

Köller, Olaf

Entwicklung und Erträge der jüngeren empirischen Bildungsforschung

Fatke, Reinhard [Hrsg.]; Oelkers, Jürgen [Hrsg.]: Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart. Weinheim; u.a. : Beltz Juventa 2014, S. 102-122. - (Zeitschrift für Pädagogik, Beiheft; 60)



Empfohlene Zitierung/ Suggested Citation:

Köller, Olaf: Entwicklung und Erträge der jüngeren empirischen Bildungsforschung - In: Fatke, Reinhard [Hrsg.]; Oelkers, Jürgen [Hrsg.]: Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart. Weinheim; u.a. : Beltz Juventa 2014, S. 102-122. - (Zeitschrift für Pädagogik, Beiheft; 60) - URN: urn:nbn:de:0111-opus-90893

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

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
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

60. Beiheft

April 2014

ZEITSCHRIFT FÜR PÄDAGOGIK

**Das Selbstverständnis
der Erziehungswissenschaft:
Geschichte und Gegenwart**

BELTZ VERLAG **JUVENTA**

Zeitschrift für Pädagogik · 60. Beiheft

Zeitschrift für Pädagogik · 60. Beiheft

Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart

Herausgegeben von
Reinhard Fatke und Jürgen Oelkers

BELTZ JUVENTA

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, bleiben dem Beltz-Verlag vorbehalten.

Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form – durch Fotokopie, Mikrofilm oder ein anderes Verfahren – reproduziert oder in eine von Maschinen, insbesondere Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. Auch die Rechte der Wiedergabe durch Vortrag, Funk- und Fernsehsendung, im Magnettonverfahren oder auf ähnlichem Wege bleiben vorbehalten. Fotokopien für den persönlichen oder sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopie hergestellt werden. Jede im Bereich eines gewerblichen Unternehmens hergestellte oder genutzte Kopie dient gewerblichen Zwecken gem. § 54 (2) UrhG und verpflichtet zur Gebührenzahlung an die VG Wort, Abteilung Wissenschaft, Goethestr. 49, 80336 München, bei der die einzelnen Zahlungsmodalitäten zu erfragen sind.

© 2014 Beltz Juventa · Weinheim und Basel

www.beltz.de · www.juventa.de

Herstellung: Lore Amann

Satz: text plus form, Dresden

Druck und Bindung: Beltz Bad Langensalza GmbH, Bad Langensalza

Printed in Germany

ISSN 0514-2717

Bestell-Nr. 443502

Inhaltsverzeichnis

Reinhard Fatke/Jürgen Oelkers

Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart. Einleitung zum Beiheft	7
---	---

Wandel des Selbstverständnisses

Klaus-Peter Horn

Pädagogik/Erziehungswissenschaft der Gegenwart – Zur Entwicklung der deutschen Erziehungswissenschaft im Spiegel ihrer disziplinären Selbstreflexion (1910–2010)	14
--	----

Patrick Bühler

Einführungen in die Pädagogik. Die Entwicklung einer Gattung	33
--	----

Michael Geiss

Die angemessene Form: Abstraktion und die Suche nach einem pädagogischen Kalkül	47
--	----

Markus Rieger-Ladich

Pädagogik als kritische Theorie? Intellektuelle Stellungskämpfe nach 1945	66
--	----

Wissenschaftspraxis und Empirie

Jürgen Oelkers

Praxis und Wissenschaft. Überlegungen zur Forschungsstruktur der Erziehungswissenschaft	85
--	----

Olaf Köller

Entwicklung und Erträge der jüngeren empirischen Bildungsforschung	102
--	-----

Felicitas Thiel/Kai S. Cortina/Hans Anand Pant

Steuerung im Bildungssystem im internationalen Vergleich	123
--	-----

Heinz-Elmar Tenorth

Politikberatung und Wandel der Expertenrolle oder: Die Expertise der Erziehungswissenschaft	139
--	-----

Die Erziehungswissenschaft und ihre Nachbarwissenschaften

Bernhard Schmidt-Hertha/Rudolf Tippelt

Erziehungswissenschaft und das Verhältnis zu ihren Bezugsdisziplinen vor dem Hintergrund der Stellen- und Besetzungspraxis (1995 bis 2012)	172
--	-----

Peter Kauder

Themenkonjunkturen im Spiegel erziehungswissenschaftlicher Dissertationen und Habilitationen 1945–2009	184
---	-----

Nicole Becker

Mehr verstehen, besser handeln? Zum Verhältnis von Pädagogik und Neurowissenschaften	208
---	-----

Essay

Roland Reichenbach

Schulkritik. Eine „metaphorologische“ Betrachtung	226
---	-----

Olaf Köller

Entwicklung und Erträge der jüngeren empirischen Bildungsforschung

1. Überblick

Kein anderes Feld in der empirisch ausgerichteten Sozialforschung dürfte in Deutschland in den letzten 20 Jahren einen solchen Aufschwung genommen haben wie die empirische Bildungsforschung. Diese Dynamik, die insbesondere im Bereich der Erziehungswissenschaft entstanden ist, ist umso bemerkenswerter, als bis zum Beginn der 1990er-Jahre das Hauptinteresse erziehungswissenschaftlicher Forschung auf der Entwicklung und Erprobung von Modellen zur Optimierung der Arbeit in Einzelschulen und dem Entwurf didaktischer Modelle und deren Einführung in die Unterrichtspraxis lag. Selbst die von Heinrich Roth Anfang der 1960er-Jahre eingeleitete „realistische Wendung“ (Roth, 1963) konnte diesen vorherrschenden Themen wenig entgegenwirken.

