

Wendt, Heike; Hußmann, Anke

Leistungsstark und gerecht? Entwicklungen von Grundschulsystemen im europäischen Vergleich

Tertium comparationis 25 (2019) 1, S. 4-27



Quellenangabe/ Reference:

Wendt, Heike; Hußmann, Anke: Leistungsstark und gerecht? Entwicklungen von Grundschulsystemen im europäischen Vergleich - In: *Tertium comparationis* 25 (2019) 1, S. 4-27 - URN: urn:nbn:de:0111-pedocs-248301 - DOI: 10.25656/01:24830

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-248301>

<https://doi.org/10.25656/01:24830>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

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

Terms of use

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

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

Kontakt / Contact:

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



Leistungsstark und gerecht? Entwicklungen von Grundschulsystemen im europäischen Vergleich

Heike Wendt

Karl-Franzens-Universität Graz

Anke Hußmann

Technische Universität Dortmund

Abstract

With the ET 2020 framework European countries have agreed to strive to increase the quality and equity of their educational systems. This study investigates to which extent countries and especially Germany have succeeded in both. Trends over time in educational quality with regard to achievement levels, heterogeneity, equality, equity, participatory and recognition justice amongst 18 European Union members states were investigated using data from PIRLS 2001 to 2016 and TIMSS 2003 to 2015. Results indicate that only 5 countries show positive developments in both equity and quality.

1. Einleitung

Allgemeiner Auftrag der Grundschule ist die Vermittlung grundlegender Bildung. Dabei Bildungsgerechtigkeit herzustellen und eine hohe Qualität schulischer Bildung zu ermöglichen sind nicht erst seit der Verabschiedung des ‚Strategischen Referenzrahmens für allgemeine und berufliche Bildung‘ (ET 2020) durch die Europäische Union (EU) im Jahr 2009 zwei (der insgesamt vier) zentrale(n) Ziele aller Bildungssysteme in Europa, sondern auch in den ‚Empfehlungen zur Arbeit in der Grundschule‘ (KMK, 2015) zu finden. Beiden Ansprüchen – Qualität schulischer Bildung und Bildungsgerechtigkeit – angemessen nachzukommen ist jedoch nicht leicht. Inwieweit gelingt es den Grundschulen daher, eine Grundlage von Bildung für alle Schülerinnen und Schüler leistungsfördernd *und* gerecht zu gestalten? Diese Frage wird in der Grundschulforschung vielfältig diskutiert (z.B. Heinzel &

Prenzel, 2002; Kucharz, Irion & Reinhoffer, 2011; Wendt, Stubbe & Schwippert, 2015). International vergleichende Schulleistungsuntersuchungen zeigen seit den 1990er-Jahren erhebliche Defizite auf, auf die zwar mit der Umsetzung bildungspolitischer Maßnahmen reagiert wurde, flächendeckend aber bislang keine positiven Auswirkungen zu verzeichnen sind, z.B. hinsichtlich des Ausbaus und der Gestaltung von Ganztagsgrundschulen (z.B. Willems, Wendt, Gröhlich, Walzeburg & Bos, 2014). Seit Jahren ist zu konstatieren, dass es Deutschland im Vergleich zu anderen Ländern kaum gelingt, bestehenden sozialen Ungleichheiten kompensierend zu begegnen (zuletzt Hußmann, Stubbe & Kasper, 2017; Stubbe, Schwippert & Wendt, 2016).

Im Folgenden wird vor diesem Hintergrund systematisch für Deutschland im Vergleich zu anderen Grundschulsystemen Europas untersucht, inwieweit es seit der Jahrtausendwende (und 10 Jahre nach Verabschiedung der gemeinsamen bildungspolitischen Rahmung ET 2020) gelungen ist, den Ansprüchen nach erhöhter Leistungsfähigkeit bei gleichzeitiger Reduktion von Defiziten in Aspekten der Bildungsgerechtigkeit näher zu kommen. Datenbasis bieten die international vergleichenden Schulleistungsstudien *Trends in International Mathematics and Science Study* (TIMSS) und *Progress in International Reading Literacy Study/ Internationale Grundschul-Lese-Untersuchung* (PIRLS/IGLU), die seit 15 Jahren an Grundschulen in Deutschland durchgeführt werden. Einleitend werden die für den Beitrag fokussierten zentralen Begriffe theoretisch gerahmt (Abschnitt 2) und Forschungsbefunde skizziert (Abschnitt 3), anschließend die leitenden Forschungsfragen (Abschnitt 4), die Datenbasis und das methodische Vorgehen (Abschnitt 5) vorgestellt. Gegliedert nach Teilfragen werden die Ergebnisse berichtet und zusammenfassend diskutiert (Abschnitt 6). Der Beitrag schließt mit einem Ausblick (Abschnitt 7).

2. Theoretische Rahmung

Wie kann es gelingen, alle Grundschülerinnen und -schüler unabhängig ihrer Vorleistungen, ihrer sozialen Herkunft oder ihres Geschlechts so zu fördern, dass sie gute kognitive Leistungsergebnisse erzielen? Diese Frage umfasst zwei, auch von der EU im Rahmen der ET 2020 definierten Ziele, die im Fokus des vorliegenden Beitrags als Qualitätsmaße zur Beschreibung von Schulsystemen herangezogen werden:

- (1) die *Leistungsfähigkeit* der Grundschulen, gemessen am Leistungsniveau der Schülerinnen und Schüler am Ende der vierten Jahrgangsstufe, und
- (2) die *Bildungsgerechtigkeit* der Grundschulen, gemessen an Aspekten verschiedener Konzeptionen: Verteilungs-, Teilhabe- und Anerkennungsgerechtigkeit.

Die *Leistungsfähigkeit* von Bildungssystemen gemäß ihrem Auftrag als eine Qualitätskomponente zu wählen, liegt einerseits in der Stärke der hier genutzten Datenbasis begründet. Insbesondere Daten aus *Large-Scale-Assessment*-Studien erfahren „erhöhte ... politische Aufmerksamkeit“ (Baumert, 2016, S. 224) und dienen häufig als Datengrundlage „für die Qualitätsentwicklung des Bildungssystems auf gesamtstaatlicher Ebene“ (ebd.). Leistung gilt andererseits im Vergleich zu vielen anderen Indikatoren als der wichtigste Prädiktor zur Messung von Qualität (Hanushek & Woessmann, 2008). Bezogen auf die gesellschaftliche Wahrnehmung hat der zunehmende Fokus auf Leistung und Leistungsfähigkeit darüber hinaus gewiss mit der Veränderung einer Industrie- hin zu einer Informations- und Wissensgesellschaft zu tun, in Deutschland ebenso wie EU-weit. Daher verwundert es nicht, dass auch die Keimzellen gesellschaftlicher Transformation, die Erziehungs- und Bildungssysteme, hiervon nicht ausgenommen sind. Qualität bemisst sich zunehmend daran, wie effektiv (und effizient) beispielsweise Schule funktioniert (Reynolds et al., 2014).

Die zweite Qualitätskomponente ist die der *Bildungsgerechtigkeit*. Im deutschsprachigen Raum lassen sich aktuell drei Konzeptionen¹ von Bildungsgerechtigkeit unterscheiden, die im Folgenden stark verkürzt beschrieben werden.

Verteilungsgerechtigkeit als Chancengleichheit. Diese Konzeption gilt insbesondere in der empirischen Bildungsforschung als weit verbreitet (z.B. vbw, 2007). Sie wurde mit der Wiederentdeckung sozialer Ungleichheit im deutschen Bildungssystem und der sich daran anschließenden Diskussion um Zugangschancen öffentlich laut: Wie hinreichend belegt, vollzieht sich die Selektion im deutschen Bildungswesen nicht alleine über Leistung, sondern durch „ungleiche Beteiligungs- und Erfolgschancen“ (ebd., S. 12), das heißt Chancen, die je nach sozialer Herkunft höher oder niedriger ausfallen. Von Vertreterinnen und Vertretern dieser Konzeption wird die Frage der gerechten Verteilung entsprechend geführt als Frage der Verteilung gerechter und gleicher Chancen: Alle Schülerinnen und Schüler sollen die gleichen Chancen bekommen, um ihre Potenziale zu entfalten und ihren Potenzialen entsprechend Wege durch das Bildungssystem gehen zu können. Die Ermöglichung gleicher Chancen wird – angelehnt an Rawls' Theorie der Gerechtigkeit (1971) – vor allem den Institutionen zugesprochen, das heißt z.B. die Schaffung chancengleicher Startbedingungen durch einen kompensatorisch wirkenden Ausbau im Bereich der Frühförderung.

Teilhabegerechtigkeit. Giesinger (2007) hingegen fordert Gerechtigkeit als Teilhabe zu fassen. Auch Teilhabegerechtigkeit zielt darauf ab, jegliche Benachteiligungen im Schulsystem auszugleichen, fokussiert dabei aber eine gesellschaftliche Teilhabe im Sinne uneingeschränkter sozialer Beteiligung und rechtlicher Gleichstellung. Im Vordergrund steht die Befähigung des Einzelnen zum Handeln. Damit

orientiert sich Giesinger in seiner Konzeption an Sen, dessen ‚Idee der Gerechtigkeit‘ (2017) gebunden ist an ein Set von Fähigkeiten (bzw. objektiven Verwirklichungschancen, den sogenannten *capabilities*), die Heranwachsende zur umfassenden Teilhabe an der Gesellschaft benötigen. Teilhabegerechtigkeit betont somit die Befähigungen und Chancen, die Schülerinnen und Schüler haben, ihr Leben als ein gutes Leben zu führen.

Anerkennungsgerechtigkeit. Stojanovs (2011) Ansatz nach – orientiert an der Anerkennungstheorie Honneths (2011) – ist Bildungsgerechtigkeit gegeben, wenn Möglichkeiten geschaffen werden, sich autonom zu entwickeln. Damit dies gelingt, braucht es in Schulen ein soziales Miteinander, in dem sowohl pädagogisch Handelnde als auch Schülerinnen und Schüler anerkennenden Umgang leben und erleben. Anders als bei der Verteilungs- und Teilhabegerechtigkeit geht es Stojanov um das Erleben „individueller Autonomie und Freiheit“ (2011, S. 22) und dabei insbesondere um „eine bestimmte Qualität schulischer Sozialbeziehungen“ (ebd., S. 57), die jedes Kind unabhängig seiner Vorleistungen oder sozialen Herkunft erfahren soll, damit seine Potenziale anerkannt werden und sich höchstmöglich entfalten können.

