hrsg.) (2003). *Technology, innovation, and educational change. A global perspective. Report of the Second Information Technology in Education Study Module 2*. Washington, DC: ISTE.
- Lai, K.-W. (2008). ICT supporting the learning process: The premise, reality, and promise. In J. Voogt & G. Knezek (Hrsg.), *International handbook of information technology in primary and secondary education* (S. 215–230). Cham: Springer.
- Law, N., Pelgrum, W. J. & Plomp, T. (Hrsg.) (2008). *Pedagogy and ICT use in schools around the world. Findings from the IEA SITES 2006 Study*. Hongkong: Springer.
- Lorenz, R., Eickelmann, B. & Gerick, J. (2015). What affects students' computer and information literacy around the world? – An analysis of school and teacher factors in high performing countries. In D. Slykhuis & G. Marks (Hrsg.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2015* (S. 1212–1219). Chesapeake, VA: AACE.
- Lorenz, R., Endberg, M. & Eickelmann, B. (2017). Unterrichtsliche Nutzung digitaler Medien durch Lehrpersonen in der Sekundarstufe I im Bundesländervergleich und im Trend von 2015 bis 2017. In R. Lorenz, W. Bos, M. Endberg, B. Eickelmann, S. Grafe & J. Vahrenhold (Hrsg.), *Schule digital – der Länderindikator 2017. Schulische Medienbildung in der Sekundarstufe I mit besonderem Fokus auf MINT-Fächer im Bundesländervergleich und im Trend von 2015 bis 2017* (S. 84–121). Münster: Waxmann.
- Muijs, D., Harris, A., Chapman, C., Stoll, L. & Russ, J. (2004). Improving schools in socioeconomically disadvantaged areas – A review of research evidence. *School Effectiveness and School Improvement*, 15(2), 149–175.
- Muthén, L. K. & Muthén, B. O. (2012). Mplus 7. Los Angeles, CA: Muthén & Muthén.
- Ottestad, G. (2013). School leadership for ICT and teachers' use of digital tools. *Nordic Journal of Digital Literacy*, 8, 107–125.
- Petko, D. (2012). Teachers' pedagogical beliefs and their use of digital media in classrooms: Sharpening the focus of the 'will, skill, tool' model and integrating teachers' constructivist orientations. *Computers & Education*, 58, 1351–1359.

- Petko, D., Egger, N. & Wespi, B. (2015). Digital media adoption in schools: Bottom-up, top-down, complementary or optional? *Computers & Education*, 84, 49–61.
- Prasse, D. (2012). *Bedingungen innovativen Handelns an Schulen*. Münster: Waxmann.
- Racherbäumer, K., Funke, C., Van Ackeren, I. & Clausen, M. (2013). Schulleffektivitätsforschung und die Frage nach guten Schulen in schwierigen Kontexten. In R. Becker & A. Schulze (Hrsg.), *Bildungskontexte. Strukturelle Voraussetzungen und Ursachen ungleicher Bildungschancen* (S. 239–268). Wiesbaden: Springer.
- Ronfeldt, M. & Owens, S. (2015). Teacher collaboration in instructional teams and student achievement. *American Educational Research Journal*, 52(3), 475–514.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Bern: Huber.
- Saudelli, M. G. & Ciampa, K. (2016). Exploring the role of TPACK and teacher self-efficacy: An ethnographic case study of three iPad language arts classes. *Technology, Pedagogy and Education*, 25(2), 227–247.
- Schaumburg, H., Prasse, D., Tschackert, K. & Blömeke, S. (2007). *Lernen in Notebook-Klassen. Endbericht zur Evaluation des Projekts „1000mal1000: Notebooks im Schulranzen“*. Bonn: Schulen ans Netz e.V.
- Scheerens, J. & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Scheiter, K. (2017). Lernen mit digitalen Medien – Potenziale und Herausforderungen aus Sicht der Lehr-Lernforschung. *Schulmanagement Handbuch*, 164(4), 33–47.
- Schuhmacher, F. (Hrsg.) (2004). *Innovativer Unterricht mit neuen Medien*. Grünwald: FWU Institut für Film und Bild.
- Schulz-Zander, R. (2001). Neue Medien als Bestandteil von Schulentwicklung. In S. Aufenanger, R. Schulz-Zander & D. Spanhel (Hrsg.), *Jahrbuch Medienpädagogik 1* (S. 263–821). Opladen: Leske + Budrich.
- Schulz-Zander, R. (2004). The school online initiative in German schools: Empirical results and recommendations to improve school development. In D. W. Chapman & L. O. Mahlck (Hrsg.), *Adapting technology for school improvement: A global perspective* (S. 269–295). Paris: UNESCO International Institute for Educational Planning.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464.
- Schweizer, W. & Horn, M. (2014). Kommt es auf die Einstellung zu digitalen Medien an? *Medien + Erziehung*, 58(6), 50–62.
- Senkeil, M., Goldhammer, F., Bos, W., Eickelmann, B., Schwippert, K. & Gerick, J. (2014). Das Konstrukt der computer- und informationsbezogenen Kompetenzen in ICILS 2013. In W. Bos, B. Eickelmann, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkeil, R. Schulz-Zander & H. Wendt (Hrsg.), *ICILS 2013. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich* (S. 83–112). Münster: Waxmann.
- Tomczyk, L., Sztokowski, R., Fabiś, A., Wąsiński, A., Chudý, Š. & Neumeister, P. (2015). Selected aspects of conditions in the use of new media as an important part of the training of teachers in the Czech Republic and Poland – Differences, risks and threats. *Education and Information Technologies*, 22(3), 747–767.
- Vennemann, M. (2018). *Individual- und Kompositionseffekte und der Kompetenzzuwachs in Mathematik und Naturwissenschaft am Ende der Grundschule*. Wiesbaden: Springer.
- Voogt, J., Knezek, G., Christensen, R. & Lai, K.-W. (2018). *International handbook of information technology in primary and secondary education* (Bd. 2). Cham: Springer.
- Wendt, H., Vennemann, M., Schwippert, K. & Drossel, K. (2014). Soziale Herkunft und computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. In W. Bos, B. Eickelmann, J. Gerick, F.

Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil, R. Schulz-Zander & H. Wendt (Hrsg.), *ICILS 2013. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich* (S. 265–296). Münster: Waxmann.

Yang, X. & Leung, F. K. S. (2015). The relationships among pre-service mathematics teachers' beliefs about mathematics, mathematics teaching, and use of technology in China. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(6), 1363–1378.

Nele McElvany & Franziska Schwabe

Gender gap in reading digitally? Examining the role of motivation and self-concept

Abstract

Reading is a core prerequisite for educational success and participation in society. However, comprehensive empirical research is needed to understand how reading may differ in a digitalized world. The current study addressed the gender gap in reading digitally. It investigated competence scores along with information on (a) reading and (b) digital motivation and self-concept in 588 elementary school students. Results revealed a gender gap in reading digitally, in reading motivation and self-concept, and in motivation and self-concept in respect to working on digital devices. Only reading motivation variables predicted reading digitally, thereby providing important information on the validity of digitally based reading tests. Reading motivation was found to fully mediate the gender effect on reading digitally. Results have important implications for research and practice.

Keywords

Reading; Motivation; Digitalization; Elementary school

Geschlechtsunterschiede im Lesen an digitalen Geräten? Analysen zur Rolle von Motivation und Selbstkonzept

Zusammenfassung

Lesen ist eine zentrale Voraussetzung für Bildungserfolg und gesellschaftliche Teilhabe. Wegen der fundamentalen Bedeutung des Lesens sind umfassende empirische Erkenntnisse zu der Frage, wie sich das Lesekonstrukt im Zuge der Digitalisierung verändert, notwendig. Vor diesem Hintergrund analysierte die vorliegende Studie Geschlechtsunterschiede im Lesen an digitalen Geräten. Neben

Prof. Dr. Nele McElvany (corresponding author) · Dr. Franziska Schwabe, TU Dortmund University, Center for Research on Education and School Development (IFS), Vogelpothsweg 78, 44227 Dortmund, Germany
e-mail: nele.mcelvany@tu-dortmund.de
franziska.schwabe@tu-dortmund.de

Lesekompetenzen an digitalen Geräten wurden intrinsische Motivation und Selbstkonzepte in den Bereichen (a) Lesen und (b) Arbeit an digitalen Geräten von 588 Grundschulkindern betrachtet. Mädchen und Jungen unterschieden sich in der Lesekompetenz, in lesebezogener Motivation und Selbstkonzept sowie in Motivation und Selbstkonzept zur Arbeit an digitalen Geräten. Ausschließlich lesebezogene motivationale Merkmale sagten die Lesekompetenz vorher, was auf die Validität des Lesetests hinwies. Außerdem mediierte die intrinsische Lesemotivation die Geschlechtsunterschiede in der Lesekompetenz vollständig. Die Ergebnisse haben wichtige Implikationen für Bildungsforschung und -praxis.

Schlagworte

Lesekompetenz; Motivation; Digitalisierung; Grundschule

1. Introduction

One core development and challenge of our modern times is the digitalization of many parts of society. This also includes education in which acquiring information by reading digitally has become a frequent activity within both school instruction and out-of-school learning (Mangen & van der Weel, 2016). However, it is not yet clear what implications digitalization is having – not only for the construct of reading literacy itself but also for the research findings on reading gathered in past decades using paper-based assessment. One core finding on reading is the school age gender gap.

Gender is a core aspect of being human; and substantial research and educational measures have focused on gender differences in reading (McGeown, Goodwin, Henderson, & Wright, 2012). Results on competence have been heterogeneous, with evidence indicating not only better performance in school-age females (Mullis, Martin, Foy, & Hooper, 2017; but see also Hyde & Linn, 1988) but also slightly better performance in adult males (Organisation for Economic Cooperation and Development [OECD], 2013). A clear pattern has emerged indicating that girls are, on average, more highly motivated to read and also read more in their free time (McElvany, Kessels, Schwabe, & Kasper, 2017; McKenna, Conradi, Lawrence, Jang, & Meyer, 2012).

However, it is necessary to verify whether these results generalize to reading digitally and digital reading. Recent research has pointed out that gender differences in digital reading still exist, but that they might be smaller than in paper-based assessments (Mullis et al., 2017). One possible explanation for this emerging pattern might be found in motivational variables (e.g., Becker, McElvany, & Kortenbruck, 2010). Some evidence has suggested that motivational constructs with regard to digital devices and their usage show either no or, at times, different gender differences compared to reading motivation and behavior when reading paper-based materials (Lorenz, Gerick, Schulz-Zander, & Eickelmann, 2014).

Against this background, the current study aimed to investigate a potential gender gap in reading digitally, the predictive value of motivational variables for working with digital devices, and the predictive value of traditional motivational variables for reading in general. It also examined the potentially mediating role of these motivational variables for gender differences in the proficiency of reading digitally.

2. Theoretical background

2.1 Reading literacy: Definition, development in elementary school age, and international assessments

Reading literacy includes the ability to extract relevant information from, understand, use, and reflect on written texts. According to Kintsch (1988), a text can be represented on three different levels: (a) the surface level representing the verbatim text structure, (b) the text level including the text basis and its elaborations representing text propositions (semantics), and (c) the situational model level integrating the text content with the readers' knowledge. Following Kintsch's (1988) construction-integration model, the reading process combines a simultaneous iteration of the bottom-up construction processes of building a propositional network based on text content with the top-down integration processes based on inferences enabled by the reader's knowledge. However, both reading activities themselves as well as the assessment of reading literacy have undergone substantial mode changes in recent years through the shift from paper-based to digital-based.

Reading literacy and its assessment on digital devices can entail two systematically different concepts (see Naumann & Sälzer, 2017; cf. OECD, 2016): It can take one of two forms: (a) It can take a text and text comprehension questions that were previously presented on paper and present them as similarly as technically possible on a digital screen without any functionality differing from a book (focus on mode change; *reading digitally*). (b) It can incorporate reading in digital environments (*digital reading*) with different tasks and challenges to those in paper-based reading (e.g., navigating; finding, selecting, and understanding information efficiently out of a choice of multiple websites; comprehending animated graphics). At least in part, digital reading constitutes a different construct of reading literacy, because it requires a specific process in which text comprehension and the operation of digital environments intersect in what is often referred to as "navigation" (Delgado, Vargas, Ackerman, & Salmerón, 2018). To avoid confounding change of mode with change of content (new task characteristics), the current article focuses on the first concept: reading on digital devices as a rather pure mode change from paper-based reading to digital based reading. Because so much reading acquisition and development is now taking place digitally, it is important to investigate which features of the new mode and which student characteristics predict reading digitally.

Most children acquire reading systematically in elementary school. Different theoretical models describe various stages of children's learning to read (for alphabetic systems such as German, see, e.g., Frith, 1985). They all include the need for children to understand the connection between a (written) sign and the sound it is assigned to, and the ability to connect various sounds to an entire word that, at an advanced stage, takes the form of orthographic reading of whole units. Additionally, and in turn, the decoded word and, at a later stage, entire sentences, paragraphs, and texts then need to be related to their meaning. During the first years of schooling, most children make great progress in automatizing their decoding skills and thus their reading literacy. Typically, students start to read fluently in 2nd- or 3rd-grade (see Chall, 1983). This automatization allows them to focus their cognitive resources on the content rather than on decoding. In her five-stage model, Wolf (2008) states that roughly 9 years is the age of moving from *the decoding reader* to *the fluent, comprehending reader*. This transition is of vital importance to children's further educational paths, because this period of their schooling marks the shift from *learning to read* to *reading to learn*. Reading turns from being an educational goal itself to being mostly an educational tool or even an educational prerequisite. Thus, it is no surprise that not only international but also many national educational monitoring studies focus on assessing reading literacy in this age group.

On an international level, the *Progress in International Reading Literacy Study* (PIRLS) is investigating elementary school students' reading literacy (International Association for the Evaluation of Educational Achievement (IEA), 2015; for older students, see Program International Student Assessment (PISA), OECD, 2016). In 2016, Germany's 4th-grade students, who are mostly at the end of their elementary schooling, attained a lower middle ranking in the international comparison, but this was not significantly lower than the mean score for all EU countries or all OECD countries (Hußmann et al., 2017). However, a major concern was the systematic differences among subgroups of students and, in particular, the lower average performance of students with a less privileged socioeconomic background, of students with a migrant background, and of boys compared to girls (Hußmann et al., 2017).