Untersuchungen dagegen, die empirisch gestützt die Beschreibung, Erklärung und Vorhersage der Erträge fachlichen Lernens und schulischer Erziehungs- und Sozialisationsprozesse in den Mittelpunkt rückten, waren unzeitgemäße Unternehmungen. Dies galt für die 1991/92 begonnene Kohortenlängsschnittstudie *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter* (BIJU; vgl. u. a. Köller, Baumert, Cortina & Trautwein, 2010) als auch für die Teilnahme Deutschlands an der *Reading Literacy Study* (RLS; Elley, 1992) und an der *Third International Mathematics and Science Study* (Baumert et al., 1997). Umso bemerkenswerter waren die fast hysterischen öffentlichen Reaktionen auf die Publikation der TIMSS-Befunde. Es folgten die Konstanzer Beschlüsse der Kultusministerkonferenz (KMK), die eine zukünftige Beteiligung Deutschlands an internationalen Schulleistungsstudien wie PISA (Baumert et al., 2001) vorsahen und den Grundstein für ein nationales Bildungsmonitoring in Deutschland legten. Zeitgleich wurde die Grundlage für den BLK-Modellversuch zur Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts erarbeitet (SINUS-Programm; BLK, 1997). Nur zwei Jahre später legte die Deutsche Forschungsgemeinschaft (DFG) das Schwerpunktprogramm *Bildungsqualität von Schule* (BiQua; vgl. Prenzel & Allolio-Näcke, 2006) auf, und es folgte eine immer noch nicht zum Abschluss gekommene Expansion der empirischen Bildungsforschung. Welche langfristigen Erträge diese Entwicklung haben wird, ist offen. Sicherlich wird auch dieser Beitrag die Frage nicht beantworten können. Was der vorliegende Beitrag vielmehr leisten will, ist eine knappe und an vielen Stellen verkürzte Rückschau auf die Entwicklung der jüngeren empirischen Bildungsforschung seit Beginn der 1990er-Jahre (zur Geschichte der empirischen Bildungsforschung seit Beginn des 20. Jahrhunderts s. die Ausführungen bei Gräsel, 2011; Ditton, 2011). Dabei soll ihr institutioneller und struktureller Ausbau be-

rücksichtigt werden, ebenso wie eine Beschreibung und Bewertung ihrer wissenschaftlichen Erträge, Letztere verzerrt durch die subjektive Sicht des Autors. Konzidiert werden muss weiterhin, dass der Beitrag keine Gesamtdarstellung anstrebt, sondern sich vorwiegend auf zentrale, in der Fachöffentlichkeit prominente und gewissermaßen strukturbildende Fördermaßnahmen der Deutschen Forschungsgemeinschaft (DFG), des Bundesministeriums für Bildung und Forschung (BMBF) und der Kultusministerkonferenz (KMK) konzentriert.

Begonnen wird mit Ausführungen zum Selbstverständnis der empirischen Bildungsforschung, die sich nicht als Teildisziplin der Pädagogik/Erziehungswissenschaft versteht, wenn auch die Expansion der empirischen Bildungsforschung an deutschen Universitäten in den letzten 20 Jahren vor allem auf Kosten erziehungswissenschaftlicher Teildisziplinen gegangen ist. So wurden nicht wenige Professuren für Allgemeine Didaktik oder Schulpädagogik in solche für empirische Bildungsforschung umgewidmet.

2. Zum heutigen Selbstverständnis der empirischen Bildungsforschung

Will man sich dem Gegenstandsbereich der empirischen Bildungsforschung annähern, so wird man zunächst auf die Definition des Deutschen Bildungsrats (1974, S. 23) zurückgreifen: „Man kann Bildungsforschung in einem weiteren und engeren Sinne auslegen. Im engeren Sinne hat es sie als Unterrichtsforschung schon immer gegeben. Im weiteren Sinne kann sie sich auf das gesamte Bildungswesen und seine Reform im Kontext von Staat und Gesellschaft beziehen, einschließlich der außerschulischen Bildungsprozesse. Wie weit oder eng aber auch die Grenzen der Bildungsforschung gezogen werden, es sollte nur dann von Bildungsforschung gesprochen werden, wenn die zu lösende Aufgabe, die Gegenstand der Forschung ist, theoretisch oder empirisch auf Bildungsprozesse (Lehr-, Lern-, Sozialisations- und Erziehungsprozesse), deren organisatorische und ökonomische Voraussetzungen oder Reform bezogen ist.“ Bildungsforschung in diesem Sinne ist ein interdisziplinäres Vorhaben von hoher gesellschaftlicher Relevanz, an der viele Disziplinen beteiligt sind: die Psychologie, die Soziologie, die Erziehungswissenschaft, die Ökonomie, die Fachdidaktiken und neuerdings die Neurowissenschaften.