Die Konzeptionen eint das Ziel von Bildungsgerechtigkeit als die Schaffung freier Entfaltung und Teilhabe am gesellschaftlichen Leben für alle Menschen. Im Hinblick auf die Funktionalität von Schulsystemen sprechen sie unterschiedliche Bereiche an: „die Selektions- und Allokationsfunktion“ (Verteilungsgerechtigkeit), „die Qualifikations-, Sozialisations- und Integrationsfunktion“ (Teilhabegerechtigkeit) und „die Personalisationsfunktion“ (Anerkennungsgerechtigkeit) (Wigger, 2015, S. 89). Darüber hinaus zeigt sich in der Zusammenschau, dass Bildungsgerechtigkeit unterschiedlichen Bedingungen unterliegt. Besonders unter Berücksichtigung von Teilhabe- und Anerkennungsgerechtigkeit rücken neben Bedingungen wie der Leistung auch jene in den Fokus, die die Qualität der Einstellungen, Haltungen und Umgangsweisen der beteiligten Personen in Schule betreffen.

3. Forschungsstand

International vergleichende Schulleistungsuntersuchungen liefern Ergebnisse von Bildungsprozessen in ausgewählten Bereichen und machen Verteilungen von Populationen ausgewählter Alters- bzw. Jahrgangsstufen sichtbar. In Berichterstattungen erfolgt in der Regel eine domänenspezifische Betrachtung von Leistungsmittelwerten und -verteilungen sowie eine getrennte Betrachtung von Disparitäten für zentrale Subgruppen im Vergleich von Teilnehmerstaaten über die Zeit. Zusammenhänge zwischen Leistungsfähigkeit und Bildungsgerechtigkeit werden bisher vornehmlich für einzelne Messzeitpunkte im Grundschulbereich (zuletzt Hußmann,

Stubbe et al., 2017; Stubbe et al., 2016) und im Sekundarstufenbereich I (z.B. Woessmann, 2004; Schleicher, 2009; OECD, 2012) analysiert, wobei lediglich Leistungsstände und Verteilungen in den Blick genommen werden.

Eine systematisch längsschnittlich-analytische Betrachtung der Frage, inwieweit es Grundschulsystemen gelingt, ihre Leistungsfähigkeit zu steigern und Bildungsungerechtigkeit über die Zeit zu reduzieren, ist bisher mit Ausnahme zweier Sekundäranalysen auf Basis von TIMSS-Daten nicht erfolgt: Mit dem *20 Years of TIMSS*-Report haben Mullis, Martin und Loveless (2016) Leistungstrends und Verteilungen in 22 Ländern von 1999 bis 2015 betrachtet. Indikator für Bildungsgerechtigkeit waren Leistungstrends für leistungsschwache und leistungsstarke Schülerinnen und Schüler. Für die europäischen Teilnehmer zeigt sich, dass in England, Portugal, Irland und Slowenien insbesondere leistungsschwächere Schülerinnen und Schüler stark von positiven Leistungstrends profitieren. In Ungarn erweisen sich die Ergebnisse hingegen stabil und in Tschechien und den Niederlanden zeigt sich, dass negative Leistungstrends insbesondere auf einen Rückgang an sehr guten Leistungen am oberen Spektrum der Skala zurückzuführen sind. Van Damme und Bellens (2017) nehmen in einer weiteren Sekundäranalyse der TIMSS-Daten einen kürzeren Zeitraum in den Blick (2003–2011). Sie untersuchen die Entwicklungen von 17 Ländern in Bezug auf ihre Leistungsfähigkeit und Facetten der Bildungsgerechtigkeit. Als Indikatoren für Bildungsgerechtigkeit werden die Leistungsdivergenz zwischen leistungsstarken und leistungsschwachen Schülerinnen und Schülern sowie Zusammenhänge der Leistungsindikatoren mit Indikatoren der sozialen Herkunft genutzt. Für keines der in dieser Analyse miteinbezogenen EU-Länder (Fläm. Gemeinschaft in Belgien, England, Ungarn, Italien, Niederlande) ließen sich in Mathematik oder Naturwissenschaften positive Entwicklungen in beiden betrachteten Qualitätsbereichen feststellen. Keine der beiden Untersuchungen ermöglicht vor dem Hintergrund des begrenzten Samples eine Interpretation der Ergebnisse in einem europäischen Referenzrahmen. Auch bieten sie keine Aussagen für Deutschland, weil Deutschland nicht in die Analysen mit einbezogen wurde.

Wie in der nationalen Berichterstattung dokumentiert, zeigen sich in Deutschland im Vergleich von 2001 zu 2016 (IGLU) sowie 2007 zu 2015 (TIMSS) konstante Leistungsmittelwerte, wobei sich für die Leseleistungen von 2006 zu 2016 ein Leistungsrückgang beobachten ließ. Während in TIMSS die Leistungsheterogenität auf einem konstant niedrigen Niveau und die sozialen Disparitäten auf einem konstant hohen Niveau zu konstatieren sind, konnte in IGLU für beide Indikatoren eine signifikante Zunahme beobachtet werden (Hußmann, Wendt et al., 2017; Wendt, Steinmayr & Kasper, 2016). Eine Einordnung entsprechender Entwicklungen in einen europäischen Rahmen bietet zudem ein besonderes analytisches Potenzial (Valtin, 2017). Vor dem Hintergrund der Bedeutung von Lesekompetenzen

am Ende der Grundschulzeit sowie der Tatsache, dass die Förderung von *Literacy* ein besonderes Handlungsfeld der EU darstellt (European Commission, 2012), scheint es gegeben, Fragen nach Leistungsfähigkeit und Bildungsgerechtigkeit studienübergreifend mit Blick auf mathematische, naturwissenschaftliche und lesebezogene Fähigkeiten zu betrachten. Zudem zeigen Analysen (Martin & Mullis, 2013; Wendt & Kasper, 2016), dass zwischen den europäischen Teilnehmerländern unterschiedliche domänenspezifische Stärken und Schwächen zu beobachten sind und Zusammenhänge mit Indikatoren auf Schüler- und Schulebene unterschiedlich ausfallen. Will man Bildungssysteme hinsichtlich ihrer Leistungsfähigkeit und Aspekten von Bildungsgerechtigkeit einschätzen, so ist es daher lohnenswert, auf ein umfassendes Indikatorenspektrum zurückzugreifen.

4. Forschungsfragen

Angesichts der bislang nur vereinzelt und nicht systematischen Zusammenstellung von Befunden zur Leistungsfähigkeit und Bildungsgerechtigkeit von Grundschulsystemen im europäischen Vergleich, ist das Anliegen des Beitrags eine systematische Zusammenstellung von Ergebnissen zu *Leistungstrends* und Veränderungen in Indikatoren der *Bildungsgerechtigkeit* nach unterschiedlichen Konzeptionen, nämlich der Verteilungs-, Teilhabe- und Anerkennungsgerechtigkeit (s. Abschnitt 2). Wie also, so die leitenden Fragen, haben sich die europäischen Bildungssysteme vor dem Hintergrund der EU im Rahmen der ET 2020 definierten Ziele hinsichtlich ihrer Leistungsfähigkeit und Bildungsgerechtigkeit entwickelt, und wie lassen sich diese Entwicklungen im europäischen Referenzrahmen bewerten? Zusammenfassend wird untersucht, welche Länder in beiden Qualitätsbereichen positive Veränderungen verzeichnen können. Vor dem Hintergrund, dass in den meisten Ländern seit der Jahrtausendwende strukturelle und curriculare Veränderungen vorgenommen wurden (Mullis, Martin, Goh & Cotter, 2016; Mullis, Martin, Goh & Prendergast, 2017), ist zu erwarten, dass zumindest zu einigen Indikatoren Veränderungen beobachtet werden können. Aus theoretischer Perspektive kann angenommen werden, dass positive Entwicklungen in beiden Qualitätsbereichen einander unterstützen; Befunde von Mullis, Martin und Loveless (2016) sowie van Damme und Belens (2016) allerdings lassen darauf schließen, dass diese Annahme nicht für alle Länder zutreffend sein wird.

5. Datenbasis und methodisches Vorgehen

5.1 Stichprobe

Die Studien TIMSS und PIRLS/IGLU dienen als Datenbasis. Als Trendstudien konzipiert wurde IGLU bislang vier Mal in den Jahren 2001, 2006, 2011 und 2016

durchgeführt, für TIMSS liegen für den Primarbereich Daten aus vier Erhebungen in 2003, 2007, 2011 und 2015 vor. Beide Studien bieten repräsentative Daten für alle teilnehmenden Staaten. Für den vorliegenden Beitrag wurde, neben Deutschland als zentraler Referenzpunkt, eine Stichprobe von Ländern gezogen, die Mitglied in der EU sind und an mindestens zwei Studienzyklen einer der beiden Schulleistungsstudien teilgenommen haben. Die Analysestichprobe umfasst insgesamt 563.657 Schülerinnen und Schüler in 18 Ländern. Informationen zu den Teilnahmebedingungen und Entwicklungen in den Bildungssystemen der Staaten und Regionen sind umfassend dokumentiert (u.a. Mullis et al., 2017). Bei den Studien handelt es sich um Querschnitterhebungen, die allerdings auf Länderebene ein längsschnittliches Design aufweisen (vgl. Gustaffsson, 2007).

5.2 Instrumente und Variablen²

Als Operationalisierung der *Leistungsfähigkeit* dienen die mittleren Leistungen, die Schülerinnen und Schüler am Ende der vierten Jahrgangsstufe auf den Gesamtskalen Lesen (IGLU) sowie Mathematik und Naturwissenschaften (TIMSS) erzielen. Verwendung finden jeweils die fünf *Plausible Values* (TIMSS-Metrik: $MW_{1995} = 500$ Punkte; $SD = 100$ Punkte). Die Vergleichbarkeit der Ergebnisse für Trendanalysen wird durch Ankeritems und eine gemeinsame Skalierung der Daten sichergestellt (Martin et al., 2016).