In the long-term perspective from 2001 to 2016, there has been no significant change in 4th-grade students' reading achievement in Germany. However, up to now, this has been assessed continuously with paper-based tests. In 2021, the PIRLS assessment will become digital. This mode change raises questions on the validity of any comparisons with previous paper-based assessments (e.g., Zehner, Goldhammer, Lubaway, & Sälzer, 2018). Whereas there is a quite large body of research on digitalization for older students and adults, it is still largely unclear how younger students handle working on digital devices during assessments in the domain of reading; and whether, for example, motivational aspects regarding working on digital devices relate to proficiency scores – either for all students or potentially even differentially for different subgroups of students.

2.2 Digitalization of children's environment, reading on digital devices, and paper versus digitally based reading

The increasing inclusion of digital devices in children's lives and particularly in their learning environments is an everyday reality. In a recent meta-analysis, Delgado et al. (2018, p. 36) concluded that "it is clear that digital-based reading is an unavoidable part of our daily lives and an integral part of the educational realm." In a German sample of 10- to 11-year-old children, 67 % reported owning a smartphone and 32 % a tablet; and in a sample of 12- and 13-year-olds, 87 % at least sometimes used a smart phone and 58 % a tablet (BitKom Research, 2017, pp. 2–3). Six- to 13-year-old students named digital games, internet, and smartphone usage more often as their favorite leisure-time activity than reading a book, and more than one-half used the internet at least once a week (Medienpädagogischer Forschungsverbund Südwest, 2016).

Due to increased access to and use of digital devices, substantial changes have also been observed in children's reading habits: Reading is now frequently performed with digital screen technologies such as laptops or tablets and e-readers (e.g., Clark, 2011; Duncan, McGeown, Griffiths, Stothard, & Dobai, 2016). Moreover, the role of digital devices (e.g., computers, tablets, e-books, and smartphones) as reading tools for and within education is also increasing substantially, because schools worldwide are beginning to shift to digital classrooms. In the last PIRLS assessment, German 4th-grade students reported using computers or tablets for 10 to 30 minutes a day to obtain information for school or to prepare texts or presentations (Hußmann et al., 2017). In contrast to the increasing use of digital devices, a comparison between 2001 and 2016 revealed that fewer students read books daily (Goy, Valtin, & Hußmann, 2017). Unfortunately, evidence on the relation between reading literacy and digital media use is inconclusive (Rosén & Gustafsson, 2016). Moreover, as well as digital devices being used either as tools for or as a target of learning, high stakes tests are increasingly being delivered digitally. The same trend can be observed in the large-scale international student assessments such as PIRLS or PISA.

The fundamental change from exclusively paper-based reading to reading in a multimedia world raises a number of theoretical as well as empirical questions (Mangen & van der Weel, 2016). Concerning the question of equivalence of computer versus paper-based tasks, Dillon (1992) reported in an earlier literature review that speed, accuracy, and comprehension were generally worse when assessed digitally. Building on this work, Noyes and Garland (2008) summarized their review results as showing mixed evidence regarding equivalence. They suggested that analyses should include further indicators such as cognitive workload and preferences. The authors concluded that although some tasks will never be equivalent, greater equivalence is being achieved today due to the general changes. Wells (2012) randomly assigned 140 middle and high school students to either a tablet or a paper condition when reading and answering comprehension questions. Fifty-four of the students also answered the MRQ (*Motivations for Reading*

Questionnaire). He found no differences between the two experimental groups in either reading comprehension or reading motivation. In contrast, a very recent meta-analysis examining research from 2000 to 2017 found that digital environments are not always best suited to foster competence development and deep comprehension in the domain of reading (Delgado et al., 2018). Both between- and within-participant designs produced the same results showing advantages of paper over digital reading. This advantage increased in time-constrained reading conditions and with recency of publication year, and was not found in studies using only narrative texts. The authors concluded that providing students with printed reading material might be a more effective way to improve comprehension. Overall, these studies could be interpreted as indicating that the change of reading habits to an increased use of digital devices is related to a lower level of reading comprehension on screen due to particular digital characteristics. However, this does not answer the question whether the change in reading habits toward more digitally based reading is leading to an overall decrease in reading competence.

Another strand of research that might add relevant information on the significance of the reading medium and the potential change in reading literacy levels is mode effect studies in the field of psychometrics. After transitioning from paper- to digital-based assessments, the OECD acknowledged the possibility that small to medium mode effects in PISA 2015 might have led to lower student scores (OECD, 2016). For Germany, Robitzsch et al. (2017) investigated the mode effect due to change in mode and tasks; and Zehner et al. (2018) examined the effect of changing the mode on open-ended tasks. Both research teams found only small effects with computer-based reading tasks being slightly more difficult compared to paper-based ones. These results might indicate that reading and reading digitally are somewhat different constructs. However, international assessments have not just changed the mode but also adapted tasks to take advantage of the new technical possibilities (digital reading). Empirical evidence on the comparability of paper-based reading and reading digitally is still scarce for German elementary school students. Singer and Alexander (2017, p. 2) pointed out that we still have only a limited understanding of how certain attributes of, for example, the learner (as well as the text, or the context) “might interact with the [reading] medium to enhance or inhibit comprehension.” One important student characteristic to consider when investigating reading and comprehension is gender.

2.3 Reading and gender

Gender is a frequently analyzed predictor of reading (Logan & Johnston, 2010; for gender differences and similarities in general, see Hyde, 2014; for gender and educational outcomes, see Steinmayr & Spinath, 2008). Concerning reading literacy, girls had a higher average achievement than boys in 4th-grade in Germany – as in all other countries included in PIRLS 2016 except Macao SAR and Portugal. Girls outperformed boys especially in literary reading literacy (in Germany: 18

points; McElvany et al., 2017). The difference was notably smaller (5 points) for informational reading literacy. Examining gender differences for older students, PISA data showed that the gender gap between 15-year-old girls and boys was even more substantial. Nonetheless, whereas the German difference in reading in PISA was 40 points in 2009 and 44 points in 2012, it fell to 21 points in 2015 (Weis et al., 2016). One possible explanation for these changes and in particular for the decrease in reading is the change in PISA assessment mode that occurred between 2012 and 2015 with the previous paper-based assessments being replaced by computer-based assessments. Moreover, with respect to the younger age group, a decrease in the gender gap was observed in several of the countries participating in ePIRLS (Mullis et al., 2017). For example, Danish students exhibited a significant gender gap in the paper-based PIRLS 2016, but no such difference between boys and girls in ePIRLS. These results are in line with previous research indicating that gender differences in reading depend in part on features of test construction (e.g., Solheim & Lundetrae, 2018).

In addition to the gender gap found in reading literacy, both reading and digital habits have been reported to differ substantially between boys and girls. Gender-specific reading preferences can be observed for both traditional reading materials (e.g., Clark, 2011) and digital reading materials (Duncan et al., 2016). In Duncan et al.'s (2016) study, gender effects favoring female students were evident for both paper-based and digital reading frequency. However, Colley and Comber (2003) found that boys used computer technology more frequently in out-of-school contexts than girls. Concerning the nature of digital activities, Ofcom (2012) reported that girls were more likely to use the internet for homework, visiting social networking and other websites, instant messaging, and Twitter, whereas boys were more likely to use it to play games or watch video clips.

This observation reveals the need to ask which mechanisms impact on potential gender differences in reading digitally. Although there is little research on gender differences in reading digitally, there are many arguments explaining gender differences in reading in general. However, evidence on most of these arguments is rather weak (e.g., cognitive differences, feminization of school, teaching methods used; see Solheim & Lundetrae, 2018), although there is frequent empirical support for the argument that boys and girls differ in their reading motivation (e.g., Schwabe, McElvany, & Trendtel, 2015).

2.4 Predictor of reading literacy: Role of motivation

Reading motivation can be defined as “the drive to read for internal purposes, such as deriving pleasure, attaining personal goals, or satisfying curiosity” (Conradi, Jang, & McKenna, 2014, p. 154). Many empirical studies have reported that girls are significantly more highly motivated to read than boys (elementary school students: Becker & McElvany, 2017; secondary school students: Brozo et al., 2014). Additionally, male students' reading motivation tends to relate more closely to

the level of their reading performance (Solheim & Lundetrae, 2018). McGeown et al. (2012) have provided an important clarification on understanding these differences by showing that students' intrinsic motivation is explained better by gender identity than by sex. The socially perceived femininity of the reading domain was also supported by results of their study indicating that a feminine identity is more closely related than a masculine identity to various aspects of reading motivation.

These mechanisms lead to another important predictor of reading literacy: the reading self-concept. The reading self-concept is "an individual's overall self-perception as a reader, including one's sense of competence and the role ascribed to reading as a part of one's personal identity" (Conradi et al., 2014, p. 154). In contrast to reading motivation, empirical evidence regarding gender differences in reading self-concept is heterogeneous and indicates either no gender gap or an advantage for girls (Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; for Germany, see Becker & McElvany, 2017).

Recently, analyses of the German PIRLS data have reconfirmed the important role of motivational variables in explaining gender differences in reading: After controlling for reading motivation, reading behavior, and reading self-concept, the effect of gender on reading literacy was no longer statistically significant (McElvany et al., 2017).

Gender differences in motivation also exist for reading digitally (see McGeown et al., 2016). In a sample of 791 elementary students, intrinsic and extrinsic motivational facets predicted systematic differences between boys and girls in their frequency of reading both traditional reading materials (i.e., factual and fiction texts) and digital reading sources (i.e., communicative digital texts). However, the authors had used the MRQ, and they themselves stated that this instrument was not perfectly suited to study motivational dynamics with respect to nonbook reading, because it had not been developed to study the motivation for reading, for example, magazines, comics, and digital texts. Hutchison, Woodward, and Colwell (2016) reported multiple gender differences in a study of more than 1,000 4th- and 5th-grade students. For example, male students preferred using the internet over reading a book, whereas female students preferred reading a book. In addition, males assumed that they would learn more from using the internet than from reading a book. Furthermore, more boys than girls considered it to be more difficult to read a book than to watch TV, whereas more girls considered it more difficult to use the internet than to read a book. International ePIRLS results supported the relevance of motivational variables for successful online informational reading. Level of self-efficacy with respect to, for example, using computers was related to average achievement. For 15-year-olds, data from PISA 2009 revealed that online reading engagement, which was conceptualized as including reading motivation as well as a positive attitude toward usage of digital devices, predicted online text reading competence even after controlling for reading competence (Naumann, 2013).

The current mode change also focuses attention on other motivational aspects: boys' and girls' motivation to handle digital devices and also their digital self-concept. One frequently formulated explanation is that boys enjoy digital devices

more; and this, in turn, is assumed to encourage achievement when working on digital tasks (Martin & Binkley, 2009). This claim seems to be supported by studies showing that males have more positive attitudes toward technology and computers than females (see also the previously reported differential use of digital technology). For example, studies have revealed that males like using computers more than females and that their self-concept of their ability to use computers is higher (Colley & Comber, 2003). For the German ICILS 2013 sample, Lorenz et al. (2014) reported gender differences in ICT self-efficacy for advanced computer skills. In a recent meta-analysis, Cai, Fan, and Du (2017) revealed that the gender gap in attitudes toward digital devices persists even though digital devices are part of everyday life in Western societies. But Cai et al. (2017) also reported that both genders exhibit mainly positive attitudes toward digital devices, indicating that test performance across modes might not be a question of liking digital devices. This is in line with Punter, Meelissen, and Glas (2017) who observed that 14-year-old girls outperformed boys in computer and information literacy skills in most countries participating in a recent large-scale survey. This finding contradicts the expectation that the more positive attitudes among males compared to females would lead to better results on computer-based tests, as assumed by, among others, Martin and Binkley (2009).

Most past research on reading, especially on reading by children, has focused on paper-based reading. Furthermore, the motivational dynamics have not yet been studied by including the aspect of motivation toward working with digital devices. Because both the children's environment and the modes of test taking are changing from paper-based to digital, it is necessary to reconsider theoretical assumptions on relevant predictors of reading and to empirically test the relevance of both established and new constructs.

3. Research aims

In light of the increasing importance of digital devices not only in everyday life but also for learning and assessment in schools, the present study focused on the gender gap in reading to be found in digital assessments of reading competencies. In addition, we were interested in whether motivational aspects related (a) to working with digital devices and (b) to reading in general predict reading digitally and whether the motivational characteristics related to digital devices differ between boys and girls. Finally, we examined whether the motivational constructs mediate the relation between gender and reading digitally. In concrete terms, we explored the following research questions:

1. Is there a gender gap in the competence of reading digitally at the end of elementary school?

2. Do motivation and the self-concept with respect to (a) working with digital devices as well as (b) reading in general predict the competence of reading digitally?
3. Do motivation and the self-concept with respect to working with digital devices differ between girls and boys?
4. Do the motivational variables involved in (a) digital devices and (b) reading in general mediate the relation between gender and the competence of reading digitally?

Combining theoretical lines of research on the advantages of girls in the reading domain, we expected to find a gender gap in the competence of reading digitally in favor of girls (Hypothesis 1). Moreover, due to the theoretical relevance of motivation and the self-concept for achievement, we expected motivation and the self-concept with respect to working with digital devices (Hypothesis 2a) as well as motivation and the self-concept with respect to reading in general (Hypothesis 2b) to be statistically significant predictors of the competence of reading digitally after controlling for gender. Whereas gender differences in reading motivation and the reading self-concept have been investigated comprehensively, little is known up to now about potential gender differences in working with digital devices at elementary school age. Due to conflicting theoretical arguments pointing to advantages for girls when reading is highlighted and advantages for boys when digital devices are stressed, we did not formulate any hypothesis for our third research question. Finally, in light of the expected relevance of motivational aspects for reading along with gender differences in the motivational constructs, we expected that motivational variables and, in particular, the aspect of reading motivation would mediate the direct path from gender to reading digitally (Hypothesis 3).

4. Method

4.1 Participants

Data was gathered in 2017 within the research project *Faire und Adaptive Lesekompetenzdiagnose* (FALKE). A sample of 604 fourth-grade students from 37 classrooms at 21 schools took part in the assessment. Six students were excluded from the analysis because they did not take part in the digital assessment of reading competence. The students' average age was 10.29 years ($SD = 0.45$) and 51.0 % were female. Data were collected in a large urban region of Germany. We assessed socioeconomic background with the *Highest International Socio-Economic Index of Occupational Status* (HISEI; Ganzeboom, de Graaf, & Treiman, 1992). The sample was fairly typical for the HISEI in Germany ($M = 50.97$, $SD = 19.87$ compared to the HISEI of a representative sample in 2016 of $M = 50.6$, $SD = 20.7$; see Stanat, Schipolowski, Rjosk, Weirich, & Haag, 2017). Twenty-two percent of

the students reported not speaking solely the language of the reading test at home. Gender groups did not differ with respect to HISEI, percentage of students with a migrant background, or age.

4.2 Instruments

Students worked on digital reading items with single multiple-choice questions on each short text (32 to 110 words each; Schwabe & McElvany, 2013). Due to the multimatrix design, not every student worked on the same booklet. Each student worked on 27 to 30 reading test items within 30 minutes. The items were presented on tablets. Students were asked to select the correct answer among four options. Texts and questions covered factual and nonfactual texts as well as lower- and higher-level reading processes. Responses were used to scale the items in an IRT framework. Item difficulties and person parameters were gathered from a 1PL model using the software package TAM (Robitzsch, Kiefer, & Wu, 2018) implemented in R. Person parameters were scored on a metric with 0 as mean and 1 as standard deviation. EAP reliability of the test was satisfactory ($EAP_{WLE} = .78$). Moreover, the test showed substantial correlations with an established standardized paper-based test of reading competence ($r = .60$; $p < .05$).

Motivational student characteristics in the domains of working with digital devices as well as reading in general were assessed with paper-based questionnaires. The reading motivation and reading self-concept scales were originally developed in the project *LESEN 3-6* conducted in Berlin by the Max Planck Institute for Human Development (see Becker et al., 2010). The items for working on digital devices were specially constructed for this study. Whereas the self-concept items were parallel versions of the reading self-concept measure, the items on motivation for working on digital devices were adapted from a scale used in the BIJU study (*Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter*; Baumert et al., 1997). For ease of presentation, we label the constructs *motivation/self-concept for working on tasks on digital devices* in the following as *digital motivation/self-concept*. The correlation pattern for the domains and constructs supported the validity of the newly constructed scales (see Table 2). Additionally, the digital motivation scale correlated moderately with digital reading behavior ($r = .40$) and the digital self-concept scale correlated highly with digital self-efficacy ($r = .72$), thereby supporting the scales' convergent validity. All scales contained four items to be answered on 4-point scales ranging from 1 (*do not agree*) to 4 (*agree totally*). Descriptive information and examples of items are reported in Table 1. Reliabilities were acceptable to good for all scales.

Table 1: Scales, examples of items, descriptive statistics, and reliabilities

	Examples of items	<i>M (SD)</i>	α
Digital motivation	I enjoy working on digital devices	2.76 (0.73)	.65
Digital self-concept	Working on digital devices is easy for me	2.93 (0.64)	.64
Reading motivation	I like to read	3.28 (0.79)	.77
Reading self-concept	Reading is easy for me	3.17 (0.70)	.82

Note. Original German items translated by the authors.

4.3 Analytical strategy

Data were analyzed with Mplus 7 (Muthén & Muthén, 1998–2012). Missing data, which ranged from 4.0 to 11.2 %, were estimated with the *Full Information Maximum Likelihood* (FIML) method. We applied Type = complex to correct for the nested data structure in all models specified. We decided to perform one-level analyses, because we assumed the underlying processes to be more important on an individual rather than a group level. Indeed, the intraclass correlation coefficients (ICC) of all variables were no higher than .07.

To answer the four research questions, we specified regression (Question 1) and structural equation models (SEMs; Questions 2 to 4). For the motivational variables, we used the answers to the items as indicators of latent constructs, while including reading competence and gender as manifest variables. The mediation model included both indirect paths from gender via motivational variables to reading competence and direct paths from gender and motivational variables to reading competence. Intercorrelations between dependent variables (Question 2), predictors (Question 3), mediators (Question 4), and individual error terms were allowed. All reported coefficients were standardized. The model fit was evaluated on the basis of CFI, RSMEA, and χ^2 .

5. Results

5.1 Preliminary analyses

Whereas mean scores on digital motivation and self-concept ranged between 2.5 and 3, indicating a slightly positive level, expressions of reading-related constructs were generally positive. Bivariate correlations between the constructs of interest are presented in Table 2.

Table 2: Bivariate manifest correlations

	Reading score	Digital motivation	Digital self-concept	Reading motivation	Reading self-concept
Reading score	1.00	–	–	–	–
Digital motivation	-.10*	1.00	–	–	–
Digital self-concept	.08	.33*	1.00	–	–
Reading motivation	.25*	-.09*	.06	1.00	–
Reading self-concept	.27*	-.06	.28*	.43*	1.00

Note. * $p < .05$.

Except for digital self-concept, all motivational variables showed statistically significant bivariate correlations with the reading score. The correlation between digital motivation and the reading score was small and negative. Correlations between the motivational constructs were significant and positive for both reading and digital devices, with the highest correlation being found between reading motivation and reading self-concept. Across domains, self-concepts correlated positively, whereas reading motivation correlated negatively with digital motivation ($r = -.09$). No significant correlations were found between different domains and different motivational constructs.

5.2 Gender differences in reading digitally

To answer the first research question, we specified a regression model with competence of reading digitally as dependent variable and gender as predictor (saturated model: CFI = 1.00, RMSEA = 0.00). Being female predicted reading digitally positively ($\beta = .09$, $SE = .04$, $p < .05$). The size of the standardized regression coefficient was rather small. Comparing the average reading competence of girls with that of boys revealed a medium effect size of Cohen's $d = 0.45$. In sum, data supported the expected gender gap in favor of girls in reading digitally in line with Hypothesis 1.

5.3. Predicting reading digitally by digital and reading-related motivation and self-concept

To examine the second research question, we specified a SEM with reading digitally as dependent variable, motivational aspects of students with respect to (a) digital devices and (b) reading as predictors, and gender as control variable. The model fit was acceptable ($\chi^2 = 237.77$, $df = 102$, $p < .05$, CFI = .94, RMSEA = .05).

Table 3: Path coefficients from SEM predicting digital reading by motivational variables

	β	<i>SE</i>	<i>p</i>
<i>Digital</i>			
Digital motivation	-.13	.10	.24
Digital self-concept	.06	.10	.58
<i>Reading</i>			
Reading motivation	.12	.06	< .05
Reading self-concept	.24	.08	< .05
Female	.02	.04	.60
<i>R</i> ²	.13		

Note. Significant intercorrelations between predictors: digital motivation with digital self-concept: $r = .67^*$; reading motivation with reading self-concept: $r = .58^*$; digital self-concept with reading self-concept: $r = .26^*$.

* $p < .05$.

Motivational variables predicted 12.5 % of the variance in competence of reading digitally in 4th-grade students (see Table 3). Neither digital motivation nor self-concept predicted reading digitally. Reading-related motivation and self-concept were statistically significant predictors of reading digitally. When the motivational constructs were included in the model, gender was no longer a statistically significant predictor of reading digitally. To summarize, the results on students' reading digitally competence were not affected by motivational variables relating to working on digital devices. In contrast, students' performance was predicted in part by reading-related motivational variables.¹ Therefore, Hypothesis 2a was rejected, whereas Hypothesis 2b was supported by the data.

5.4 Gender differences in digital motivation and digital self-concept

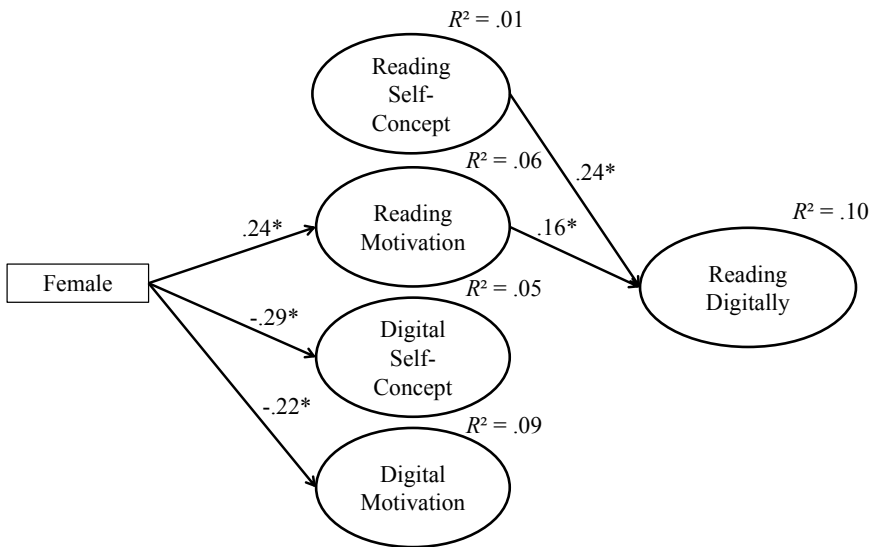
To investigate the third research question, we specified a SEM that included direct paths between gender and motivational aspects of students with respect to digital devices. The model fit was acceptable ($\chi^2 = 51.05$, $df = 16$, $p < .05$, CFI = .95, RMSEA = .06) and revealed significant paths between gender and digital motivation ($\beta = -.27$, $SE = .05$, $p < .05$) and between gender and digital self-concept ($\beta = -.19$, $SE = .04$, $p < .05$). The amounts of explained variance were rather small (digital motivation: $R^2 = .07$; digital self-concept: $R^2 = .03$). Girls possessed less positive motivation and self-concept in the digital domain compared to boys. Therefore, the data indicated a difference between boys and girls in digital motivation and self-concept.

¹ The pattern was robust when gender was not included.

5.5 Mediation of gender differences in reading digitally by motivational variables

To answer the final research question on whether motivational aspects with respect to (a) digital devices and (b) reading mediated the relationship between gender and reading digitally, we specified a SEM. This model included direct paths between reading digitally and both gender and motivational aspects of students with respect to digital devices. In addition, indirect effects of gender on reading digitally were specified via motivational aspects (see Figure 1). The model fit was acceptable ($\chi^2 = 200.00$, $df = 91$, $p < .05$, $CFI = .95$, $RMSEA = .05$).

Figure 1: Mediation model



Note. Indirect path from gender via reading motivation to reading digitally: .04*; Significant correlations between: digital motivation and self-concept: $r = .65$; reading motivation and self-concept: $r = .57$; digital and reading self-concept: $r = .23$.
* $p < .05$

The mediation model revealed that the gender gap in reading digitally was mediated fully by reading motivation. Alongside the statistically significant and substantial direct paths from gender to reading motivation and from reading motivation to digital reading, there was a statistically significant indirect effect (see Figure 1). All other motivational variables did not mediate the gender gap: Gender had no direct effect on students' reading self-concept, which was a positive predictor of reading. Again, gender had a statistically significant negative relation to digital motivational variables, but these had no direct effects on reading digitally. In sum, Hypothesis 3 was supported by the data on reading motivation.

6. Discussion

Reading is a core competence for lifelong learning. However, digitalization has led to fundamental changes in the ways information and texts are delivered (Mangen & van der Weel, 2016; see also political statements on digital media education as a new goal of education in Kultusminister Konferenz (KMK), 2016). This indicates the need for comprehensive knowledge on reading digitally. Because the end of elementary schooling is a crucial time point in reading development, the current study focused on reading digitally in younger students. Students' gender was at the center of the investigation, because, on the one hand, girls are known to be more motivated readers (e.g., McElvany et al, 2017), whereas, on the other hand, boys are often expected to hold more positive attitudes toward digital devices (Cai et al., 2017). We wanted to know how these conflicting aspects would play out empirically when *reading digitally* is measured.

The analyses revealed a statistically significant gender gap in reading digitally for 4th-grade students. Further analyses showed that reading digitally was substantially predicted by reading-related motivation and self-concept but not by digital motivation and self-concept. In addition, gender differences were found in digital motivation and self-concept, with boys possessing more positive values. Finally, the analyses pointed to the mediating role of reading motivation for the gender effect on reading digitally that was not found for any of the other motivational facets.

Thus, our study confirms that the girls' advantage also applies to reading digitally. This gender gap in reading digitally is in line with previous paper-based (e.g., McElvany et al., 2017) or digital (e.g., Mullis et al., 2017) findings for this age group. Our findings add to this body of research by focusing explicitly on the change from paper to tablet without any change in the items themselves. Moreover, our results suggest that gender-related differences in reading do not vanish as a result of digitalization. Quite on the contrary, it is new aspects such as motivation and self-concept in respect to working on digital devices that differ between boys and girls. Our observation of lower levels of digital motivation and self-concept in girls is in line with previous research on older students. For example, results from the ICILS 2013 showed lower self-efficacy on challenging ICT tasks in girls (Lorenz et al., 2014). Moreover, this observation relates to the well-known STEM gender gap (e.g., Liben & Coyle, 2014).

The lack of a relation between either digital motivation or between digital self-concept and reading digitally provides very important information for test developers, test takers, educators, and educational administrators. Our study provides first evidence that digitally based testing of reading competence is already possible in grade 4 without confounding the reading results with construct-irrelevant influences of motivational aspects related to the test mode *digital*. Nonetheless, this pattern of results needs to be replicated in larger student samples. Furthermore, results on the reading-related motivation variables predicting reading digitally can be interpreted on the basis of comprehensive theoretical work and previous em-

irical evidence as supporting the content validity of testing reading digitally. This result differs somewhat from findings on digital reading in which digital reading engagement was found to predict digital reading competence (Naumann, 2013). Hence, a *mode-only* change might not alter the construct of reading.

Finally, the present results are in line with research on the crucial significance of intrinsic reading motivation for reading competence (Becker & McElvany, 2017; Brozo et al., 2016). Our findings contradict recent findings from McGeown et al. (2016), who reported no relation between intrinsic motivation and the amount of digital reading, which might be a proxy for reading competence. The differences in results might stem from the usage of different measures of intrinsic reading motivation. Our broader measure of reading motivation does not explicitly exclude reading digitally or digital reading from the reading construct.

Due to the cross-sectional design of our study, we cannot interpret our findings causally. Moreover, the newly constructed measures of digital motivation and self-concept fail to show very good reliabilities in this young age group. Hence, more research is needed to replicate the findings reported here.

Future studies could follow up on the gender gap and use a parallel administration of the same test items in both modes to investigate how the size of the gender gap reported here relates to the gender gap in a paper-based test (see for older students, Zehner et al., 2018). Future studies could also include reading behavior with both traditional and digital materials. This might shed light on the mechanisms underlying the link between reading motivation and reading digitally. Also more specifically the construct of *digital reading motivation* could be investigated.