Die unterschiedlichen Gerechtigkeitsperspektiven wurden mittels unterschiedlicher Indikatoren untersucht. Leitend für die Auswahl war neben theoretischen Perspektiven vornehmlich die Verfügbarkeit geeigneter Indikatoren mit konstant guten Messeigenschaften für den europäischen Trendvergleich (für eine Übersicht siehe Wendt, Bos, Goy, Jusufi & Kasper, 2017; sowie Hußmann, Wendt, Bos & Rieser, i.V.). Die Konzeption von *Verteilungsgerechtigkeit als Chancengleichheit* wird mittels fünf Indikatoren untersucht. Zwei Indikatoren bilden die *Heterogenität in den Schülerleistungen* ab: (1.) die durchschnittliche Streuung der Werte um den Mittelwert (*Standardabweichung*) und (2.) die *Leistungsdifferenz* zwischen den mittleren Schülerleistungen sehr leistungsschwacher (10. Perzentil) und sehr leistungsstarker Schülerinnen und Schüler (90. Perzentil). Ein dritter Indikator gibt Auskunft über die *Kopplung zwischen sozialer Herkunft und gezeigten Schülerleistungen*. Zur Messung der sozialen Herkunft wird auf die Frage nach der Anzahl der Bücher im Haushalt zurückgegriffen. Sie ist eine der wenigen Indikatoren, die in allen Studienzyklen unverändert im Schülerfragebogen erhoben wurde und konstant gute Messeigenschaften aufweist (vgl. Schwippert, 2019). Die Variable erfasst primär das objektivierte Kulturkapital und liefert einen Hinweis auf das in einer Familie zur Verfügung stehende ökonomische Kapital. Indirekt misst sie das inkorporierte Kulturkapital, da ein Zusammenhang zwischen dem Bildungsniveau

einer Familie und der Anzahl der Bücher, die diese besitzt (Bildungsnähe des Elternhauses), besteht. Die Frage wird mit einer fünfstufigen Ordinalskala erfasst. Für die vorgenommenen Analysen wurden die Angaben in zwei Gruppen ‚maximal 100 Bücher‘ und ‚mehr als 100 Bücher‘ dichotomisiert. Als Indikator für *Verteilungsgerechtigkeit* dient die Höhe der Differenz in den Leistungsmittelwerten beider Gruppen, wobei ein geringer Wert auf eine geringe Kopplung zwischen sozialer Herkunft und gezeigten Leistungen hindeutet.

Ein vierter Indikator gibt Auskunft über Leistungsunterschiede nach *Geschlecht und gezeigten Schülerleistungen*. Auch hier dient die Höhe der Differenz in den Leistungsmittelwerten beider Gruppen als Indikator für Verteilungsgerechtigkeit. Als fünfter Indikator wird die Veränderung in den mittleren Leistungen am 10. und 90. Perzentil im Vergleich zu den mittleren Leistungstrends eines Landes betrachtet. Der Idee von *Chancengleichheit* folgend wäre eine positive Entwicklung zu konstatieren, wenn über die Zeit ein deutlich höherer Zuwachs am unteren Ende des Leistungsspektrums zu beobachten wäre als am oberen Ende.

Als Operationalisierung der Konzeption von *Teilhabegerechtigkeit* werden zunächst die prozentualen Anteile an Schülerinnen und Schülern betrachtet, die auf den jeweiligen Leistungsskalen Werte unterhalb von 475 Punkten (unter Kompetenzstufe III) erzielen. Für diese Gruppe von Kindern ist zu erwarten, dass sie in der Sekundarstufe I mit erheblichen Schwierigkeiten beim Lernen in allen Fächern konfrontiert sein wird, wenn es nicht gelingt, sie maßgeblich zu fördern. Als zweiter Indikator dient eine Skala aus Antworten aus dem Schülerfragebogen in IGLU 2006, 2011 und 2016 zur Lesefreude und zum Leseverhalten. Die Skala umfasst vier Items (z.B. *Ich lese, um Sachen zu erfahren, die ich gerne lernen will.* – 4 Antwortkategorien, vgl. Hußmann et al., i.V.) und zeigt akzeptable Messeigenschaften ($\alpha_{16} = .702$; $\alpha_{11} = .616$, $\alpha_{06} = .601$). Als Indikator für eine geringe *Teilhabegerechtigkeit* werden Anteile von Schülerinnen und Schülern betrachtet, die sich über alle Items hinweg negativ äußern. Der Idee von *Teilhabegerechtigkeit* folgend und der Annahme, dass erworbene Lesekompetenzen eine autonome Lebensgestaltung ermöglichen, wäre eine positive Entwicklung zu konstatieren, wenn sich über die Zeit der Anteil an Schülerinnen und Schülern verringern würde, der von sehr geringer Lesefreude berichtet.

Für die Operationalisierung der Konzeption von *Anerkennungsgerechtigkeit* stehen in beiden Studien kaum geeignete Indikatoren zur Verfügung. Als Annäherung an das Konstrukt werden aus Items, die in IGLU 2011 und 2016 administriert wurden, drei Skalen gebildet, die Aspekte von Anerkennung in der Schule im Sinne von ‚sich wohl fühlen‘ sowie positiver Beziehungen zu Lehrkräften und Mitschülerinnen und Mitschülern beinhalten. Die erste Skala zu ‚sich wohl fühlen‘ umfasst drei Items (z.B. *An dieser Schule habe ich das Gefühl, dazuzugehören.* – 4 Ant-

wortkategorien, vgl. ebd.) und zeigt akzeptable Messeigenschaften ($\alpha_{16} = .678$; $\alpha_{11} = .662$). Die zweite Skala zur Beziehungsqualität zu Lehrkräften umfasst vier Items (z.B. *Mich interessiert, was meine Lehrerin/mein Lehrer sagt.* – 4 Antwortkategorien, vgl. ebd.) mit ebenfalls akzeptablen Messeigenschaften ($\alpha_{16} = .711$; $\alpha_{11} = .677$). Betrachtet werden für beide Skalen Anteile von Schülerinnen und Schülern, die über alle Items negative Wahrnehmungen äußern.

Die dritte Skala zur Beziehungsqualität unter Peers umfasst fünf Items, die sich auf physisches, relationales und verbales Mobbing beziehen (z.B. *Wie oft haben in diesem Jahr andere Schüler deiner Schule eins der folgenden Dinge getan (persönlich oder auch per Kurznachricht (SMS), E-Mail oder über das Internet)? – mich beim Spielen oder anderen Dingen nicht mitmachen lassen,* – 4 Antwortkategorien, vgl. ebd.) mit ebenfalls akzeptablen Messeigenschaften ($\alpha_{16} = .788$; $\alpha_{11} = .757$). Betrachtet werden Anteile von Schülerinnen und Schülern, die davon berichten, regelmäßig – d.h. in allen Kategorien – mindestens einmal monatlich von Mobbing betroffen zu sein. Der Konzeption von *Anerkennungsgerechtigkeit* folgend wäre für alle drei Indikatoren eine positive Entwicklung zu konstatieren, wenn sich über die Zeit der Anteil an Schülerinnen und Schülern verringern würde, der von geringen schulischen Sozialbeziehungen und einem wenig anerkennenden Umgang berichtet.

5.3 Analysestrategie

Alle Ergebnisse wurden mit gängigen deskriptiven Verfahren unter Nutzung des *IDB Analyzers* ausgewertet. Für die Analysen wurden für beide Studien getrennt nach Erhebungszyklus und Land die jeweiligen Kennwerte ermittelt. Hierfür wurden GewichtungsvARIABLEN genutzt und für die Schätzung der Standardfehler die *Jackknife Repeated Replication Technique* angewandt (JRR; Joncas & Foy, 2011).

Als Indikator für eine Veränderung über die Zeit dient die Differenz in den Kennwerten zwischen 2015/2016 und 2001/2003 bzw. 2006/2007. Die Signifikanz von Unterschieden wurde unter Nutzung üblicher Verfahren der Berechnung (Gonzales, 2014) und Konventionen für IGLU und TIMSS (Wendt, Tarelli, Bos, Frey & Vennemann, 2012) unter Annahme einer Irrtumswahrscheinlichkeit von $I = .05$ bewertet.

6. Ergebnisse

6.1 Leistungsfähigkeit

Wie hat sich die *Leistungsfähigkeit* der europäischen Bildungssysteme im Trend verändert? Bei der Betrachtung der Differenzen in den mittleren Leistungen (vgl. Tab. 1) fällt zunächst auf, dass sich die Länder deutlich unterscheiden: Von 18

Ländern, für die Indikatoren für die Lesekompetenz vorliegen, zeigen 8 Länder positive Entwicklungen, 4 Länder negative Entwicklungen und 6 Länder konstante Leistungsmittelwerte.

Tabelle 1: Veränderung in den Leistungsmittelwerten zwischen IGLU 2016 und 2001 bzw. 2006 sowie TIMSS 2015 und 2003 bzw. 2007

Teilnehmer	2001/2003 – 2015/2016			2006/2007 – 2015/2016		
	Lesen	Mathe	Nawi	Lesen	Mathe	Nawi
	M ₁₆ -M ₀₁ (SE)	M ₁₅ -M ₀₃ (SE)	M ₁₅ -M ₀₃ (SE)	M ₁₆ -M ₀₆ (SE)	M ₁₅ -M ₀₇ (SE)	M ₁₅ -M ₀₇ (SE)
Slowenien	41 (2,8)	41 (2,9)	52 (5,0)	21 (3,1)	18 (2,4)	24 (2,9)
Tschechien	6 (3,3)				42 (3,4)	19 (3,4)
Ungarn	11 (3,6)	1 (4,8)	12 (5,7)	3 (4,2)	19 (4,8)	6 (4,8)
Litauen	5 (3,8)	1 (3,3)	16 (3,9)	11 (3,2)	6 (3,5)	13 (3,1)
Lettland	13 (2,7)			17 (3,0)		
Schweden	-6 (3,3)			6 (3,0)	16 (3,9)	15 (4,5)
England	6 (3,9)	15 (4,8)	-4 (5,8)	19 (3,1)	5 (4,4)	-6 (3,9)
Slowakei	17 (4,1)			4 (4,0)	2 (5,6)	-5 (6,0)
Dänemark	1 (2,2)			1 (3,0)	16 (3,6)	10 (3,6)
Spanien				15 (3,2)		
Bulgarien	1 (5,7)			5 (6,5)		
Österreich				2 (2,9)		
Belgien (Franz. Gem.)				-2 (3,4)		
Italien	7 (3,3)	4 (5,0)	1 (5,8)	-3 (3,4)	0 (4,3)	-19 (4,3)
Deutschland	-2 (3,4)			-10 (3,6)	-4 (2,8)	1 (3,2)
Frankreich	-14 (3,5)			-10 (3,1)		
Belgien (Fläm. Gem.)			-7 (4,0)	-22 (2,6)		
Niederlande	-9 (3,1)	-11 (2,6)	-8 (4,6)	-2 (2,4)	-5 (2,6)	-6 (3,5)

Fettdruck: Differenz im Leistungsmittelwert zwischen 2015/2016 und 2001/2003 bzw. 2015/2016 und 2006/2007 statistisch signifikant ($p < .05$).