Our study raises several issues: We focused on gender as an important individual predictor. Looking at social predictors, especially in Germany and in many other Western countries, social status relates closely to reading competence (e.g., Hußmann et al., 2017). Differential availability and, furthermore, differential use of digital devices as well as familiarity with reading (informational) texts on digital devices might influence coping with demands of reading digitally in educational and testing situations. Therefore, for example, socioeconomic status seems to be an important aspect to consider when investigating reading digitally as well as digital reading.

Our particular interest was in an achievement test of reading digitally. However, the investigation also provides important insights for educational practice. Whereas the gender gap in reading digitally is itself obviously a matter of concern for educators, boys' higher motivation for digital devices might, at the same time, offer a starting point for effective support measures. Because elementary school students currently report that their out-of-school environment is involved much more strongly with digital media than their school learning environment, increasing interesting and effective ways to use digital devices might help boys to further advance their reading digitally (see also Delgado et al., 2018). Thus, more research will be needed on how and when digital devices can be included productively in not only assessment but also learning. For example, a recent study using a randomized controlled trial showed that a computer-assisted reading program such as

Accelerated Reader had a statistically significant positive impact on student reading gains when compared with traditional reading instruction (Shannon, Styers, Wilkerson, & Peery, 2015). At the same time, the generally low student motivation regarding working on digital devices indicates the need for good concepts to prepare elementary school students for the growing digitalization of their educational environment. In particular, the low motivation in girls requires attention and calls for an awareness for gender-sensitive approaches.

The results presented here provide important information on the substantial changes in the context of reading in the 21st century. Because reading is of such fundamental relevance on both the individual and the societal level, it is very important for modern societies to analyse and understand these changes.

References

- Baumert, J., Gruehn, S., Heyn, S., Köller, O., Schnabel, K. U., Leven, I., Plabmeier, N., Schleußer, R., Tauscher, N., & Thebis, F. (1997). *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU): Dokumentation-Band 1. Skalen Längsschnitt I, Wellen 1–4* [Educational careers and psychosocial development in youth: Documentation of scales for longitudinal waves 1 to 4]. Unpublished manuscript, Max-Planck-Institut für Bildungsforschung, Berlin, Germany.
- Becker, M., & McElvany, N. (2017). The interplay of gender and social background: A longitudinal study of interaction effects in reading attitudes and behaviour. *British Journal of Educational Psychology, 88*(4), 529–549.
- Becker, M., McElvany, N., & Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy: A longitudinal study. *Journal of Educational Psychology, 102*, 773–785.
- BitKom Research (2017). *Kinder und Jugend in der digitalen Welt* [Children and youth in the digital world]. Retrieved from <https://www.bitkom.org/Presse/Anhaengen-PIs/2017/05-Mai/170512-Bitkom-PK-Kinder-und-Jugend-2017.pdf>
- Brozo, W. G., Sulkunen, S., Shiel, G., Garbe, C., Pandian, A., & Valtin, R. (2014). Reading, gender, and engagement. Lessons from five PISA countries. *Journal of Adolescent & Adult Literacy, 57*(7), 584–593.
- Cai, Z., Fan, X., & Du, J. (2017). Gender and attitudes toward technology use: A meta-analysis. *Computers & Education, 105*, 1–13.
- Chall, J. (1983). *Stages of reading development*. New York, NY: McGraw Hill.
- Clark, C. (2011). *Setting the baseline: The National Literacy Trust's first annual survey into young people's reading – 2010*. London, England: National Literacy Trust. Retrieved from <https://files.eric.ed.gov/fulltext/ED541400.pdf>
- Colley, A., & Comber, C. (2003). Age and gender differences in computer use and attitudes among secondary school students: What has changed? *Educational Research, 45*(2), 155–165.
- Conradi, K., Jang, B. G., & McKenna, M. C. (2014). Motivation terminology in reading research: A conceptual review. *Educational Psychology Review, 26*(1), 127–164.
- Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review, 25*, 23–38.
- Dillon, A. (1992). Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics, 35*(10), 1297–1326.
- Duncan, L. G., McGeown, S. P., Griffiths, Y. M., Stothard, S. E., & Dobai, A. (2016). Adolescent reading skill and engagement with digital and traditional literacies as

- predictors of reading comprehension. *British Journal of Psychology*, 107(2), 209–238.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 301–330). London, England: Erlbaum.
- Ganzeboom, H. B. G., de Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.
- Goy, M., Valtin, R., & Hußmann, A. (2017). Leseselbstkonzept, Lesemotivation, Leseverhalten und Lesekompetenz [Reading self-concept, motivation, behavior, and competence]. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E. Lankes, N. McElvany, T. C. Stubbe, & R. Valtin (Eds.), *IGLU 2016: Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 143–175). Münster, Germany: Waxmann.
- Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E.-M., McElvany, N., Stubbe, T. C., & Valtin, R. (Eds.) (2017). *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* [PIRLS 2016. Reading literacy of elementary school students in Germany in an international comparison]. Münster, Germany: Waxmann.
- Hutchison, A. C., Woodward, L., & Colwell, J. (2016). What are preadolescent readers doing online? An examination of upper elementary students' reading, writing, and communication in digital spaces. *Reading Research Quarterly*, 51(4), 435–454.
- Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, 65, 373–398.
- Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, 104(1), 53–69.
- IEA – International Association for the Evaluation of Educational Achievement (2015). *PIRLS 2016 Assessment Framework*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.
- Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, 73(2), 509–527.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163–182.
- KMK – Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2016). *Bildung in der digitalen Welt*. [Education in the digital world]. Retrieved from https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2016/Bildung_digitale_Welt_Webversion.pdf
- Liben, L. S., & Coyle, E. F. (2014). Developmental interventions to address the STEM gender gap: Exploring intended and unintended consequences. *Advances in Child Development and Behavior*, 47, 77–115.
- Logan, S., & Johnston, R. (2010). Investigating gender differences in reading. *Educational Review*, 62(2), 175–187.
- Lorenz, R., Gerick, J., Schulz-Zander, R., & Eickelmann, B. (2014). Computer- und informationsbezogene Kompetenzen von Jungen und Mädchen im internationalen Vergleich [Computer- and information-related competencies of boys and girls in an international comparison]. In W. Bos, B. Eickelmann, J. Gerick, F. Goldhammer, K. Schwippert, H. Schaumburg, & M. Senkbeil (Eds.), *ICILS 2013: Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern der 8. Jahrgangsstufe im internationalen Vergleich* (pp. 231–264). Münster, Germany: Waxmann.
- Mangen, A., & van der Weel, A. (2016). The evolution of reading in the age of digitisation: An integrative framework for reading research. *Literacy*, 50(3), 116–124.

- Martin, R., & Binkley, M. (2009). Gender differences in cognitive tests: A consequence of gender-dependent preferences for specific information presentation formats? In F. Scheuermann & J. Björnsson (Eds.), *The transition to computer-based assessment: New approaches to skills assessment and implications for large-scale testing* (pp. 75–82). Ispra, Italy: Joint Research Centre.
- McElvany, N., Kessels, U., Schwabe, F., & Kasper, D. (2017). Geschlecht und Lesekompetenz [Gender and reading competence]. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, N., McElvany, T. C. Stubbe, & R. Valtin (Eds.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 177–194). Münster, Germany: Waxmann.
- McGeown, S., Goodwin, H., Henderson, N., & Wright, P. (2012). Gender differences in reading motivation: Does sex or gender identity provide a better account? *Journal of Research in Reading*, 35(3), 328–336.
- McGeown, S. P., Osborne, C., Warhurst, A., Norgate, R., & Duncan, L. G. (2016). Understanding children's reading activities: Reading motivation, skill and child characteristics as predictors. *Journal of Research in Reading*, 39(1), 109–125.
- McKenna, M. C., Conradi, K., Lawrence, C., Jang, B. G., & Meyer, J. P. (2012). Reading attitudes of middle school students: Results of a U.S. survey. *Reading Research Quarterly*, 47(3), 283–306.
- Medienpädagogischer Forschungsverbund Südwest (2016). *KIM-Studie. Kindheit, Internet, Medien. Basisuntersuchung zum Medienumgang 6–13-Jähriger* [KIM study: childhood, internet, media: Basic study of media usage in 6- to 13-year-olds]. Retrieved from https://www.mpfs.de/fileadmin/files/Studien/KIM/2016 / KIM_2016_Web-PDF.pdf
- Mullis, I. V., Martin, M. O., Foy, P., & Hooper, M. (2017). *ePIRLS 2016: International results in online informational reading*. Retrieved from <https://files.eric.ed.gov/fulltext /ED580351.pdf>
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Naumann, J. (2013). Online-Leseengagement: Vorhersage von Navigationsverhalten und Textverstehen bei Online-Texten [Online reading engagement: Prediction of navigation behavior and text comprehension in online texts]. In C. Rosebrock & A. Bertschi-Kaufmann (Eds.), *Literalität erfassen: Individuell, kulturell, bildungspolitisch* (pp. 188–202). Weinheim, Germany: Juventa.
- Naumann, J., & Sälzer, C. (2017). Digital reading proficiency in German 15-year olds: Evidence from PISA 2012. *Zeitschrift für Erziehungswissenschaft*, 20(4), 585–603.
- Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics*, 51(9), 1352–1375.
- Ofcom (2012). *Children and parents: Media use and attitudes report*. Retrieved from https://www.ofcom.org.uk/___data/assets/pdf_file/0020/56324/main.pdf
- OECD – Organisation for Economic Co-operation and Development (2013). *Technical report of the survey of adult skills (PIAAC)*. Retrieved from https://www.oecd.org/skills/piaac /_Technical%20Report_17OCT13.pdf
- OECD – Organisation for Economic Co-operation and Development (2016). *PISA 2015 results (volume I): Excellence and equity in education*. Paris, France: Organisation for Economic Co-operation and Development Publishing.
- Punter, R. A., Meelissen, M. R., & Glas, C. A. (2017). Gender differences in computer and information literacy: An exploration of the performances of girls and boys in ICILS 2013. *European Educational Research Journal*, 16(6), 762–780.
- Robitzsch, A., Kiefer, T., & Wu, M. (2018). *TAM: Test analysis modules. R package version 2.10-24*. Retrieved from <https://CRAN.R-project.org/package=TAM>

- Robitzsch, A., Lüdtke, O., Köller, O., Kröhne, U., Goldhammer, F., & Heine, J.-H. (2017). Herausforderungen bei der Schätzung von Trends in Schulleistungsstudien [Challenges when estimating trends in educational assessments]. *Diagnostica*, 63(2), 148–165.
- Rosén, M., & Gustafsson, J.-E. (2016). Is computer availability at home causally related to reading achievement in grade 4? A longitudinal difference in differences approach to IEA data from 1991 to 2006. *Large-scale Assessments in Education*, 4, 1–19.
- Schwabe, F., & McElvany, N. (2013). *Development of a computerized adaptive test of reading competence*. Talk at symposia at the 15th Biennial Conference of the European Association of Research on Learning and Instruction (EARLI), Munich, Germany.
- Schwabe, F., McElvany, N., & Trendtel, M. (2015). The school age gender gap in reading achievement: Examining the influences of item format and intrinsic reading motivation. *Reading Research Quarterly*, 50, 1–14.
- Shannon, L. C., Styers, M. K., Wilkerson, S. B., & Peery, E. (2015). Computer-assisted learning in elementary reading: A randomized control trial. *Computers in the Schools*, 32(1), 20–34.
- Singer, L. M., & Alexander, P. A. (2017). Reading across mediums: Effects of reading digital and print texts on comprehension and calibration. *Journal of Experimental Education*, 85(1), 155–172.
- Solheim, O. J., & Lundetræ, K. (2018). Can test construction account for varying gender differences in international reading achievement tests of children, adolescents and young adults? – A study based on Nordic results in PIRLS, PISA and PIAAC. *Assessment in Education: Principles, Policy & Practice*, 25(1), 107–126.
- Stanat, P., Schipolowski, S., Rjosk, C., Weirich, S., & Haag, N. (Eds.) (2017). *IQB-Bildungstrend 2016 – Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* [IQB trends in student achievement 2016 – The second national assessment of German and mathematics proficiencies at the end of 4th grade]. Münster, Germany: Waxmann.
- Steinmayr, R., & Spinath, B. (2008). Sex differences in school achievement: What are the roles of personality and achievement motivation? *European Journal of Personality*, 22(3), 185–209.
- Weis, M., Zehner, F., Sälzer, C., Strohmaier, A., Artelt, C., & Pfof, M. (2016). Lesekompetenzen in PISA 2015: Ergebnisse, Veränderungen und Perspektiven [Reading competence in PISA 2015: Results, changes, and perspectives]. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme, & O. Köller (Eds.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (pp. 249–283). Münster, Germany: Waxmann.
- Wells, C. L. (2012). *Do students using electronic books display different reading comprehension and motivation levels than students using traditional print books?* (Unpublished doctoral dissertation). Liberty University, Lynchburg, VA.
- Wolf, M. (2008). *Proust and the squid: The story and science of the reading brain*. Cambridge, England: Icon Books.
- Zehner, F., Goldhammer, F., Lubaway, E., & Sälzer, C. (2018). Unattended consequences: How text responses alter alongside PISA's mode change from 2012 to 2015. *Education Inquiry*, 9(1), 1–22.

Olaf Köller, Jennifer Meyer, Steffani Saß & Jürgen Baumert

New analyses of an old topic: Effects of intelligence and motivation on academic achievement

Abstract

This study focuses on a topic with a long tradition in educational psychology. In a large data set with several achievement measures we investigated the effects of intelligence and motivation on academic achievement in three domains, namely, German, mathematics, and English, using three different achievement measures (standardized tests, grades, and final written exams) in a sample of upper secondary students ($N = 3\,775$; Grade 13; 54.8 % female; age $M = 19.92$ years) in Germany. Furthermore, we focused on grade point average (GPA) as a general achievement indicator at the end of upper secondary school. First, we aimed to replicate previous results on the predictive power of intelligence and motivation for achievement. Second, we aimed to extend the large body of existing research by adding final written exams – school-based performance tests – as an additional measure. Our findings indicate that motivation had stronger effects on achievement than intelligence did. This was particularly true for the domain-specific achievement measures. Motivation had the strongest effects on grades, followed by final exams. The effects of intelligence were comparatively stronger for standardized achievement tests. Overall, the findings suggest that both intelligence and motivation are important predictors of achievement and that this is true for all kinds of achievement measures.