Für die Bereiche Mathematik und Naturwissenschaften weisen jeweils 6 von 12 Ländern positive Entwicklungen, jedoch auch 2 (Mathematik) bzw. 3 Länder (Naturwissenschaften) signifikant negative Veränderungen in den Leistungsmittelwerten auf. Zusammenfassend lassen sich 4 Ländergruppen unterscheiden: (1.) Länder mit signifikant positiven Zuwächsen in allen drei Kompetenzbereichen (Slowenien, Tschechien), (2.) Länder mit signifikanten Zuwächsen in einzelnen Kompetenzbereichen (Ungarn, Litauen, Lettland, Schweden, England, Slowakei, Dänemark, Spanien), (3.) Länder mit konstanten Leistungsmittelwerten (Bulgarien, Österreich, Franz. Gemeinschaft in Belgien) sowie (4.) Länder mit negativen Trends in einzelnen Leistungsmittelwerten (Niederlande, Fläm. Gemeinschaft in Belgien, Frankreich, Deutschland, Italien). Lediglich für die Niederlande lässt sich im Zeitraum 2001/2003 bis 2015/2016 ein signifikant negativer Trend in allen drei Leistungsindikatoren beobachten.

Tabelle 2: Veränderung in der Heterogenität in den Schülerleistungen zwischen IGLU 2016 und 2001 bzw. 2006 sowie TIMSS 2015 und 2003 bzw. 2007 gemessen an der Veränderung der Standardabweichung und den Leistungsunterschieden zwischen den mittleren Schülerleistungen sehr leistungsschwacher (10. Perzentil) und sehr leistungsstarker Schülerinnen und Schüler (90. Perzentil)

Teilnehmer	2001/2006 – 2016										2003/2007 – 2015														
	Lesen					Mathe					Nawi														
	D ₁₀ ^{16*} (SE)	SD ₁₀ ^{16*} (SE)	D ₉₀ ^{16*} (SE)	SD ₉₀ ^{16*} (SE)	D ₁₀ ^{16*} (SE)	SD ₁₀ ^{16*} (SE)	D ₉₀ ^{16*} (SE)	SD ₉₀ ^{16*} (SE)	D ₁₀ ^{15*} (SE)	SD ₁₀ ^{15*} (SE)	D ₉₀ ^{15*} (SE)	SD ₉₀ ^{15*} (SE)	D ₁₀ ^{16*} (SE)	SD ₁₀ ^{16*} (SE)	D ₉₀ ^{16*} (SE)	SD ₉₀ ^{16*} (SE)									
Italien	-14	(8,6)	-6	(2,0)	-3	(2,0)	-9	(8,6)	-27	(9,0)	-10,5	(2,6)	-13	(8,7)	-5,4	(2,5)	-47	(9,3)	-18,4	(2,3)	-39	(8,3)	-14,9	(2,3)	
England	-22	(9,0)	-7	(2,1)	-8	(2,1)	-22	(7,6)	-11	(9,8)	-3,7	(2,9)	-5	(9,5)	-2,3	(2,7)	-34	(9,6)	-13,5	(2,7)	-24	(8,3)	-10,5	(2,5)	
Slowenien	1	(6,9)	0	(1,9)	1	(1,5)	4	(6,5)	-27	(8,7)	-9,2	(1,9)	-7	(6,5)	-2,7	(1,8)	-20	(8,0)	-7,6	(2,1)	-16	(6,8)	-6,6	(1,9)	
Tschechien	4	(8,0)	3	(2,4)																					
Spanien					-6	(2,2)	-12	(6,8)					-5	(7,5)	-1,6	(1,9)					-17	(8,2)	-6,0	(2,1)	
Dänemark					-2	(1,9)	-4	(7,8)					12	(8,9)	4,3	(2,1)					-18	(9,4)	-7,7	(2,3)	
Frankreich	-5	(7,4)	-1	(2,1)	3	(1,7)	4	(6,5)																	
Belgien (Franz. Gem.)					0	(1,8)	0	(8,2)																	
Osterreich					2	(2,0)	3	(7,8)																	
Litauen	3	(7,6)	1	(1,8)	0	(1,7)	2	(7,1)																	
Bulgarien	9	(11,9)	3	(3,9)	3	(3,6)	4	(12,6)																	
Schweden	5	(7,4)	1	(1,7)	3	(1,8)	9	(7,6)					6	(8,8)	2,6	(2,2)					-3	(9,5)	-0,3	(2,8)	
Slowakei	21	(11,3)	11	(4,0)	7	(4,2)	11	(12,0)					-6	(11,6)	-5,4	(4,5)					2	(11,3)	-2,2	(5,2)	
Deutschland	21	(8,6)	11	(3,4)	11	(3,4)	29	(8,7)					1	(6,8)	-2,8	(1,7)					-18	(7,0)	-9,5	(1,9)	
Litauen	10	(8,5)	5	(2,3)	12	(2,2)	27	(7,5)	-9	(8,8)	-2,6	(2,2)	-11	(9,5)	-4,5	(2,4)	4	(7,1)	2,8	(1,9)	8	(6,7)	4,0	(1,7)	
Belgien (Fläm. Gem.)					5	(1,4)	13	(5,8)	6	(6,0)	1,8	(1,6)					18	(6,5)	6,9	(1,6)					
Niederlande	8	(6,6)	2	(1,8)	7	(1,5)	18	(5,3)	6	(6,2)	1,4	(1,8)	-13	(6,5)	-5,3	(1,7)	19	(7,8)	6,9	(1,7)	1	(7,9)	0,3	(1,9)	
Ungarn	27	(8,2)	9	(2,0)	4	(2,4)	15	(9,1)	25	(10,2)	10,7	(3,1)	-8	(11,1)	-3,2	(3,3)	4	(11,4)	3,4	(3,2)	-6	(12,3)	-2,0	(3,4)	

Fettdruck: Differenz statistisch signifikant ($p < .05$).

Tabelle 3: Veränderung in den Leistungsmittelwerten am 10. und 90. Perzentil zwischen IGLU 2016 und 2001 bzw. 2006 sowie TIMSS 2015 und 2003 bzw. 2007

Teilnehmer	2001/2003 – 2015/2016						2006/2007 – 2015/2016																	
	Lesen			Mathe			Lesen			Mathe			Nawi											
	P10 ₀₁ ⁰¹ (SE)	P90 ₀₁ ⁰¹ (SE)	Δ	P10 ₀₃ ⁰³ (SE)	P90 ₀₃ ⁰³ (SE)	Δ	P10 ₀₆ ¹⁵ (SE)	P90 ₀₆ ¹⁵ (SE)	Δ	P10 ₀₇ ¹⁵ (SE)	P90 ₀₇ ¹⁵ (SE)	Δ	P10 ₀₃ ⁰³ (SE)	P90 ₀₃ ⁰³ (SE)	Δ	P10 ₀₇ ⁰⁷ (SE)	P90 ₀₇ ⁰⁷ (SE)	Δ						
Slowenien	39	(5,5)	40	(4,1)	58	(7,2)	31	(4,9)	64	(6,3)	45	(5,0)	18	(5,1)	21	(3,9)	23	(4,6)	16	(4,6)	34	(5,4)	18	(4,2)
Tschechien	7	(6,8)	10	(4,2)									21	(5,9)	9	(3,3)	45	(6,1)	40	(4,5)	27	(6,2)	10	(5,4)
Spanien													3	(6,7)	-1	(3,9)	9	(7,0)	21	(5,6)	19	(8,4)	1	(4,2)
Dänemark																								
Italien	15	(7,3)	1	(4,6)	19	(6,7)	-8	(5,9)	24	(7,3)	-23	(5,8)	-1	(7,0)	-10	(4,9)	7	(7,4)	-6	(4,4)	0	(6,9)	-39	(4,7)
England	18	(7,4)	-3	(5,1)	22	(7,7)	11	(6,0)	15	(7,6)	-18	(5,8)	32	(5,9)	10	(4,7)	9	(7,4)	4	(6,0)	7	(6,5)	-17	(5,2)
Osterreich													0	(6,4)	3	(4,4)								
Bulgarien	-3	(10,8)	5	(5,0)									2	(11,4)	6	(5,4)								
Litauen	12	(5,8)	15	(5,0)									15	(5,4)	16	(4,6)								
Belgien (Franz. Gem.)													-2	(6,4)	-1	(5,2)								
Frankreich	-13	(5,5)	-18	(4,9)									-13	(4,6)	-10	(4,5)								
Schweden	-11	(5,6)	-6	(4,8)									0	(6,2)	9	(4,4)	11	(7,6)	18	(4,3)	15	(8,4)	12	(4,5)
Litauen	1	(7,1)	10	(4,6)	8	(6,8)	-1	(5,6)	15	(5,9)	19	(3,9)	-2	(6,4)	25	(3,9)	11	(7,4)	0	(5,9)	9	(5,4)	18	(3,9)
Slowakei	3	(10,1)	24	(5,1)									-3	(11,2)	9	(4,4)								
Ungarn	-5	(7,0)	22	(4,2)	-13	(6,7)	11	(5,2)	9	(9,9)	13	(5,7)	-7	(6,2)	8	(4,0)	23	(10,0)	14	(4,8)	8	(10,8)	2	(5,8)
Deutschland Belgien (Fläm. Gem.)	-13	(7,7)	8	(3,8)									-27	(7,8)	2	(3,7)	-4	(5,7)	-3	(3,8)	10	(5,8)	-8	(4,0)
Niederlande	-15	(5,5)	-7	(3,8)	-14	(4,8)	-8	(3,9)	-18	(6,3)	1	(4,6)	-12	(4,4)	6	(2,9)	2	(4,8)	-11	(4,3)	-7	(6,1)	-6	(5,0)

Felddruck: Differenz statistisch signifikant ($p < .05$).

6.2 Verteilungsgerechtigkeit als Chancengleichheit

Wie hat sich die *Heterogenität in den Schülerleistungen* im Trend verändert? Bei Betrachtung der dargestellten Veränderungen in den *Standardabweichungen* (SD) sowie *Leistungsunterschieden* (D) zwischen den mittleren Kompetenzwerten am 10. und 90. Perzentil ist festzustellen (vgl. Tab. 2), dass sich Veränderungen in der Leistungsheterogenität weniger im Bereich Mathematik, mehr aber in Naturwissenschaften und im Lesen beobachten lassen. Für 6 Länder (Italien, England, Slowenien, Tschechien, Spanien, Dänemark) ist in einzelnen Indikatoren eine Reduktion der Leistungsheterogenität zu beobachten. Ein eindeutig positiver Trend besteht für Italien und England. Ein eindeutig negativer Trend ist für Ungarn im Lesen und Mathematik festzustellen. Auch die Slowakei, Litauen und die Flämische Gemeinschaft in Belgien zeigen eine signifikante Zunahme der Heterogenität. Weitere 7 Länder zeigen konstante Ergebnisse. Für Deutschland lässt sich, wie auch für die Niederlande und Dänemark, ein ambivalentes Ergebnis festhalten. Es zeigt sich ein konstantes Ergebnis im Bereich Mathematik, eine signifikante Zunahme an Leistungsheterogenität im Bereich Lesen, jedoch eine signifikante Reduktion im Bereich Naturwissenschaften.

Wie haben sich die *mittleren Leistungen leistungsstarker und leistungsschwacher Kinder* im Vergleich zu den mittleren Leistungstrends eines Landes im Trend verändert? Wie Tabelle 3 zeigt, lassen sich 5 Gruppen von Ländern unterscheiden. (1.) Für Slowenien, Tschechien, Spanien und Dänemark lässt sich im Sinne der Herstellung von Chancengleichheit insofern eine positive Entwicklung beobachten, als dass die durchschnittliche positive Leistungsentwicklung leistungsstarke und leistungsschwache Schülerinnen und Schüler betrifft, wobei leistungsschwache insbesondere von dieser Entwicklung profitieren. (2.) In Italien und England zeigen sich positive Entwicklungen im unteren, jedoch auch negative Trends im oberen Leistungsspektrum. (3.) Gleichförmige bzw. konstante Entwicklungen in positiver oder negativer Hinsicht sind in Österreich, Bulgarien, Lettland, Frankreich sowie der Französischen Gemeinschaft in Belgien zu finden. (4.) In Schweden, Litauen und der Slowakei werden positive Trends insbesondere von den leistungsstärkeren Schülerinnen und Schülern getragen. (5.) Für Deutschland, Ungarn, die Niederlande und die Flämische Gemeinschaft in Belgien ist in der Tendenz ein Schereneffekt zu beobachten.

Welche Veränderungen zeigen sich bezüglich der *Kopplung von sozialer Herkunft und gezeigten Schülerleistungen* im Trend? In Tabelle 4 sind die Veränderungen im Leistungsvorsprung von Kindern mit mehr als 100 Büchern zu Hause vor denen mit weniger als 100 Büchern im Trend dargestellt. In keinem Land lässt sich in dieser Betrachtung eine Reduktion der Leistungsdisparitäten feststellen. 13 der 18 Teilnehmerländer zeigen sogar eine signifikante Vergrößerung. Deutschland

gehört mit Spanien, Frankreich, Dänemark und Bulgarien zu den wenigen Ländern mit konstanten Leistungsdisparitäten.

Tabelle 4: Veränderung im Leistungsvorsprung von Kindern mit mehr als 100 Büchern zu Hause vor denen mit weniger als 100 Büchern zwischen IGLU 2016 und 2001 bzw. 2006 sowie TIMSS 2015 und 2003 bzw. 2007

Teilnehmer	2001/2003 – 2015/2016						2006/2007 – 2015/2016					
	Lesen		Mathe		Nawi		Lesen		Mathe		Nawi	
	Δ_{16-01}	(SE)	Δ_{15-03}	(SE)	Δ_{15-03}	(SE)	Δ_{16-06}	(SE)	Δ_{15-07}	(SE)	Δ_{15-07}	(SE)
Spanien							-4	(3,1)				
Frankreich	-3	(3,1)					2	(2,9)				
Dänemark							2	(3,3)	9	(4,7)	-2	(5,6)
Deutschland	2	(4,1)					6	(4,3)	-2	(3,7)	-7	(4,6)
Bulgarien	7	(6,4)					13	(6,5)				
Niederlande	7	(3,0)	-2	(3,0)	3	(3,8)	0	(2,2)	1	(4,1)	-6	(5,2)
Schweden	7	(3,6)					9	(3,6)	11	(4,1)	8	(5,4)
Belgien (Franz. Gem.)							10	(4,0)				
Österreich							14	(3,3)				
England	11	(3,8)	3	(5,0)	3	(4,3)	3	(3,3)	-1	(5,3)	1	(5,1)
Tschechien	15	(3,3)							6	(3,9)	7	(4,8)
Lettland	12	(3,1)										
Slowenien	12	(2,9)	8	(3,2)	7	(3,7)	6	(3,2)	-1	(3,9)	-2	(4,3)
Ungarn	20	(3,7)	24	(4,5)	21	(4,5)	7	(4,4)	8	(5,0)	4	(5,4)
Italien	2	(3,3)	20	(4,7)	13	(4,9)	1	(3,9)	14	(4,7)	6	(5,2)
Belgien (Fläm. Gem.)			12	(2,9)	18	(3,0)	8	(3,1)				
Litauen	9	(3,8)	10	(3,8)	19	(3,8)	11	(3,2)	17	(4,6)	23	(5,7)
Slowakei	21	(4,7)					10	(4,8)	23	(4,5)	27	(5,2)

Fettdruck: Leistungsvorsprung in 2015/2016 signifikant ($p < .05$) höher als in 2001/2003 bzw. 2006/2007. Abweichungen zu Hußmann, Stubbe et al. (2017) ergeben sich, da hier Schülerangaben genutzt werden.

Welche Veränderungen zeigen sich bezüglich des *Zusammenhangs von Geschlecht und gezeigten Schülerleistungen* im Trend? Bis auf wenige Ausnahmen zeigen sich in keinem Land signifikante Veränderungen (ohne Tab., zu Kennwerten vgl. Wendt, Steinmayr & Kasper, 2016; McElvany, Kessels, Schwabe & Kasper, 2017). Festzustellen aber ist, dass sich mit Ausnahme der Niederlande und Deutschland in allen anderen europäischen Teilnehmerstaaten in keinem Leistungsindikator signifikante Veränderungen in den geschlechtsspezifischen Disparitäten beobachten lassen. In den Niederlanden ist lediglich im Bereich der Naturwissenschaften zwischen 2007 und 2015 eine auf den ersten Blick signifikant positive Reduktion der Disparität zu beobachten, die jedoch nicht auf eine Leistungssteigerung der Mädchen, sondern eine geringere Leistung der Jungen zurückzuführen ist. Eine vergleichbare Entwicklung zeigt sich für die Mathematik in Deutschland, wobei die Leistungsdisparitäten im Lesen konstant geblieben sind. Lediglich in den Naturwis-

senschaften ist für Deutschland eine positive Entwicklung festzustellen: Mädchen können zwischen 2007 und 2015 ihre Leistungen verbessern.

6.3 Teilhabegerechtigkeit

Welche Veränderungen zeigen sich in Bezug auf den Anteil an Schülerinnen und Schülern mit Leistungen auf den Kompetenzstufen I und II im Trend? Tabelle 5 zeigt, dass für die meisten Länder (14 von 18) eine positive Entwicklung im Sinne einer signifikanten Reduktion des Anteils an leistungsschwachen Schülerinnen und Schülern in mindestens einer Domäne zu beobachten ist. Deutschland gehört mit den Niederlanden, Frankreich und der Flämischen Gemeinschaft in Belgien zu den wenigen Ländern mit einer negativen Entwicklungstendenz im Sinne eines eher steigenden Anteils, insbesondere an leseschwachen Schülerinnen und Schülern. Eine Entwicklung, die einer genaueren Untersuchung bedarf.