Keywords

Intelligence; Motivation; Domain-specific achievement; Upper secondary education; Achievement measures

Olaf Köller (corresponding author) · Jennifer Meyer · Steffani Saß, Leibniz Institute for Science and Mathematics Education, Olshausenstraße 62, 24118 Kiel, Germany
e-mail: koeller@ipn.uni-kiel.de
jmeyer@ipn.uni-kiel.de
sass@ipn.uni-kiel.de

Jürgen Baumert, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany
e-mail: sekbaumert@mpib-berlin.mpg.de

Neue Analysen zu einem alten Thema: Effekte von Intelligenz und Motivation auf Schulleistungen

Zusammenfassung

Die vorliegende Arbeit behandelt ein Thema mit langer Tradition in der pädagogischen Psychologie. Wir nutzen einen großen Datensatz mit verschiedenen Leistungsmaßen (Zeugnisnoten, standardisierte Leistungstests und schriftliche Abiturprüfungen), um die Effekte von Intelligenz und Motivation auf Schulleistungen in drei Fächern (Deutsch, Mathematik und Englisch) zu untersuchen. Zusätzlich werden Effekte auf die Abiturdurchschnittsnote analysiert. Die Stichprobe besteht aus 3 775 Schülerinnen und Schülern (54.8 % weiblich; mittleres Alter $M = 19.92$ Jahre) am Ende der 13. Jahrgangsstufe in deutschen Gymnasien. In einem ersten Schritt werden frühere Befunde zur prädiktiven Kraft von Intelligenz und Motivation für Schulleistungen (Tests und Noten) repliziert. Darüber hinaus werden die umfangreichen früheren Befunde dahingehend ergänzt, dass wir schriftliche Abiturleistungen analysieren. Insgesamt zeigen unsere Auswertungen deutlich stärkere Effekte von motivationalen Maßen im Vergleich zu Intelligenzindikatoren. Die stärksten motivationalen Effekte zeigten sich in fachspezifischen Analysen auf Zeugnisnoten, gefolgt von zentralen Abschlussprüfungen. Die Effekte von Intelligenz sind bei standardisierten Schulleistungstests höher. Insgesamt belegen die Befunde aber die Relevanz beider Prädiktoren für die Erklärung von Schulleistungen, und zwar für alle Leistungsindikatoren.

Schlagworte

Intelligenz; Motivation; fachspezifische Schulleistungen; gymnasiale Oberstufe; Schulleistungsmaße

1. Introduction

There is an ongoing debate on the predictive power of intelligence and motivation for academic achievement and learning (for recent studies, see Jansen, Lüdtke, & Schroeders, 2016; Lotz, Schneider, & Sparfeldt, 2018). Despite the fact that both characteristics have been shown to predict academic outcomes, research is still examining the relative contribution of each characteristic regarding its influence on achievement and learning. Previous research has revealed that the type of achievement measure probably moderates the effect sizes of motivation and intelligence, that is, while motivational factors often strongly predict school grades, intelligence seems more powerful in predicting standardized test scores (Helmke, 1992; Lotz et al., 2018; Spinath, Spinath, Harlaar, & Plomin, 2006; Steinmayr & Spinath, 2009). Furthermore, some evidence suggests that the effects of intelligence and motivation differ depending on the domain (e.g., particularly strong effects of intelli-

gence on mathematics achievement; cf. Roth, Becker, Romeyke, Schäfer, Domnick, & Spinath, 2015) and on whether achievement indicators are domain-specific or more general, that is, *grade point average* (GPA) as a global measure of achievement (Poropat, 2009) versus subject- or domain-specific achievement scores. The work presented here aimed to address all these potential moderating effects when investigating the effects of motivation and intelligence on academic achievement. We based the statistical analyses on a large sample of upper secondary students in their final year of schooling in Germany (Leucht, Kampa, & Köller, 2016). In the following, we present previous research on the relationship between intelligence, motivation, and academic achievement.

1.1 Intelligence and achievement

Intelligence is one of the most important factors for explaining individual differences in educational achievement. Many cross-sectional studies have found substantial correlations between intelligence test scores and educational achievement (e.g., Benson, Kranzler & Floyd, 2016; Kaufman, Reynolds, Liu, Kaufman & McGrew, 2012). A recent meta-analysis by Roth et al. (2015) revealed a population correlation of $\rho = .54$ for intelligence and various school grades. The findings were robust (with small moderator effects) across different intelligence measures, different school subjects, different grade levels, and gender. In addition, the correlations between intelligence and standardized achievement test scores in large-scale assessments (LSA) of mathematics achievement ranged from $r = .38$ to $r = .72$ (Saß, Kampa, & Köller, 2017). Lotz et al. (2018) also reported coefficients ranging between .60 and .70 and referred to the study by Baumert, Lüdtke, Trautwein, and Brunner (2009) in which the correlation between intelligence and standardized test scores was even higher. It is generally expected that brighter students learn more easily, are better in dealing with new material, and easily transfer existing knowledge and skills to new learning situations. All of this is intrinsic to intelligence (Jensen, 1998).

Some research, however, also suggests that standardized tests and intelligence tests partly measure the same skill, particularly in mathematics. Saß et al. (2017) pointed out that inductive and deductive reasoning, which is part of intelligence, is also essential to understand mathematical relations, draw conclusions, and apply mathematical knowledge (for similar arguments see Kaufman et al., 2012). The authors thus argue that items in standardized mathematics tests often measure both mathematics achievement and intelligence.

Irrespective of construct validity, many studies have also confirmed the long-term predictive power of intelligence for educational achievement (e.g., Laidra, Pullmann, & Allik, 2007; Schneider & Niklas, 2017). For example, Deary, Strand, Smith, and Fernandes (2007) found that intelligence at age 11 influenced final school exams at age 16 in diverse subjects (e.g., languages, mathematics, sciences, arts, and music).

1.2 Motivation and achievement

The question as to whether motivation predicts students' academic behavior has a very long history in educational psychology. Based on pioneer work by Murray (1938) and Atkinson (1957), numerous studies have investigated the relationship between motivation and achievement. In particular, studies measuring motivation by means of questionnaires found stable positive associations between achievement motivation and academic achievement (cf. Fink, 1962; Gough & Fink, 1964; Schneider & Green, 1977). Subsequent studies used diverse theoretical approaches such as achievement goals (Elliot & McGregor, 2001), intrinsic motivation (Deci & Ryan, 2002), self-efficacy (Kriegbaum, Jansen, & Spinath, 2015), interest (Köller, Schnabel, & Baumert, 2001; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005), or academic self-concept (Guay, Ratelle, Roy, & Litalien, 2010; Lotz et al., 2018; Marsh et al., 2005) to study the relationship between motivational variables and academic achievement.

The *expectancy-value theory* (EVT; Eccles et al., 1983; Wigfield & Eccles, 2000) provides a powerful framework to better understand the role that domain-specific motivational variables play in predicting academic achievement and academic choices. Wigfield and Eccles (2000) argue that expectation of succeeding in a task and the value assigned to the task influence achievement-related behavior. Students who expect to be successful and who value a task highly spend more effort and time working on the task, experience less anxiety, and, in the end, have a higher probability of solving the task. Academic choices, for example, course selection, are also predicted by both components. Students prefer courses for which their self-perceived competence is high and which they value.

In empirical studies, domain-specific self-concept often serves as an indicator of the expectancy component and domain-specific interest represents the value component (cf. Marsh et al., 2005). Both self-concept and interest usually show substantial correlations with achievement. The size of this correlation is moderated by the type of achievement measure; correlations with grades are stronger than with standardized achievement tests. For example, in their meta-analysis, Möller, Pohlmann, Köller, and Marsh (2009) found a correlation of $r = .50$ between mathematics self-concept and grades, while the coefficient was $r = .37$ for test scores. For verbal self-concept, the coefficients were $r = .40$ and $r = .34$, respectively. Marsh, Trautwein, Lüdtke, Köller, and Baumert (2006) found a similar pattern in a large sample (more than 4,000 students from upper secondary schools).

The reciprocal effects model (REM; Marsh & Craven, 2006; Möller, Zimmermann, & Köller, 2014) posits that subsequent achievement is not only affected by prior self-concept, but that subsequent self-concept is also affected by prior achievement. Students' results in academic tests and exams and their grades provide information for social upward and downward comparisons with their classmates. Success often leads to downward comparisons, thereby enhancing academic self-concept, and failure provokes upward comparison, thereby lowering academic self-concept. Achievement is thus not only a result but also a cause of self-con-

cept. Such a REM has also been applied to the relationship between domain-specific interest and achievement (e.g., Köller et al., 2001). The associations found, however, are usually weaker (about $r = .30$, Schiefele, Krapp, & Winteler, 1992) than for self-concept. In a recent study, Jansen et al. (2016) ran secondary analyses of a large German data set comprising the achievement and interest data of more than 39,000 9th-graders from German secondary schools. Five different domains (German, mathematics, biology, chemistry, and physics) were considered. Correlations between interest and grades varied between $r = .27$ (biology) and $r = .38$ (mathematics) with a high degree of domain-specificity; interest and achievement were substantially correlated within one domain but coefficients across domains were small. No clear picture was found for standardized achievement tests; in general, correlations were small or close to zero. Comparing the effects of self-concept and interest on achievement, Lotz et al. (2018) concluded that self-concept seems to be the stronger predictor of achievement and that the effect of interest on achievement equals zero when self-concept is controlled for. Note that, in this context, the correlation between self-concept and interest is usually very high (above $.50$; cf. Marsh et al., 2005) and authors argue that feeling competent (high academic self-concept) increases interest (e.g., Marsh et al., 2005). Interest, on the other hand, seems more important in the case of academic choices (see Köller et al., 2001, for course selection).

Whereas the EVT specifically describes the role that domain-specific motivational factors play in academic achievement, other research approaches have tried to find more general predictors of global achievement measures such as GPA. On the basis of self-concept research, Guay et al. (2010) carried out a repeated measurement study that examined the relationship between academic self-concept and academic achievement in terms of a cumulative achievement measure. The correlations between self-concept and achievement were $r = .60$ (time point 1) and $r = .52$ (time point 2). In line with research on domain-specific self-concept (*reciprocal-effects model* (REM); Marsh & Craven, 2006), it has been argued that this correlation is the result of reciprocal effects between global academic self-concept and global achievement.

In a totally different line of research, the relation between personality factors and global achievement measures has been investigated (see De Raad & Schouwenburg, 1996; Poropat, 2009; Stumm, Hell, & Chamorro-Premuzic, 2011). In particular, conscientiousness has consistently been shown to be associated with academic achievement beyond cognitive ability in secondary education, with effect sizes of $\rho = .23$ (Poropat, 2009). To understand why conscientiousness predicts academic achievement, it can be useful to consider its connection to traits and types of behavior that are known to be crucial for school performance. For example, conscientiousness is closely related to motivational variables such as self-discipline, ambition, persistence, diligence, dutifulness, and grit (Credé, Tynan, & Harms, 2017; Dumfart & Neubauer, 2016; Ivcevic & Brackett, 2014; Schmidt, Fleckenstein, Retelsdorf, Eskreis-Winkler, & Möller, 2017). Further, conscientiousness has been linked to learning behavior that results in good grades (see Credé & Kuncel, 2008;

Kling, Nofle, & Robins, 2013), such as self-regulated learning (Bidjerano & Dai, 2007), goal orientation (Sorić, Penezić, & Burić, 2017), systematic studying and methodical learning styles (e.g., Geisler-Brenstein, Schmeck, & Hetherington, 1996; Komarraju, Karau, Schmeck, & Avdic, 2011), as well as academic effort, which refers to an individual's care and persistence regarding school work (see Trautwein, Lüdtke, Roberts, Schnyder, & Niggli, 2009; Trautwein, Lüdtke, Schnyder, & Niggli, 2006). In summary, previous research indicates that students who are more attentive to their school assignments (i.e., high in conscientiousness) tend to perform better at school (Poropat, 2009).

1.3 Simultaneous effects of intelligence and motivation on achievement

Although theory and research suggest that intelligence and motivation have unique as well as common effects on achievement, a large body of research has tried to find out which of the two factors is more important, that is, which explains more variance in student achievement (see the recent studies by Jansen et al., 2016; Lotz et al., 2018). A relatively old but methodologically sound study was carried out by Helmke (1992). In a sample of lower secondary school students, he found that cognitive variables accounted for 38 % of the variance in a standardized mathematics test, whereas domain-specific self-concept only explained 32 % of the variance. The picture changed for mathematics grades, where the predictive power of self-concept (57 % explained variance) was much higher than for cognitive variables (only 20 %). A similar pattern was reported by Lotz et al. (2018), who additionally had domain-specific interest as a second motivational variable. Interest, however, did not show substantial effects on grades or on test results after controlling for self-concept and intelligence. Jansen et al. (2016) only used domain-specific interest as a motivational measure and also found a pattern of results that was quite similar to Helmke (1992). In a repeated measurement study, Kriegbaum et al. (2015) ran cross-sectional and longitudinal path analyses to predict mathematics achievement in the *Programme for International Student Achievement* (PISA). The authors used different motivational measures and compared their effects on achievement after controlling for intelligence (in cross-sectional models) and prior achievement (in longitudinal models). The strongest motivational predictor in the cross-sectional analyses was task-specific self-efficacy with exactly the same regression weight as intelligence ($\beta = .41$). When prior achievement was controlled for, the effect of intelligence ($\beta = .20$) was slightly stronger than that of self-efficacy ($\beta = .15$).

Concerning global achievement measures, a lot of studies have focused on the effects of personality traits such as conscientiousness on GPA above and beyond intelligence. In a large study, Laidra, Pullmann, and Allik (2007) used personality measures and intelligence to predict GPA in several grade levels. Their results clearly show that intelligence had effects in all grade levels ($\beta > .40$). Furthermore, even after controlling for intelligence, conscientiousness had a substantial and

unique effect on GPA in secondary school ($\beta = .21$). In addition to this work, other studies have combined domain-specific and global personality measures and examined the extent to which they predict academic achievement (cf. Spinath, Freudenthaler, & Neubauer, 2010). They found that domain-specific achievement was more strongly influenced by domain-specific motivational measures than by personality traits (see also Marsh et al., 2006).

Taken together, existing empirical studies clearly suggest that both factors, intelligence and motivation, contribute to academic achievement. This relationship, however, is moderated by the type of achievement measure used and by the domain investigated.

2. The present study

Much research has already been conducted in an attempt to better understand the predictive power of motivation and intelligence for academic achievement. There is broad consensus that both factors uniquely contribute to achievement and that their relative power depends on the domain and the type of achievement measure. Our study contributes to the existing literature. In a first step, we examined the predictive power of intelligence and domain-specific motivation on academic achievement in English, mathematics, and German, contrasting three achievement indicators: standardized achievement tests, final written exams, and school grades. In a second step, we examined the extent to which the GPA at the end of upper secondary school was predicted by intelligence and global motivational and personality measures, that is, academic self-concept and conscientiousness. We decided to use global motivational measures because they represent more general motivational tendencies that influence achievement-related behavior across domains and can therefore be expected to be predictive of such a global measure as GPA, which covers achievement outcomes in all off the different school subjects in upper secondary school.