Tabelle 5: Veränderung im Anteil von Schülerinnen und Schülern mit Leistungen auf den Kompetenzstufen I und II zwischen IGLU 2016 und 2001 bzw. 2006 sowie TIMSS 2015 und 2003 bzw. 2007

Teilnehmer	2001/2003 – 2015/2016						2006/2007 – 2015/2016					
	Lesen		Mathe		Nawi		Lesen		Mathe		Nawi	
	Δ_{16-01}	(SE)	Δ_{15-03}	(SE)	Δ_{15-03}	(SE)	Δ_{16-06}	(SE)	Δ_{15-07}	(SE)	Δ_{15-07}	(SE)
Slowenien	-15,9	(1,5)	-20,8	(1,9)	-23,5	(1,8)	-7,3	(1,5)	-8,8	(1,5)	-10,5	(1,4)
Spanien							-7,9	(1,7)				
Slowakei	-4,8	(1,9)					-0,9	(1,9)	-1,9	(2,3)	1,3	(2,2)
Italien	-4,3	(1,6)	-4,0	(2,2)	-4,6	(2,5)	0,1	(1,6)	-1,4	(2,1)	3,6	(2,2)
Lettland	-3,3	(1,3)					-4,2	(1,4)				
England	-3,2	(1,5)	-5,4	(2,0)	-2,2	(2,0)	-7,1	(1,4)	-1,1	(1,7)	-0,6	(1,6)
Tschechien	-2,0	(1,5)							-19,9	(1,9)	-9,3	(1,9)
Litauen	-0,7	(1,6)	-1,7	(1,7)	-5,4	(1,7)	0,0	(1,4)	-3,3	(1,8)	-4,2	(1,9)
Österreich							-0,5	(1,6)				
Dänemark							-0,4	(1,4)	-4,0	(1,8)	-6,0	(2,0)
Ungarn	0,0	(1,5)	1,1	(2,1)	-4,2	(2,1)	1,2	(1,7)	-7,8	(2,2)	-2,2	(2,2)
Bulgarien	0,4	(2,2)					-0,4	(2,4)				
Belgien (Franz. Gem.)	1,0	(2,1)					1,0	(2,1)				
Schweden	2,5	(1,3)					-0,1	(1,3)	-7,3	(2,2)	-6,4	(2,2)
Deutschland	2,0	(1,6)					5,7	(1,6)	1,7	(1,6)	-2,1	(1,8)
Niederlande	4,0	(1,4)	5,7	(1,5)	7,0	(1,8)	3,3	(1,1)	0,8	(1,6)	3,1	(1,9)
Frankreich	5,5	(1,7)					4,4	(1,6)				
Belgien (Fläm. Gem.)			2,1	(1,2)	6,0	(1,7)	9,8	(1,6)				

Fettdruck: Differenz im Anteil zwischen 2015/2016 und 2001/2003 bzw. 2015/2016 und 2006/2007 statistisch signifikant ($p < .05$).

Welche Veränderungen zeigen sich in Bezug auf die Lesefreude und die damit verbundene Annahme, dass erworbene Lesekompetenzen eine autonome Lebensgestaltung ermöglichen? Tabelle 6 ist zu entnehmen, dass sich drei Gruppen von Ländern unterscheiden lassen: (1.) In England, den Niederlanden, Österreich, Spanien, Tschechien und Italien lässt sich eine signifikante Reduktion an Kindern mit geringer Lesefreude beobachten. Während sich (2.) in Frankreich, Ungarn, der Flämischen Gemeinschaft in Belgien und Bulgarien keine signifikanten Veränderungen zeigen, lassen sich (3.) in den weiteren 8 Ländern, darunter auch Deutschland, kleine signifikante Zuwächse an Schülerinnen und Schülern beobachten, die von geringer Lesefreude berichten und von denen anzunehmen ist, dass sie ihre erworbenen Lesekompetenzen nicht für eine autonome Lebensgestaltung nutzen können.

Tabelle 6: Veränderung im Anteil von Schülerinnen und Schülern mit geringer Lesefreude zwischen IGLU 2016 und 2001 bzw. 2006

Teilnehmer	2001/2006 – 2016									
	2001	(SE)	2006	(SE)	2016	(SE)	Δ_{16-06}	(SE)	Δ_{16-01}	(SE)
England	40,9	(1,4)	44,2	(1,4)	29,2	(1,0)	-15,0	(1,7)	-11,7	(1,8)
Niederlande	49,8	(1,1)	52,2	(1,2)	44,4	(1,3)	-7,8	(1,7)	-5,4	(1,7)
Österreich			32,1	(1,0)	24,3	(1,1)	-7,8	(1,5)		
Spanien			25,1	(1,1)	17,6	(0,7)	-7,5	(1,3)		
Tschechien	34,7	(1,2)			29,3	(0,9)			-5,5	(1,5)
Italien	30,9	(1,1)	24,6	(1,0)	27,3	(0,9)	2,7	(1,4)	-3,6	(1,4)
Frankreich	23,7	(0,9)	24,8	(0,8)	24,2	(0,9)	-0,6	(1,2)	0,5	(1,3)
Ungarn	27,7	(1,0)	24,8	(1,2)	27,1	(1,2)	2,3	(1,7)	-0,5	(1,6)
Belgien (Fläm. Gem.)			42,8	(1,1)	39,9	(0,9)	-2,8	(1,4)		
Bulgarien	16,0	(1,4)	21,2	(1,4)	18,1	(1,4)	-3,1	(1,9)	2,1	(1,9)
Deutschland	32,9	(0,9)	29,7	(0,8)	32,5	(1,2)	2,8	(1,4)	-0,3	(1,5)
Dänemark			34,5	(1,2)	38,8	(1,2)	4,3	(1,7)		
Belgien (Franz. Gem.)			26,9	(0,8)	31,3	(1,2)	4,4	(1,5)		
Lettland	22,5	(1,0)	33,4	(1,3)	28,5	(1,2)	-4,9	(1,8)	6,0	(1,6)
Slowenien	19,9	(1,1)	27,0	(0,9)	27,3	(1,2)	0,3	(1,5)	7,3	(1,6)
Schweden	34,2	(1,0)	42,0	(1,2)	40,9	(1,4)	-1,1	(1,8)	6,7	(1,7)
Slowakei	26,0	(1,2)	27,5	(1,0)	33,1	(1,2)	5,6	(1,6)	7,2	(1,7)
Litauen	14,2	(0,9)	17,7	(0,8)	22,2	(1,0)	4,5	(1,3)	8,0	(1,4)

Footdruck: Differenz im Anteil zwischen 2016 und 2001 bzw. 2016 und 2006 statistisch signifikant ($p < .05$).

Tabelle 7: Veränderung im Anteil von Schülerinnen und Schülern mit geringem Wohlfühlen in der Schule sowie wenig positiv wahrgenommener Beziehungsqualität zu Lehrkräften und Mitschülerinnen und Mitschülern zwischen IGLU 2016 und 2011

Teilnehmer	geringes Wohlfühlen in der Schule			geringe Beziehungsqualität zu Lehrkräften			geringe Beziehungsqualität zu Peers		
	2011 (SE)	2016 (SE)	Δ_{16-11} (SE)	2011 (SE)	2016 (SE)	Δ_{16-11} (SE)	2011 (SE)	2016 (SE)	Δ_{16-11} (SE)
Spanien	9,2 (0,6)	7,2 (0,4)	-2,1 (0,7)	10,1 (0,7)	5,4 (0,4)	-4,7 (0,8)	11,5 (0,6)	7,3 (0,5)	-4,2 (0,8)
Ungarn	10,7 (0,6)	11,1 (0,8)	0,4 (1,0)	6,8 (0,5)	5,5 (0,5)	-1,4 (0,7)	10,7 (0,5)	5,0 (0,4)	-5,7 (0,6)
Österreich	15,3 (0,9)	14,5 (0,7)	-0,8 (1,1)	12,7 (0,9)	10,1 (0,7)	-2,6 (1,2)	8,3 (0,6)	5,5 (0,5)	-2,8 (0,8)
Niederlande	12,8 (0,9)	10,5 (0,9)	-2,3 (1,3)	16,5 (1,2)	13,5 (0,8)	-3,0 (1,4)	6,5 (0,5)	6,0 (0,4)	-0,5 (0,6)
England	13,5 (0,9)	12,0 (0,7)	-1,5 (1,1)	9,9 (0,8)	8,2 (0,6)	-1,7 (1,0)	9,9 (0,7)	7,5 (0,4)	-2,5 (0,8)
Dänemark	7,5 (0,5)	7,7 (0,6)	0,3 (0,8)	14,3 (0,8)	15,9 (0,9)	1,6 (1,2)	2,9 (0,3)	2,8 (0,4)	0,0 (0,5)
Frankreich	9,6 (1,0)	9,2 (0,8)	-0,3 (1,3)	8,1 (0,9)	9,0 (0,8)	0,9 (1,2)	4,7 (0,4)	3,9 (0,4)	-0,7 (0,6)
Schweden	6,9 (0,6)	7,7 (0,7)	0,8 (0,9)	10,6 (0,9)	8,7 (0,6)	-1,9 (1,1)	2,4 (0,3)	2,9 (0,3)	0,6 (0,4)
Bulgarien	6,2 (0,5)	5,5 (0,5)	-0,7 (0,7)	3,8 (0,5)	3,3 (0,4)	-0,5 (0,6)	8,3 (0,6)	7,8 (0,7)	-0,4 (0,9)
Litauen	7,1 (0,6)	7,1 (0,6)	0,0 (0,8)	7,2 (0,5)	6,2 (0,6)	-1,0 (0,7)	7,1 (0,5)	5,8 (0,5)	-1,3 (0,7)
Tschechien	12,2 (0,8)	14,9 (0,8)	2,6 (1,1)	9,4 (0,9)	8,1 (0,6)	-1,3 (1,1)	8,2 (0,7)	5,0 (0,3)	-3,2 (0,7)
Slowakei	9,2 (0,7)	12,1 (0,6)	2,9 (0,9)	8,7 (0,6)	10,0 (0,6)	1,3 (0,9)	8,9 (0,6)	7,3 (0,6)	-1,6 (0,8)
Italien	8,7 (0,8)	11,7 (0,7)	3,0 (1,1)	9,7 (0,7)	7,1 (0,5)	-2,6 (0,9)	7,0 (0,5)	7,6 (0,5)	0,6 (0,7)
Belgien (Franz. Gem.)	12,6 (0,8)	17,4 (1,1)	4,9 (1,4)	9,9 (0,8)	10,1 (0,9)	0,2 (1,3)	16,2 (1,2)	12,4 (0,6)	-3,8 (1,3)
Slowenien	11,7 (0,8)	14,3 (0,9)	2,6 (1,2)	6,5 (0,5)	7,9 (0,8)	1,3 (1,0)	8,9 (0,8)	7,6 (0,5)	-1,2 (0,9)
Deutschland	12,9 (0,8)	16,8 (0,9)	3,9 (1,2)	8,6 (0,7)	9,6 (0,6)	1,0 (1,0)	6,2 (0,5)	5,9 (0,5)	-0,3 (0,7)

Felddruck: Differenz statistisch signifikant ($p < .05$).