Our first research goal was to replicate previous findings in a large and rich data set of students at the end of upper secondary school. In contrast to previous studies, the sample was academically positively selected, consisting only of those students prepared for university studies. Our second research goal was to add final written exams as a further achievement measure when examining the effects of intelligence and motivation. Only a small number of studies up until now have investigated the effects of intelligence and motivation on final written exams. These exams at the end of upper secondary education in Germany have important consequences for students as they make up one third of the final GPA and, consequently, are important for college admission. Students' preparation before the exams is crucial and dependent on students' learning motivation and behavior. In our sample of Grade 13, the students had to retrieve content that had been learned over at least two school years. However, in contrast to grades as an accumulated mea-

sure over a longer period of time, the final exam is a performance situation lasting four to six hours for the individual student. It is strongly characterized by a certain amount of pressure as all the knowledge and competencies have to be retrieved on one particular occasion with little room for compensation in the case of failure. Therefore, these final examinations are of particular interest when investigating the effects of intelligence and motivation.

Overall, we tested six hypotheses in a sample of cognitively selected upper secondary students:

- 1) We expected to find stronger effects of intelligence on mathematics achievement than on English and German achievement.
- 2) In the case of motivational variables, we did not have any different expectations for the different domains. However, we expected to find large positive effects of motivation on all domain-specific achievement measures.
- 3) Regarding domain-specific motivational measures, we predicted that academic self-concept would clearly outperform academic interest in predicting achievement in all three domains.
- 4) We further investigated whether the relationships between intelligence, motivation, and achievement differed as a function of the achievement indicator (achievement test scores vs. report card grades vs. final written exams)? We predicted that intelligence would have the strongest positive effects on standardized test scores and would have moderate effects on grades and final written exams.
- 5) For motivation, we expected to find a different pattern, namely, large positive effects on grades and final written exams and moderate effects on test scores.
- 6) Finally, we predicted that intelligence, conscientiousness, and academic self-concept, as a global achievement measure, would have positive effects on the GPA.

3. Method

The present study is based on secondary analyses of the *LISA 6* (Lernergebnisse an beruflichen und allgemeinbildenden Gymnasien in Schleswig-Holstein [Educational Outcomes of Students from Vocational and Academic Upper Secondary Schools]) large-scale assessment study. The study was conducted in the German Federal State of Schleswig-Holstein ($N = 3\,775$; 54.8 % female; see Leucht et al., 2016). In Schleswig-Holstein, there are two school tracks in upper secondary education (Grades 11-13): the vocational and the academic track. The academic track refers to the traditional *Gymnasium*, which provides general preuniversity education, whereas the vocational track focuses on more applied subjects such as technical and economics courses, in addition to compulsory education (e.g., languages and mathematics). In both tracks, students can obtain the general higher education entrance qualification (*Abitur*) after successful completion of classes and final exams

at the end of Grade 13. All participants worked on achievement tests and student questionnaires.

3.1 Procedures

Whereas participation in the questionnaires was voluntary, participation in the achievement tests was mandatory for all students at the randomly drawn academic-track schools ($N = 1433$ students from 17 schools) and at all of the vocational-track schools ($N = 2342$ students from 27 schools). Before commencing the study, written informed consent was obtained from all the participants and their parents. The study was carried out in accordance with the ethical guidelines for research with human participants as proposed by the American Psychological Association (APA). All of the study materials and procedures were approved by the *Ministry of Education, Science and Cultural Affairs of the Federal State of Schleswig-Holstein*. Data collection was managed by the *International Association for the Evaluation of Educational Achievement (IEA)*. Experienced university students who were trained by the IEA administered both tests and questionnaires. Although all students worked on achievement tests, data from the questionnaires were available for only $N = 2234$ students.

3.2 Measures

3.2.1 Academic achievement

Three domain-specific measures plus one global measure of academic achievement were used in the present investigation.

Report card grades. Domain-specific end-of-school-year report card grades (in German, mathematics, and English) were collected via school administration lists for Grade 13. In Germany, report card grades at upper secondary school range from zero to 15 points, with higher values indicating higher achievement.

Standardized tests: Data from standardized tests were available for mathematics and English, but not for German. Mathematics achievement was assessed using a 20-item mathematics test from the *National Educational Panel Study (NEPS)*, which was based on the literacy concept and designed in line with the German educational standards and with international frameworks used in PISA (for details see Neumann et al., 2013; Kampa, Köller, Schmidt, & Leucht, 2016). Following the NEPS framework, mathematical competencies can be described with two dimensions: a) content areas in the field of mathematics (quantity [4 items], change and relationship [6 items], space and shape [3 items], and data and chance [7 items]), and b) the cognitive component of mathematical competence, covering processes related to solving mathematical problems. Five cognitive processes are measured in the test: technical abilities and skills (9 items), modeling (1 item), mathemati-

cal problem solving (4 item), using representational forms (5 item), and mathematical communication (1 item). The results of the *LISA 6* study (see also Kampa et al., 2016) provided evidence for the validity of the test. Techniques from the item response theory (IRT; one-parameter model with a background model), implemented in the computer software ConQuest 3.0 (Adams, Wu, & Wilson, 2012) were applied to compute five plausible values (PVs). The PV reliability was .91.

English achievement was measured with two different tests. The first one with listening and reading comprehension exercises consisted of items from the *German National Assessment* (GNA; e.g., Stanat, Böhme, Schipolowski, & Haag, 2016). The test items were designed to monitor the implementation of educational standards in Germany (see Köller, Knigge, & Tesch, 2010; see also Leucht, Fleckenstein, & Köller, 2016) and therefore represent competencies based on curricula for English language classrooms. Three to four tasks consisting of different items were presented in four 15-minute blocks. Blocks were balanced in difficulty and rotated in eight different booklets to control for position effects and performance decline with test duration (*multimatrix design*). The sufficient reliability and validity of the test has been shown in previous studies; results can be linked to similar standardized tests such as PISA (see Fleckenstein, Leucht, Pant & Köller, 2016). Five PVs were computed for each student to obtain reliable proficiency scores for both skills.

In addition, students' proficiency in English was assessed by a short paper-and-pencil version of the *Test of English as a Foreign Language* (TOEFL), which was developed for the *Institutional Testing Program* (ITP) of the *Educational Testing Service* (ETS) in Princeton (cf. Köller & Trautwein, 2004). The test consists of three sufficiently reliable (internal consistency above .80) subscales, Listening Comprehension (LC); Structure and Written Expression (SWE) and Vocabulary and Reading Comprehension (VRC) which are usually combined to form a global proficiency score. Five PVs were estimated for each subtest and for the global score.

To get a combined English test score from the GNA and the TOEFL, we applied principal component analysis to the five subscales and used the first component (covering more than 70 % of the variance with all loadings above .70) for all further analyses. Note that all other components had Eigenvalues < 1 so that all five subtests were sufficiently represented by the first component. Because we had five PVs per dimension, we computed five global scores that reached a reliability of .90.

Final written exams. We collected information on the grades received in final written exams (*Abitur*) in all three domains via school administration lists. These grades also range from zero to 15 points, with higher values indicating better performance. Centralized *Abitur* tasks were provided by the *Ministry of Education, Science and Cultural Affairs of the Federal State of Schleswig-Holstein* for both school types. In mathematics, competencies in different fields were captured, for example, calculus and geometry, with coherent superordinate assignments consisting of several subtasks (KMK, 2002a). In English and German, assignments consisted of text comprehension (e.g., fictional and nonfictional texts) combined with essay writing tasks (e.g., KMK, 2002b).

Students had to sit final exams in only two of the three subjects. Therefore, not all students took final exams in all three subjects. As a consequence, in English, final exam results were available for $N = 2,984$ students, in mathematics for $N = 3,074$ students, and in German for $N = 3,051$ students.

*GPA*s. Students' GPA scores were collected from school administration lists at the end of upper secondary school. Students' GPAs in German upper secondary schools are calculated on the basis of a weighted combination of score-card grades of the last two school years (66.7 %) and grades in final exams (33.3 %). Scores range between 1 (*excellent*) and 4 (*sufficient*).

3.2.2 Intelligence

General cognitive ability was assessed using the figural (25 items) and verbal (20 items) reasoning subscales of the cognitive ability test (KFT4-12 R+; Heller & Perleth, 2000). Both subscales are indicators of fluid intelligence. The reliabilities of the subscales, as provided in the manual according to the Kuder-Richardson Formula 20, were satisfactory, ranging from $\alpha = .68$ (verbal) to $.81$ (figural). Five PVs were calculated (PV reliability of $.79$) to obtain total scores for the students.

3.2.3 Motivation

Measures of domain-specific motivation, general academic motivation, and personality were used.

Domain-specific measures: We included self-concept and interest measures for the three domains, English, mathematics, and German, based on the EVT. Self-concept was measured by means of four items that have been used in several previous studies (e.g., Möller et al., 2014) and have been shown to be reliable and valid. An example item is "I have always been good in English/mathematics/German". Students responded on a four-point scale ranging from 1 (*totally disagree*) to 4 (*totally agree*). Reliabilities (Cronbach's alpha) reached values of $.83$ (English), $.84$ (German), and $.89$ (mathematics).

Interest was measured by means of four items for English, four items for mathematics, and three items for German. These items had also been used in previous studies (e.g., Marsh et al., 2005) and showed good psychometric properties. An example item is "I would like to have more lessons in English/mathematics/German". Students again responded on a four-point scale ranging from 1 (*totally disagree*) to 4 (*totally agree*). Reliabilities (Cronbach's alpha) reached values of $.77$ (English), $.82$ (German), and $.88$ (mathematics).

Global measures: Four items, again with a four-point response format, were used to assess academic self-concept. The items were taken from Jopt (1978). An example is "I often think that I am not as bright as my classmates". The items reached an internal consistency of $\alpha = .83$.

The four items on conscientiousness were taken from the German short version of the Big Five Inventory (BFI-K; Rammstedt & John, 2005). An example item is “I do tasks thoroughly”. Students had to respond on a five-point scale ranging from 1 (*totally incorrect*) to 5 (*totally correct*). The internal consistency was sufficient (Cronbach’s alpha = .67).

3.2.4 Covariates

Information on gender and school type was collected from school administrations. School type was used as a covariate because previous analyses of the *LISA 6* data have revealed that academic-track students show higher academic achievement as well as higher SES and cognitive ability than vocational-track students (see Leucht & Köller, 2016). Students’ SES was operationalized by the *International Socio-Economic Index of Occupational Status* (ISEI; Ganzeboom, De Graaf, Treiman, & De Leeuw, 1992). Data were collected from both parents, and the highest ISEI in the family (HISEI) was computed as an indicator of SES. Higher HISEI values indicate a high SES.

3.3 Statistical analyses

We ran a series of path analyses using *Mplus* (Version 8.0; Muthén & Muthén, 1998–2017) to examine our research questions. The path analyses were conducted domain-specifically, that is, separately for the three domains and for GPA. Due to the different metrics of the dependent and independent variables, completely standardized solutions were inspected to investigate the predictive power of intelligence and motivation. We interpreted nonoverlapping 95 % confidence intervals as significant differences between the standardized regression coefficients. We decided to run manifest analyses without correcting for measurement errors because most measures had high reliabilities and some measures did not have any reliability information (grades and final exams).

All models were based on maximum likelihood estimation. Because of the hierarchical data structure with students nested in classes and schools, it was necessary to control for dependencies in the data. Thus, we took the data structure into account by computing robust estimates of the model parameters and their standard errors (Type = Complex in *Mplus*; see Muthén & Satorra, 1995) in all models.

Missing data were a serious problem in the present investigation. Table 1 provides missing information for all the measures. Especially the student questionnaire, which was voluntary, showed substantial numbers of missings.¹ We thus decided to apply multiple imputation (MI) techniques to generate a large body of

¹ Note that the voluntary participation of students in questionnaire studies is regulated by law, which often leads to substantial percentages of missing data in such studies. Participation in standardized tests is usually mandatory.

complete data sets. Since full data sets were available for the test measures (PVs), we ran nested imputations (e.g., Harel & Schafer, 2003; Rubin, 2003; Weirich, Haag, Hecht, Böhme, Siegle, & Lüdtke, 2014), that is, we generated 20 full data sets without missings for each of the five PV vectors, resulting in 100 data sets overall that formed the basis for all analyses presented in the Results section. We used *Mplus* to analyze all data sets. *Mplus*, however, only provides combined results for nonnested MI data sets. This can result in slightly biased standard errors of estimated parameters. Therefore, we decided to set a more conservative alpha level of $p < .01$.

Table 1: Descriptive statistics (means, standard deviations, and ranges for nonstandardized variables) and percentage of missing data in the present study

Variable	<i>M</i>	<i>SD</i>	Range	Percentage of missings
<i>Mathematics</i>				
Test	0.00	1.00		-
Grade	8.37	3.27	0 – 15	3.0
Final exam	6.47	3.55	0 – 15	18.6
Self-concept	2.43	0.88	1 – 4	45.1
Interest	2.14	0.88	1 – 4	45.2
<i>English</i>				
Test	0.00	1.00		-
Grade	8.78	2.74	0 – 15	3.2
Final exam	8.15	2.81	0 – 15	21.0
Self-concept	2.63	0.76	1 – 4	44.5
Interest	2.42	0.73	1 – 4	44.9
<i>German</i>				
Grade	8.69	2.51	0 – 15	3.0
Final exam	7.06	2.65	0 – 15	19.2
Self-concept	2.53	0.72	1 – 4	45.5
Interest	2.48	0.82	1 – 4	45.2
<i>General</i>				
Intelligence	0.00	1.00		-
Self-concept	3.17	0.65	1 – 4	45.0
Conscientiousness	3.65	0.68	1 – 5	55.6

4. Results

4.1 Descriptive and correlational findings

Table 1 provides descriptive information on all the variables, including the percentages of missing data. Correlational findings for the different domains and the general measures are presented in Tables 2 and 3. Large correlations were found between the different achievement measures, particularly between report card grades and final written exams. Intelligence had high correlations with standardized test results in English and the highest correlations with test results in mathematics. Domain-specific self-concept and interest had very strong associations, with correlations above .70. However, the correlations between self-concept and achievement variables were stronger than those between interest and achievement variables, thereby supporting findings from previous studies (e.g., Lotz et al., 2018; Marsh et al., 2005).

Concerning the general measures, intelligence, self-concept, and conscientiousness showed more or less the same relationships with GPA. The somewhat smaller coefficient between conscientiousness and GPA is due to the lower reliability of the conscientiousness scale.

Table 2: Correlations among variables; measures for English above the diagonal; for mathematics below the diagonal

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Test (1)		.51	.56	-.07	.37	.18	.46	.47	.33
Grade (2)	.44		.77	.06	.15	.19	.21	.60	.40
Final exam (3)	.52	.70		.02	.22	.19	.27	.56	.34
Sex ¹ (4)	-.45	.03	-.07		-.06	-.07	-.15	.05	.05
School type ² (5)	.33	.13	.15	-.06		.24	.37	.07	.03
HISEI (6)	.14	.10	.13	-.07	.24		.11	.13	.11
Intelligence (7)	.64	.31	.37	-.15	.37	.11		.06	.01
Self-concept (8)	.44	.59	.57	-.15	.05	.01	.29		.71
Interest (9)	.36	.44	.45	-.16	-.01	.00	.23	.83	

Note. ¹Reference is female; ²Reference is academic-track school; analyses were based on 100 nested imputed data sets; combined coefficients; coefficients above .09 and below -.09 are significant ($p < .01$).

Table 3: Correlations among variables; measures for German above the diagonal; general measures below the diagonal

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Grade (1)		.64		.10	.16	.15	.19	.43	.31
Final exam (2)				.09	.16	.12	.21	.38	.21
GPA (3)									
Sex ¹ (4)				-.05	-.06	-.07	-.15	.25	.19
School Type ² (5)				-.14		.24	.37	.04	-.04
HISEI (6)				-.18			.11	.03	.00
Intelligence (7)				-.30				-.06	-.11
Self-concept (8)				-.30	-.19	.03	-.12	.20	.78
Interest (9)									
Conscientiousness (10)				-.24	.24	-.07	-.01	-.06	.18

Note. ¹Reference is female; ²Reference is academic-track school; analyses were based on 100 nested imputed data sets; combined coefficients; coefficients above .09 and below -.09 are significant ($p < .01$).

4.2 Findings of path analyses

Table 4 provides the results of the path analyses that were conducted separately for each domain and for the general measures. The findings represent the combined results from 100 imputed data sets (see Method section).

Table 4: Intelligence and motivation as predictors of achievement (score-card grade, final written exam, and standardized test); standardized regression coefficients/standard errors from path analyses

	Dependent Variable								
	English			Mathematics			German		GPA
	Grade	Exam	Test	Grade	Exam	Test	Grade	Exam	
Intelligence	.17*/.02	.20*/.02	.35*/.02	.14*/.02	.20*/.02	.48*/.02	.19*/.02	.21*/.02	-.26*/.02
Self-concept	.61*/.02	.60*/.02	.40*/.02	.66*/.03	.55*/.03	.24*/.03	.45*/.03	.52*/.03	-.21*/.02
Interest	-.05/.03	-.10*/.03	.04/.03	-.11*/.03	-.04/.03	.00/.03	-.03/.03	-.18*/.03	
Conscientiousness									-.20*/.02
<i>Covariates</i>									
Sex ¹	.06*/.02	.04*/.02	-.03/.02	.13*/.02	.04/.02	-.34*/.01	.04/.02	.04*/.02	-.09*/.02
School-type ²	.03/.02	.09*/.02	.21*/.02	.03/.03	.03/.03	.12*/.02	.05/.03	.04/.03	-.02/.03
HISEI	.10*/.02	.08*/.03	.03/.02	.08*/.02	.10*/.02	.04/.02	.11*/.02	.08*/.02	-.13*/.02
R ²	51.1	39.0	44.4	39.4	38.3	61.2	25.0	22.4	22.3

Note. * $p < .01$; R²: Percentage of explained variance; ¹Reference is female; ²Reference is academic-track school; analyses were based on 100 nested imputed data sets; combined coefficients.

English. Both intelligence and motivational measures had substantial effects on all three measures. However, once we controlled for English self-concept, the effects of interest became close to zero or even negative, which can be explained by multicollinearity (high correlation between self-concept and interest). Intelligence had the largest effect on test scores and effects on grades and written exams were smaller. The picture for self-concept is quite different: it had very large effects on grades and written exams but somewhat smaller effects on test scores. Although the predictive power of intelligence and self-concept was nearly the same for test scores, self-concept clearly outperformed intelligence in the case of grades and written exams. The effects of all the covariates were small: Female students slightly outperformed male students; high SES-students performed a little better than low SES-students and students from academic-track schools outperformed those from vocational-track schools.

Mathematics. To some extent the picture in mathematics corresponds to the one in English with one exception: intelligence outperformed self-concept in predicting standardized tests scores. The substantial effect of gender on test scores indicates that male students outperformed female students. Interestingly, the gender effect for grades indicates that female students (after controlling for all other measures in Table 4) got better grades than male students.

German. Again, self-concept outperformed intelligence but both predictors had substantial effects on grades and written exams. The effects for self-concept were a little bit smaller than in English and mathematics. The negative effect of interest on written exams is again a result of multicollinearity. Regarding the covariates, a small but significant effect was found in favor of high SES-students. Overall, less variance was explained by all of the predictors.

GPA². Intelligence, self-concept, and conscientiousness had effects of similar size. Again, the effects of the covariates were small, with female students having slightly better GPAs than male students and high SES-students getting better GPAs than low SES-students. Congruent to German, the amount of explained variance was smaller than in English and mathematics.

In summary, we found evidence that, in most cases, motivational variables outperformed intelligence in predicting achievement. However, intelligence had additional substantial effects on all achievement measures; however, this may have been due to the fact that the data were collected in a cognitively selected sample with reduced variance in intelligence.

5. Discussion

The major goal of the present study was to expand our knowledge of the differential predictive value of intelligence and motivation for academic achievement in upper secondary schools. We considered domain-specific academic outcomes in

2 Note that high values for GPA indicate low overall achievement (1 = *excellent*; 4 = *poor*).

English, mathematics, and German, as well as GPA to find out whether the pattern of results was moderated by the domain, the type of achievement indicator, and the domain-specificity. Concerning the different domains, our results indicate relatively high consistency in the pattern of findings: Domain-specific self-concept was the dominant predictor for written exams and grades. These regression weights became smaller when looking at standardized tests. Interest, as the second domain-specific motivational predictor, was positively correlated with all achievement measures. However, in the path analyses, the regression weights came close to zero or were even slightly negative. Intelligence showed consistently moderate effects on grades and written exams. In line with previous literature (Lotz et al., 2018), the effects were higher with respect to standardized tests in English and in mathematics. Regarding the prediction of GPA, academic self-concept, conscientiousness, and intelligence showed effects of similar size. Thereby, it is important to note that a personality trait was predictive beyond the effect of academic self-concept. Even though our study was limited to a cross-sectional design and thus did not make it possible to study effects on changes in achievement, it provides some new insights and raises questions about the detailed relationships of the different predictors with academic achievement.

5.1 Predictive power of intelligence

Many previous studies have consistently shown the predictive power of intelligence for academic achievement. Our study replicates these findings. However, the effects were relatively small for grades and written exams. One reason for these lower coefficients could be the selectivity of our sample. As we only sampled students at the end of upper secondary school, more than 50 % of the age cohort was not part of the sample because they left secondary school after Grade 9 or Grade 10 to commence vocational education. Research findings suggest that students who leave school earlier have lower intelligence scores than those who attend upper secondary schools (see the recent study by Guill, Lüdtke, & Köller, 2017, on intelligence differences in different school types in Germany). Therefore, a shrinkage of variance in intelligence may have been the reason for the lower correlations and regression weights.

The effects of intelligence on standardized tests scores were somewhat higher and support the argument that the type of achievement measure moderates the predictive effect of intelligence. Although this is fully in line with previous work (Lotz et al., 2018), we would like to point out, as mentioned above, that some authors (e.g., Kaufman et al., 2012; Saß et al., 2017) have recently argued that academic achievement tests partly measure intelligence. If this is true, one might argue that the higher predictive power of intelligence on standardized tests is merely a consequence of the fact that both tests, at least to some extent, measure the same construct.

5.2 Predictive power of domain-specific self-concept and interest

A recent study by Lotz et al. (2018) revealed that mathematics-specific self-concept had substantial effects on mathematics achievement, but that mathematics-specific interest did not have any effects once self-concept was controlled for. Our results strongly support this finding and confirm the argument by Lotz et al. (2018) that, in the EVT, self-concept is the dominant predictor when achievement is the dependent variable, and that interest seems to be more important when academic choices are predicted. Another explanation for the nonsignificant to negative effects of interest might be that we only used cross-sectional data with performance measures. Domain-specific interest with its intrinsic character might be more important when learning (i.e., change in achievement) is the central outcome in longitudinal studies. Köller et al. (2001) provided support for the assumption that mathematics interest predicts mathematics learning when highly interested students have the chance to opt for advanced courses in which they have more mathematics lessons of higher instructional quality while low-interest students opt for basic courses with a low quantity and quality of instruction.

Compared to many previous studies, we found even higher standardized regression coefficients for domain-specific self-concept. This was particularly true for grades and final exams. Again, we have to mention that our sample was cognitively positively selected, that is, our sample only comprised students from upper secondary schools. As a consequence of this cognitively more homogeneous group, motivation in terms of domain-specific self-concept became a more important predictor of students' academic outcomes. Obviously, high self-concept students put more effort into their final written exams. The strong correlations between grades and self-concept might not only be the result of higher performance but also of more participation and supportive behavior during classroom lessons. There is much evidence (e.g., Kaiser, Retelsdorf, Südkamp, & Möller, 2013) that teachers reinforce such behavior by providing better grades. We are, however, aware that the coefficients of self-concept typically become smaller in longitudinal research when prior grades are included as predictors of subsequent grades (see Limitations section). Finally, our results provide support for our hypothesis that the type of achievement measure moderates the predictive power of motivation, with the finding that the regression weights of self-concept were somewhat lower for standardized tests.

5.3 Prediction of GPA by general measures

Finally, we analyzed the effects of academic self-concept, conscientiousness, and intelligence on GPA. Because GPA represents a global measure of academic achievement, we did not focus on domain-specific motivational factors as predictors. Using self-concept and conscientiousness as global motivational measures, we found evidence for their predictive power above and beyond intelligence. It is noteworthy,

however, that the motivational measures showed effect sizes similar to those for intelligence. In contrast to the domain-specific achievement measures, motivation was not a clearly predominant predictor. These findings suggest that the predominance of motivation disappears once intelligence and motivation are measured at the same aggregation level (aggregated across all domains).

5.4 Limitations

Although our study provides new insights into the relationship between intelligence, motivation, and achievement, there are some limitations that require additional research. First of all, the cross-sectional design did not allow an analysis of the effects on changes in achievement. Much evidence from previous research shows that once prior achievement is controlled for, the effects of all motivational and other cognitive predictors decrease substantially but still remain positive and significant (e.g., Marsh et al., 2005). To inspect such effects in our study, we ran some additional analyses predicting achievement in final written exams in mathematics, German, and English while controlling for report card grades. The effects of domain-specific self-concept and intelligence decreased but remained significant ($p < .01$; standardized path coefficients: mathematics/German/English: self-concept: .199/.279/.210; intelligence: .128/.104/.096), again providing evidence for their predictive power.

A methodological limitation of the present study is that all analyses were based on multiple data sets from nested imputation procedures but that the software package used (*Mplus*) ignores the nested character of all data sets. This could have resulted in somewhat biased estimates of standard errors. Being aware of this problem, we used a more conservative alpha level ($p < .01$). Our impression, however, is that much more research in handling data from the multiple nested imputation of missing values is necessary. We nevertheless feel that our strategy is defensible due to the large amount of missing data from the student questionnaire and the fact that plausible values were available for the standardized tests.

Finally, we restricted our analyses to manifest variables and did not control for measurement error. We are sure that the pattern of results would have been stable even after controlling for measurement error. As the reliabilities of most measures were very high, one would expect slightly higher regression coefficients after controlling for measurement error but not a substantial change in the overall picture of findings.

5.5 Conclusions

A long research tradition has addressed the effects of intelligence and motivation on academic achievement. However, there is still a lack of studies that simultaneously analyze different domains, different achievement outcomes, and different

motivation measures. The current study thus contributes substantially to the literature. The strong effects of domain-specific self-concept on all the different achievement measures highlight the great importance of motivation in explaining differences in achievement at the end of upper secondary school. Although the effects of intelligence were somewhat smaller, our findings also underline the important role that intelligence plays in the school context beyond motivation. This is true, however, for different domains and different achievement measures.

References

- Adams, R. J., Wu, M. L., & Wilson, M. R. (2012). *ACER ConQuest 3.0.1*. Camberwell, England: ACER.
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*, 359-372.
- Baumert, J., Lüdtke, O., Trautwein, U., & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition. Evidence in support of the distinction between intelligence and student achievement. *Educational Research Review*, *4*, 165-176.
- Benson, N. F., Kranzler, J. H., & Floyd, R.G. (2016). Examining the integrity of measurement of cognitive abilities in the prediction of achievement: Comparisons and contrasts across variables from higher-order and bifactor models. *Journal of School Psychology*, *58*, 1-19.
- Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, *17*, 69-81.
- Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, *3*, 425-453.
- Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*, *11*, 492-511.
- Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*, 13-21.
- De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality*, *10*, 303-336.
- Dumfart, B., & Neubauer, A. C. (2016). Conscientiousness is the most powerful non-cognitive predictor of school achievement in adolescents. *Journal of Individual Differences*, *37*, 8-15.
- Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). San Francisco, CA: Freeman.
- Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*, 501-519.
- Fink, M. B. (1962). Objectification of data used in underachievement self-concept studies. *California Journal of Educational Research*, *13*, 105-112.
- Fleckenstein, J., Leucht, M., Pant, H. A., & Köller, O. (2016). Proficient beyond borders: Assessing non-native speakers in a native speakers' framework. *Large-scale Assessments in Education*, *4*(19).