6.4 Anerkennungsgerechtigkeit

Welche Veränderungen zeigen sich hinsichtlich des *Wohlfühlens am Schulstandort, der Beziehungsqualität von Lehrkräften zu Mitschülerinnen und Mitschülern und der Beziehungsqualität unter Peers* von 2011 zu 2016? Bei Betrachtung von Tabelle 7 ist festzustellen, dass in Europa zwischen etwa 7 und 17 Prozent aller Schülerinnen und Schüler von wenig positiven Beziehungen zu ihren Lehrkräften und Peers berichten und zum Ausdruck bringen, sich in der Schule eher nicht wohl zu fühlen. Vier Entwicklungen sind zu beobachten: (1.) In einer Gruppe von Ländern (Spanien, Ungarn, Österreich, den Niederlanden, England) zeigen sich positive Entwicklungen im Sinne eines sinkenden Schüleranteils, (2.) in einer weiteren Gruppe von Ländern (Dänemark, Frankreich, Schweden, Bulgarien, Litauen) zeigen sich keine Veränderungen, wohingegen (3.) Slowenien und Deutschland die einzigen Länder mit steigendem Anteil im Indikator Wohlfühlen am Schulstandort sind. (4.) In einer vierten Gruppe von Ländern (Tschechien, Slowakei, Italien, Franz. Gemeinschaft in Belgien) zeigen sich ambivalente Entwicklungen: Während sich die Beziehungsqualität insbesondere zu Mitschülerinnen und Mitschülern signifikant verbessert, berichtet ein signifikant höherer Anteil von Schülerinnen und Schülern in 2016 davon, sich in der Schule nicht wohlzufühlen.

6.5 Zusammenfassung der Ergebnisse

Wie haben sich nun, in zusammenfassender Perspektive der vorgestellten Befunde, die Bildungssysteme in Europa vor dem Hintergrund der EU im Rahmen der ET 2020 definierten Ziele hinsichtlich ihrer Leistungsfähigkeit und Bildungsgerechtigkeit entwickelt? Wie lassen sich diese Entwicklungen im europäischen Referenzrahmen bewerten? Tabelle 8 zeigt die in den Abschnitten 6.1 bis 6.5 anhand von Einzelindikatoren beschriebenen Ergebnisse zusammengeführt. Auf der y-Achse finden sich die Länder hinsichtlich ihrer Entwicklung in Bezug auf die Leistungsfähigkeit (vgl. Tab. 1) verortet, auf der x-Achse jene hinsichtlich ihrer Entwicklungen mit Blick auf die Bildungsgerechtigkeit. Genutzt wurde die Anzahl an positiven und negativen Entwicklungen in beiden Qualitätskomponenten. Als Ergebnis ist festzuhalten, dass 10 der hier betrachteten 18 Länder positive Entwicklungen in Bezug auf die Leistungsfähigkeit verzeichnen können, während sich für nur 7 Länder positive Entwicklungen hinsichtlich der Verbesserung der Bildungsgerechtigkeit beobachten lassen. Für 5 EU-Länder (Spanien, Tschechien, England, Dänemark, Slowenien) sind dabei positive Entwicklungen in beiden Bereichen festzustellen. Dass eine positive Entwicklung hinsichtlich der Leistungsfähigkeit auch mit einer steigenden Bildungsungerechtigkeit einhergehen kann, zeigt sich an der Entwicklung in 5 weiteren EU-Ländern (Slowakei, Ungarn, Litauen, Lettland,

Schweden). Weitestgehend konstante Ergebnisse in beiden Bereichen zeigen Bulgarien, Österreich und die Französische Gemeinschaft in Belgien. Während Frankreich bei einer eher negativen Entwicklung in den Leistungen der Schülerinnen und Schüler konstante Ergebnisse in allen betrachteten Indikatoren der Bildungsgerechtigkeit aufweist, gehen diese in Deutschland, Belgien und den Niederlanden mit einer eher steigenden Bildungsungerechtigkeit einher. Einen bemerkenswerten Fall stellt Italien dar, welches als einziges Land trotz eher negativer Entwicklungen in der Leistungsfähigkeit eine Vielzahl an positiven Entwicklungen in Aspekten der Bildungsgerechtigkeit verzeichnen kann.

Deutschland zählt damit zu einer Reihe von Ländern, in denen sich konstante Leistungsstände bei negativen Entwicklungen in Bezug auf Aspekte von Bildungsgerechtigkeit beobachten lassen. Die Ergebnisse verweisen darauf, dass sowohl hierzulande als auch in vielen anderen Ländern der EU das Thema Bildungsgerechtigkeit derzeit immer noch nicht wirksam angegangen wird. Jene Länder, die in beiden Dimensionen positive Entwicklungen zeigen, sind vermutlich nicht zufällig solche, in denen groß angelegte, landesweite Programme umgesetzt wurden.

Tabelle 8: Entwicklungstendenzen europäischer Teilnehmerländer an IGLU und TIMSS in Bezug Leistungsfähigkeit und Bildungsgerechtigkeit

		Leistungsfähigkeit				
		negativ	eher negativ	konstant	eher positiv	positiv
Bildungsgerechtigkeit	positiv		Italien (5/3)		Spanien (5/0) England (5/2)	Tschechien (5/2)
	eher positiv			Österreich (2/1)	Dänemark (3/1)	Slowenien (3/2)
	konstant		Frankreich (0/0)	Bulgarien (0/0)		
	eher negativ	Niederlande (3/4)	Deutschland (2/6)	Belgien (franz. Gem.) (1/3)	Slowakei (2/3) Ungarn (2/3) Litauen (1/4) Lettland (1/2)	
	negativ		Belgien (Fläm. Gem.) (0/4)		Schweden (1/4)	

Werte in Klammern geben die Anzahl an Indikatoren der Bildungsgerechtigkeit an, in denen sich (positive/negative) Entwicklungen zeigen.

7. Zusammenfassende Diskussion und Ausblick

Mehr als die Hälfte der Länder zeigen positive Veränderungen mit Blick auf die Qualitätsdimension Leistung, aber nur für etwa ein Viertel der Länder lassen sich positive Entwicklungen in Bezug auf Aspekte der Herstellung von Bildungsgerechtigkeit feststellen. Positive Entwicklungen in beiden Dimensionen zeigen sich le-

diglich in 5 Ländern. Damit ist für die meisten Grundschulsysteme Europas festzuhalten, dass sie 10 Jahre nach der Verabschiedung des ‚Strategischen Referenzrahmens für allgemeine und berufliche Bildung‘ (ET 2020) noch nicht auf dem richtigen Weg sind, zwei ihrer vier zentralen Ziele umzusetzen. Dies betrifft insbesondere die Verbesserung von Aspekten der Bildungsgerechtigkeit.

Im europäischen Vergleich zeigt sich, dass Deutschland gemeinsam mit Belgien und den Niederlanden zu den wenigen Ländern gehört, in denen sich in beiden Qualitätsdimensionen negative Entwicklungen abzeichnen. Bei der Bewertung und Generalisierung der Ergebnisse ist jedoch Vorsicht geboten. Zunächst ist festzuhalten, dass in dieser Studie nur Daten für 18 der 28 Mitgliedstaaten der Europäischen Union betrachtet wurden. Zudem liegen für die meisten Länder keine kompletten Datensätze vor, entsprechend könnten positive oder negative Entwicklungen einzelner Länder unter- oder überschätzt werden. Ebenso zeigt sich, dass eine differenzielle Betrachtung nach unterschiedlichen Kompetenzdomänen durchaus zu unterschiedlichen Bewertungen führen kann. Für die Bewertung einer Veränderung wurde vornehmlich die Signifikanz von Differenzwerten berücksichtigt. Zukünftige Studien sollten differenzieller das jeweilige Ausgangsniveau und in den einzelnen Indikatoren zusätzlich die Relevanz von signifikanten Unterschieden stärker kritisch in den Blick nehmen. Beispielsweise wäre für Deutschland zu diskutieren, ob die Bewertung einer tendenziell negativen Veränderung mit Blick auf die Leistungsfähigkeit vor dem Hintergrund eines geringen aber signifikanten Mittelwertunterschieds zwischen IGLU 2006 und 2016 gerechtfertigt ist, oder ob nicht vor dem Hintergrund der sonst konstanten Leistungswerte diese Entwicklung als Ausreißer zu bewerten ist.

Eine weitere Einschränkung ergibt sich aus der für die vorliegenden Analysen gewählten Zeitspanne. Um eine Vergleichbarkeit der Studien und Mindestanzahl an Vergleichsländern herzustellen, haben wir uns entschieden, Studienergebnisse seit der Jahrtausendwende zu berücksichtigen. Für einige Länder, wie Tschechien, für die auch Daten aus TIMSS 1995 vorgelegen hätten, wären wir – wie die Analysen von Martin, Mullis und Loveless (2016) zeigen – hinsichtlich der Bewertung der Entwicklung der Leistungsfähigkeit zu einer abweichenden, negativeren Bewertung gekommen.

Im vorliegenden Beitrag haben wir zudem den Versuch unternommen, das Set an Indikatoren zur Betrachtung von Aspekten der *Bildungsgerechtigkeit* auszuweiten und nach unterschiedlichen Konzeptionen, nämlich der Verteilungs-, Teilhabe- und Anerkennungsgerechtigkeit theoretisch zu verorten. Insgesamt bestätigte sich, dass eine Erweiterung des Indikatorensets durchaus eine differenzierte Beschreibung und daran anschließend auch Bewertung von Entwicklungen ermöglicht: So zeigt sich beispielsweise, dass Länder mit einer insgesamt eher negativen Entwick-

lung durchaus in einzelnen Indikatoren der Anerkennungsgerechtigkeit positive Entwicklungen verzeichnen. Vor diesem Hintergrund wäre es wünschenswert, wenn zukünftige Untersuchungen mit Blick auf die Bewertung der Leistungsfähigkeit von Bildungssystemen noch stärker auch nicht kognitive Outcomes wie soziale oder interkulturelle Kompetenzen oder *well-beings* und *capabilities* berücksichtigen würden (Abschnitt 2).