- Ganzeboom, H. B. G., Graaf, P. M. de, & Treiman, D. J., & De Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, *21*, 1–56.
- Geisler-Brenstein, E., Schmeck, R. R., & Hetherington, J. (1996). An individual difference perspective on student diversity. *Higher education*, *31*(1), 73–96.
- Gough, H. G., & Fink, M. B. (1964). Scholastic achievement among students of average ability, as predicted from the California Psychological Inventory. *Psychology in the Schools*, *1*(4), 375–380.
- Guay, F., Ratelle, C. F., Roy, A., & Litalien, D. (2010). Academic self-concept, autonomous academic motivation, and academic achievement: Mediating and additive effects. *Learning and Individual Differences*, *20*, 644–653.
- Guill, K., Lüdtke, O. & Köller, O. (2017). Academic tracking is related to gains in students' intelligence over four years: Evidence from a propensity score matching study. *Learning and Instruction*, *47*, 43–52.
- Harel, O., & Schafer, J. L. (2003). *Multiple imputation in two stages*. Washington DC: Paper presented at the Proceedings of Federal Committee on Statistical Methodology Research Conference.
- Heller, K. A., & Perleth, C. (2000). *KFT 4-12+ R. Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision* [Cognitive ability test for Grade 4]. Göttingen, Germany: Beltz.
- Helmke, A. (1992). *Selbstvertrauen und schulische Leistungen* [Self-confidence and academic achievement]. Göttingen, Germany: Hogrefe.
- Ivcevic, Z., & Brackett, M. (2014). Predicting school success: Comparing conscientiousness, grit, and emotion regulation ability. *Journal of Research in Personality*, *52*, 29–36.
- Jansen, M., Lüdtke, O., & Schroeders, U. (2016). Evidence for a positive relation between interest and achievement: Examining between-person and within-person variation in five domains. *Contemporary Educational Psychology*, *46*, 116–127.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Jopt, U. J. (1978). *Selbstkonzept und Ursachenerklärung in der Schule* [Self-concept and causal attribution in school]. Bochum, Germany: Kamp.
- Kaiser, J., Retelsdorf, J., Südkamp, A., & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction*, *28*, 73–84.
- Kampa, N., Köller, O., Schmidt, F. T. C., & Leucht, M. (2016). Kompetenzen im Fach Mathematik. [Competencies in mathematics]. In M. Leucht, N. Kampa, & O. Köller (Eds.), *Fachleistungen beim Abitur. Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein* (pp. 119–143). Münster, Germany: Waxmann.
- Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive *g* and academic achievement *g* one and the same *g*? An exploration on the Woodcock–Johnson and Kaufman tests. *Intelligence*, *40*, 123–138.
- Kling, K. C., Noffle, E. E., & Robins, R. W. (2013). Why do standardized tests underpredict women's academic performance? The role of conscientiousness. *Social Psychological and Personality Science*, *4*, 600–606.
- KMK – Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany] (2002a). Einheitliche Prüfungsanforderungen in der Abiturprüfung Englisch [Standardized requirements for final examinations in English at the end of upper secondary school] (Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 24.05.2002). Bonn, Germany: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. Retrieved from http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Englisch.pdf
- KMK – Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [Standing Conference of the Ministers of Education and Cultural

- Affairs of the Länder in the Federal Republic of Germany] (2002). *Einheitliche Prüfungsanforderungen in der Abiturprüfung Mathematik* [Standardized requirements for final examinations in mathematics at the end of upper secondary school] (Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 24.05.2002). Bonn, Germany: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. Retrieved from http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Mathe.pdf
- Köller, O., Knigge, M., & Tesch, B. (Eds.). (2010). *Sprachliche Kompetenzen im Ländervergleich* [Verbal competencies in a national comparison among federal states]. Münster, Germany: Waxmann.
- Köller, O., Schnabel, K., & Baumert, J. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *Journal for Research in Mathematics Education*, 32, 448–470.
- Köller, O., & Trautwein, U. (2004). Englischleistungen von Schülerinnen und Schülern an allgemein bildenden und beruflichen Gymnasien [English achievement of students from academic track and vocational upper secondary schools]. In O. Köller, R. Watermann, U. Trautwein, & O. Lüdtke (Eds.), *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien* (pp. 285–326). Opladen, Germany: Leske + Budrich.
- Komaraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, 51(4), 472–477.
- Kriegbaum, K., Jansen, M., & Spinath, B. (2015). Motivation: A predictor of PISA's mathematical competence beyond intelligence and prior achievement. *Learning and Individual Differences*, 43, 140–148.
- Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, 42, 441–451.
- Leucht, M., Fleckenstein, J., & Köller, O. (2016). Erreichen kriterialer Leistungsstandards in der ersten Fremdsprache Englisch. [Reaching proficiency levels in English as a foreign language]. In M. Leucht, N. Kampa, & O. Köller (Eds.), *Fachleistungen beim Abitur. Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein* (pp. 171–199). Münster, Germany: Waxmann.
- Leucht, M., Kampa, N., & Köller, O. (Eds.) (2016). *Fachleistungen beim Abitur. Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein* [Academic achievement at the end of upper secondary school. A comparison of vocational and academic upper secondary schools]. Münster, Germany: Waxmann.
- Leucht, M., & Köller, O. (2016). Anlage und Durchführung der Studie [Design and method of the study]. In M. Leucht, N. Kampa, & O. Köller (Eds.), *Fachleistungen beim Abitur. Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein* (pp. 79 – 98). Münster, Germany: Waxmann.
- Lotz, C., Schneider, R., & Sparfeldt, J. (2018). Differential relevance of intelligence and motivation for grades and competence tests in mathematics. *Learning and Individual Differences*, 65, 30–40.
- Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, 1, 133–163.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, 76, 397–416.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct valida-

- tion and relations to well-being and achievement. *Journal of Personality*, 74, 403–456.
- 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, 1129–1176.
- Möller, J., Zimmermann, F., & Köller, O. (2014). The reciprocal internal/external frame of reference model using grades and test scores. *British Journal of Educational Psychology*, 84, 591–611.
- Murray, H. A. (1938). *Explorations in personality*. Oxford, England: Oxford University Press.
- Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide. Eighth Edition*. Los Angeles, CA: Muthén & Muthén.
- Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, 25, 267–316.
- 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, 80–109.
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135, 322–338.
- Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K) [Short version of the Big Five Inventory (BFI-K): Development and validation of an economic inventory for assessment of the five factors of personality]. *Diagnostica*, 51, 195–206.
- Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, 53, 118–137.
- Rubin, D.B. (2003). Nested multiple imputation of NMES via partially incompatible MCMC. *Statistica Neerlandica*, 57, 3–18.
- Saß, S., Kampa, N., & Köller, O. (2017). The interplay of g and mathematical abilities in large-scale assessments across grades. *Intelligence*, 63, 33–44.
- Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 183–212). Hillsdale, NJ: Erlbaum.
- Schmidt, F. T. C., Fleckenstein, J., Retelsdorf, J., Eskreis-Winkler, L., & Möller, J. (2017). Measuring grit: A German validation and a domain-specific approach to grit. *European Journal of Psychological Assessment*, Advanced Online Publication, 1–12. Retrieved from <https://econtent.hogrefe.com/doi/abs/10.1027/1015-5759/a000407>
- Schneider, F. W., & Green, J. E. (1977). Need for affiliation and sex as moderators of the relationship between need for achievement and academic performance. *Journal of School Psychology*, 15, 269–277.
- Schneider, W., & Niklas, F. (2017). Intelligence and verbal short-term-/working memory: Their interrelationships from childhood to young adulthood and their impact on academic achievement. *Journal of Intelligence*, 5, 26.
- Sorić, I., Penezić, Z., & Burić, I. (2017). The Big Five personality traits, goal orientations, and academic achievement. *Learning and Individual Differences*, 54, 126–134.
- Spinath, B., Spinath, F. M., Harlaar, N., & Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence*, 34, 363–374.
- Spinath, B., Freudenthaler, H. H., & Neubauer, A. C. (2010). Domain-specific school achievement in boys and girls as predicted by intelligence, personality and motivation. *Personality and Individual Differences*, 48, 481–486.

- Stanat, P., Böhme, K., Schipolowski, S. & Haag, N. (Eds.) (2016). *IQB Bildungstrend 2015: Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich*. [IQB Educational Trend 2015: Second national assessment on Language competencies at the end of grade 9]. Münster, Germany: Waxmann.
- Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences, 19*, 80–90.
- Stumm, S., von Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science, 6*, 574–588.
- Trautwein, U., Lüdtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different forces, same consequence: Conscientiousness and competence beliefs are independent predictors of academic effort and achievement. *Journal of Personality and Social Psychology, 97*, 1115–1128.
- Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: Support for a domain-specific, multilevel homework model. *Journal of Educational Psychology, 98*, 438–456.
- Weirich, S., Haag, N., Hecht, M., Böhme, K., Siegle, T., & Lüdtke, O. (2014). Nested multiple imputation in large-scale assessments. *Large-scale Assessments in Education, 2*(9).
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68–81.



Johannes Bellmann,
Hans Merkens (Hrsg.)

Bildungsgerechtigkeit als Versprechen

Zur Rechtfertigung und
Infragestellung eines
mehrdeutigen Konzepts

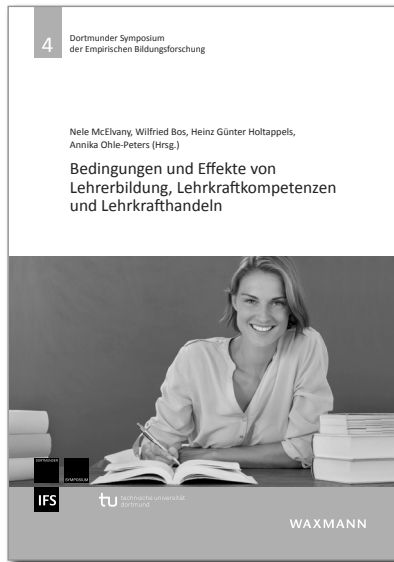
2018, 288 Seiten, br., 34,90 €,
ISBN 978-3-8309-3958-0

E-Book: 30,99 €,
ISBN 978-3-8309-8958-5

Dieser Band vereint grundlagentheoretisch und empirisch orientierte Beiträge, in denen Bildungsgerechtigkeit als ein mehrdimensionales Konzept erkundet wird. Ein gemeinsamer Ausgangspunkt ist die Skepsis gegenüber verbreiteten eindeutigen Auskünften, was denn unter Bildungsgerechtigkeit zu verstehen sei. Unter Bezugnahme auf theoretische Diskussionen und empirische Befunde bemühen sich die Beiträge um eine stärkere Differenzierung der Debatte. Indem gezeigt wird, dass das Versprechen von Bildungsgerechtigkeit alles andere als eindeutig ist, wird zugleich die dilemmatische und aporetische Struktur einer Orientierung an diesem Versprechen ins Blickfeld gerückt. Immer wieder erweist sich Bildungsgerechtigkeit als ein Versprechen, das weder einlösbar noch aufgebbar ist, und gerade deshalb geht es in den Beiträgen sowohl um die Rechtfertigung als auch um die Infragestellung eines mehrdeutigen Konzepts.



UNSERE BUCHEMPFEHLUNG



Nele McElvany, Wilfried Bos,
Heinz Günter Holtappels,
Annika Ohle-Peters
(Hrsg.)

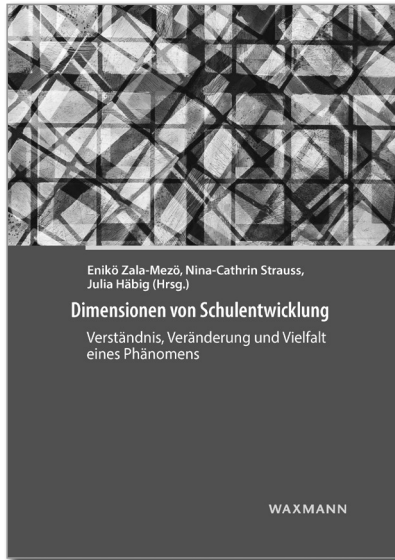
Bedingungen und Effekte von Lehrerbildung, Lehrkraftkompetenzen und Lehrkrafthandeln

*Dortmunder Symposium der
Empirischen Bildungsforschung,
Band 4, 2019, 164 Seiten, br., 29,90 €,
ISBN 978-3-8309-3952-8
E-Book: 26,99 €,
ISBN 978-3-8309-8952-3*

Lehrkräften kommt als Gestalterinnen und Gestaltern von Lern- und Umgebungen eine zentrale Rolle für erfolgreiche Lern- und Bildungsprozesse zu. Dieser vierte Band der Reihe „Dortmunder Symposium der Empirischen Bildungsforschung“ nimmt zunächst die Anforderungen unterschiedlicher Disziplinen der Empirischen Bildungsforschung an die Lehrkraftausbildung in den Blick und diskutiert zu vermittelnde Kompetenzen, die Lehrkräfte benötigen, um erfolgreich unterrichten zu können. Dabei wird auch der Blick auf Lehr-Lernprozesse in inklusiven Klassensettings gerichtet und der Frage nachgegangen, welche Bedeutung theoretische und praktische Anteile der Lehrerbildung zur Kompetenzentwicklung haben. Darüber hinaus werden in diesem Band aktuelle Studien und Forschungsergebnisse zu Bedingungen und Effekten von Lehrkraftkompetenzen präsentiert.



UNSERE BUCHEMPFEHLUNG



Enikő Zala-Mező,
Nina-Cathrin Strauss,
Julia Häbig
(Hrsg.)

Dimensionen von Schulentwicklung Verständnis, Veränderung und Vielfalt eines Phänomens

2018, 240 Seiten, br., 34,90 €,
ISBN 978-3-8309-3835-4

E-Book: 30,99 €,
ISBN 978-3-8309-8835-9

Das Buch versteht sich als eine Auseinandersetzung mit dem Thema Schulentwicklung. Die Beiträge der Autorinnen und Autoren aus der Schweiz, Deutschland und Österreich behandeln verschiedene Aspekte von Schulentwicklung und in den meisten Artikeln werden eigene Forschungsprojekte vorgestellt. Das Themenspektrum reicht dabei von Partizipation der Schülerinnen und Schüler, Förderung von Gesundheit und Inklusion, Schulführung, Gestaltung von Lerngesprächen, Digitalisierung und Schulmanagement bis hin zu regionaler Schulentwicklung. Die Zusammenstellung ergibt einen attraktiven, internationalen Mix aus Themen, Perspektiven, Methoden und Fragestellungen. Diese bewusst gewählte Vielfalt wird von den Herausgeberinnen in einem Synthesekapitel zusammengeführt. Sie vergleichen die Beiträge nach zentralen Merkmalen der Schulentwicklung und schärfen dadurch den Blick auf das Phänomen Schulentwicklung.