Für die Bewertung des erweiterten Indikatorensets ist zugleich kritisch anzumerken, dass diese nicht optimal die theoretischen Konstrukte abbilden, sondern maximal eine erste Annäherung darstellen. Darüber hinaus zeigen einige gewählte Indikatoren lediglich akzeptable Messeigenschaften. Insbesondere zeigen sich auch Grenzen in den gewählten Indikatoren zur Betrachtung von Aspekten der Verteilungsgerechtigkeit, da hier im Trend nur sehr begrenzt für beide Studien vergleichbare Indikatoren zur Verfügung stehen. Die Forschung zeigt jedoch deutlich, dass sich durchaus je nach gewähltem Indikator für die soziale Herkunft Ergebnisse unterscheiden können (Hußmann, Stubbe et al., 2017). Auch lässt sich sicherlich für den europäischen Kontext und vor dem Hintergrund gesellschaftlicher und struktureller Veränderungen in den Gesellschaften und Bildungssystemen bemängeln, dass Aspekte von Migration und Inklusion nicht berücksichtigt werden konnten. Wünschenswert wäre, wenn zukünftige Untersuchungen, vielleicht durch die Ergänzung der Fragebögen in einer europäischen Dimension, hier ein stärkeres Augenmerk auf die Schaffung geeigneter Trendindikatoren legen würden.

Anmerkungen

1. Zum Verhältnis der drei Konzeptionen in kritischer Auseinandersetzung vgl. z.B. Hopf (2017) und Wigger (2015).
2. Deskriptive Verteilungen zu einzelnen hier vorgestellten Variablen und Ländern im Trend sind umfassend in den Kapiteln der nationalen Berichtsbände zu TIMSS (Wendt et al., 2016) und IGLU (Hußmann, Wendt et al., 2017) nachzulesen.

Literatur

- Baumert, J. (2016). Leistungen, Leistungsfähigkeit und Leistungsgrenzen der empirischen Bildungsforschung. *Zeitschrift für Erziehungswissenschaft*, 19 (Suppl. 1), 215–253.
- European Commission. (2012). *EU high level group of experts on literacy. Final report. September 2012*. Verfügbar unter: http://ec.europa.eu/dgs/education_culture/repository/education/policy/school/doc/literacy-report_en.pdf [20.11.2017].
- Giesinger, J. (2007). Was heißt Bildungsgerechtigkeit? *Zeitschrift für Pädagogik*, 53, 362–381.
- Gonzales, E. (2014). Calculating standard errors of sample statistics when using international large-scale assessment data. In R. Strietholt, W. Bos, J.E. Gustafsson & M. Rosén (Eds.), *Educational policy evaluation through international comparative assessment* (pp. 59–74). Münster: Waxmann.

- Gustafsson, J.E. (2007). Understanding causal influences on educational achievement through analysis of differences over time within countries. In T. Loveless (Ed.), *Lessons learned: What international assessments tell us about math achievement* (pp. 37–63). Washington, D.C.: Brookings Institution Press.
- Hanushek, E.A. & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46 (3), 607–668.
- Heinzel, F. & Prenzel, A. (2002). *Heterogenität, Integration und Differenzierung in der Primarstufe* (Jahrbuch Grundschulforschung, Bd. 6). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Honneth, A. (2011). *Das Recht der Freiheit. Grundriß einer demokratischen Sittlichkeit*. Berlin: Suhrkamp.
- Hopf, W. (2017). Von der Gleichheit der Bildungschancen zur Bildungsgerechtigkeit für alle – ein Abschied auf Raten vom Gleichheitsideal? In M.S. Baader & T. Freytag (Hrsg.), *Bildung und Ungleichheit in Deutschland* (S. 23–37). Wiesbaden: Springer Fachmedien.
- 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. 195–218). Münster: Waxmann.
- 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.
- Hußmann, A., Wendt, H., Bos, W. & Rieser, S. (Hrsg.). (i.V.). *IGLU 2016. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen*. Münster: Waxmann.
- Joncas, M. & Foy, P. (2011). Sample design in TIMSS and PIRLS. In M.O. Martin & I.V.S. Mullis (Eds.), *TIMSS and PIRLS methods and procedures*. Verfügbar unter: <http://timssandpirls.bc.edu/methods/t-sample-design.html> [28.02.2019].
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2015). *Empfehlungen zur Arbeit in der Grundschule*. Beschluss der Kultusministerkonferenz vom 02.07.1970 in der Fassung vom 11.06.2015. Verfügbar unter: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1970/1970_07_02_Empfehlungen_Grundschule.pdf [28.02.2019].
- Kucharz, D., Irion, T. & Reinhöffer, B. (2011). *Grundlegende Bildung ohne Brüche* (Jahrbuch Grundschulforschung, Bd. 15). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Martin, M.O. & Mullis, I.V.S. (2013). *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade – implications for early learning*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Martin, M.O., Mullis, I.V.S. & Hooper, M. (Eds.). (2016). *Methods and procedures in TIMSS 2015*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Verfügbar unter: <http://timssandpirls.bc.edu/publications/timss/2015-methods.html> [28.02.2019].
- McElvany, N., Kessels, U. Schwabe, F. & Kasper, D. (2017). Geschlecht und Lesekompetenz. 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. 177–194). Münster: Waxmann.

- Mullis, I.V.S., Martin, M.O. & Foy, P. (2008). *TIMSS 2007 international mathematics report: Findings from IEAs Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I.V.S., Martin, M.O., Goh, S. & Cotter, K. (Eds.). (2016). *TIMSS 2015 encyclopedia: Education policy and curriculum in mathematics and science*. Verfügbar unter: <http://timssandpirls.bc.edu/timss2015/encyclopedia/> [28.02.2019].
- Mullis, I.V.S., Martin, M.O., Goh, S. & Prendergast, C. (Eds.). (2017). *PIRLS 2016 encyclopedia: Education policy and curriculum in reading*. Verfügbar unter: <http://timssandpirls.bc.edu/pirls2016/encyclopedia/> [28.02.2019].
- Mullis, I.V.S., Martin, M.O., Gonzalez, E.J. & Kennedy, A.M. (2003). *PIRLS 2001 international report: IEA's study of reading literacy achievement in primary schools*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I.V.S., Martin, M.O. & Loveless, T. (2016). *20 years of TIMSS: International trends in mathematics and science achievement, curriculum, and instruction*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- OECD (Organisation for Economic Co-operation and Development). (2012). *Equity and quality in education: Supporting disadvantage students and schools*. Paris: OECD Publishing.
- Rawls, J. (1971). *A theory of justice*. Cambridge: Harvard University Press.
- Reynolds, D., Sammons, P. Fraine, B., van Damme, J., Townsend, T., Teddlie, C. & Strungfield, S. (2014). Educational effectiveness research (EER): A state-of-the-art review. *School Effectiveness and School Improvement*, 25 (2), 197–230.
- Schleicher, A. (2009). Securing quality and equity in education: Lessons from PISA. *Prospects*, 39, 251–263.
- Schwippert, K. (2019). Was wird aus den Büchern? Sozialer Hintergrund von Lernenden und Bildungsungleichheit aus Sicht der international vergleichenden Erziehungswissenschaft. *Journal for Educational Research Online*, 11 (1), 92–117. Verfügbar unter: <https://www.waxmann.com/artikelART102938> [28.07.2019].
- Sen, A. (2017). *Die Idee der Gerechtigkeit*. München: dtv.
- Stojanov, K. (2011). *Bildungsgerechtigkeit. Rekonstruktionen eines umkämpften Begriffs*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- 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 im internationalen Vergleich* (S. 299–317). Münster: Waxmann.
- Valtin, R. (2017). Einordnung der IGLU-2016-Befunde in das europäische Rahmenkonzept für gute Leseförderung. 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. 315–326). Münster: Waxmann.
- van Damme, J. & Bellens, K. (2017). Countries strive towards more quality and equity in education: Do they show success or failure? Evidence from TIMSS 2003 and 2011, for grade 4. In M. Rosén, K. Yang Hansen & U. Wolff (Eds.), *Cognitive abilities and educational outcomes. A Festschrift in honour of Jan-Eric Gustafsson* (pp. 127–148). Cambridge: Springer.

- vbw (Vereinigung der Bayerischen Wirtschaft e.V.). (Hrsg.). (2007). *Bildungsgerechtigkeit. Jahrgutachten 2007*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Wendt, H., Bos, W., Goy, M., Jusufi, D. & Kasper, D. (2017). *TIMSS 2015. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen*. Münster: Waxmann.
- Wendt, H. & Kasper, D. (2016). Subject-specific strength and weaknesses of fourth-grade students in Europe: A comparative latent profile analysis of multidimensional proficiency patterns based on PIRLS/TIMSS combined 2011. *Large Scale Assessment in Education, 14* (4). doi:10.1186/s40536016-0026-2
- Wendt, H., Steinmayr, R. & Kasper, D. (2016), Geschlechterunterschiede in mathematischen und naturwissenschaftlichen Kompetenzen. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 257–298). Münster: Waxmann.
- Wendt, H., Stubbe, T.C., Schwippert, K. & Bos, W. (Hrsg.). (2015). *10 Jahre international vergleichende Schulleistungsforschung in der Grundschule. Vertiefende Analysen zu IGLU und TIMSS 2001 bis 2011*. Münster: Waxmann.
- Wendt, H., Tarelli, I., Bos, W., Frey, K. & Vennemann, M. (2012). Ziele, Anlage und Durchführung der Trends in International Mathematics and Science Study (TIMSS 2011). In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 27–68). Münster: Waxmann.
- Wigger, L. (2015). Bildung und Gerechtigkeit – Eine Kritik des Diskurses um Bildungsgerechtigkeit aus bildungstheoretischer Sicht. In V. Manitius, B. Hernstein, N. Berkemeyer & W. Bos (Hrsg.), *Zur Gerechtigkeit von Schule. Theorien, Konzepte, Analysen* (S. 72–92). Münster: Waxmann.
- Willems, A.S., Wendt, H., Gröhlich, C., Walzebug, A. & Bos, W. (2014). Mehr Chancengerechtigkeit durch die Ganztagschule? Ein Vergleich selektionsbedingter Ungleichheiten in der Bildungsteilnahme bei Schülerinnen und Schülern an Ganz- und Halbtagsgrundschulen auf Basis der IGLU und TIMSS 2011 Daten. In H.G. Holtappels, A.S. Willems, M. Pfeifer, W. Bos, & N. McElvany (Hrsg.), *Jahrbuch der Schulentwicklung, Bd. 18. Daten, Beispiele und Perspektiven* (S. 62–100). Weinheim: Beltz.
- Woessmann, L. (2004). *How equal are educational opportunities? Family background and student achievement in Europe and the United States*. Bonn: IZA.