Einmal abgesehen von Teilen der Erziehungswissenschaft und der Fachdidaktiken, finden sich heute unter dem Dach der Bildungsforschung Disziplinen mit stark sozial- und naturwissenschaftlicher Tradition, alle im Übrigen in ihrem wissenschaftstheoretischen Verständnis dem kritischen Rationalismus (Popper, 1963) verpflichtet, alle mit dem Anspruch, ihre Modelle bzw. die aus Theorien abgeleiteten Hypothesen in empirischen Untersuchungen – seien sie quantitativer oder qualitativer Natur – scheitern zu lassen. Ziel ist ausdrücklich nicht die Selbstbestätigung der theoretischen Annahmen, sondern die Prüfung von Theorien mit dem expliziten Ziel, sie dem Scheitern auszusetzen: „Es geht nicht um Enttheoretisierung, sondern Entideologisierung der Pädagogik, um die Stärkung ihres Sachverstandes zur Minderung ihrer Ideologie-Anfälligkeit“ (Roth, 1963, S. 117).

Bezogen auf die obige Definition des Bildungsrates legt die empirische Bildungsforschung ihre Arbeiten im weiteren Sinne aus und versteht institutionelle ebenso wie informelle und non-formale Bildungsprozesse über die Lebensspanne als ihren Gegenstand. Zu ihren Forschungsthemen gehören frühkindliche Erziehungs- und Bildungsprozesse ebenso wie instruktionspsychologische Fragestellungen, Studien zur Effektivität und Effizienz von Weiterbildungsmaßnahmen im Erwachsenenalter oder Arbeiten zu Lernprozessen mithilfe von Internetangeboten. Die thematische und disziplinäre Breite der empirischen Bildungsforschung ist Stärke und Schwäche zugleich. Sie führt zu einem verbesserten multiperspektivischen Verständnis von Bildungsprozessen, erschüttert aber gleichzeitig die disziplinäre Identität ihrer Akteure. Diese „Identitätsprobleme“, aber auch das damit verbundene Selbstverständnis der empirischen Bildungsforschung schließen aus, dass sie sich in ihrer Gänze unter dem Dach der Pädagogik/Erziehungswissenschaft verorten kann und will. Dies mag auch erklären, warum sich die verschiedenen Disziplinen im Jahre 2012 in der *Gesellschaft für empirische Bildungsforschung* (GEBF) zusammengefunden haben.

3. Bildungsforschung und die empirische Wende der 1990er-Jahre

Zu Beginn der 1990er-Jahre spielten sich empirische Projekte zu Themen der Bildungsforschung im Wesentlichen außerhalb der Erziehungswissenschaft und Fachdidaktiken ab. Die fest etablierte Pädagogische Psychologie als angewandte Disziplin der Psychologie bearbeitete u. a. Fragen des Wissenserwerbs im schulischen Unterricht (z. B. Weinert, Schrader & Helmke, 1989). Weitere bedeutende empirische Arbeiten lagen im Bereich der Motivationsforschung (vgl. Heckhausen, 1980). Die empirisch arbeitende Bildungssoziologie war sehr stark durch die Lebensverlaufsstudien am Max-Planck-Institut für Bildungsforschung (MPI für Bifo) geprägt (vgl. u. a. Mayer, 1990). Im Bereich der Ökonomie war 1983 das sozio-oekonomische Panel (SOEP; Kroh, 2013) begonnen worden, eine repräsentative Haushaltsbefragung, die bis heute fortgeführt wird und eine breite empirische Basis für die Beantwortung bildungsökonomischer Forschungsfragen bietet. In der Erziehungswissenschaft existierten vor allem die großen längsschnittlichen empirischen Studien Helmut Fends (im Überblick Fend, 1991), die sich im Schnittbereich zwischen Erziehungswissenschaft und Entwicklungspsychologie verorteten, sowie Forschungsarbeiten im Bereich der pädagogischen Diagnostik, die seit 1971 am Zentrum für Empirische Pädagogische Forschung (zefp) in Landau durchgeführt wurden. Rainer Lehmann an der Universität Hamburg (später Humboldt-Universität zu Berlin) und Jürgen Baumert am Institut für die Pädagogik der Naturwissenschaften (IPN) arbeiteten an den internationalen Schulleistungsstudien der *International Association for the Evaluation of Educational Achievement* (IEA) zum Leseverstehen und zu den mathematisch-naturwissenschaftlichen Kompetenzen (s. o.). Stark beeinflusst durch Fends Arbeiten zu individuellen Entwicklungsprozessen unter den institutionellen Rahmenbedingungen von Schule, begann 1991 die BIJU-Studie (s. o.) als Kooperationsprojekt zwischen dem MPI für Bifo und dem IPN. Ein Kernziel dieser Studie

lag in der Beschreibung und Erklärung von Entwicklungsverläufen schulischer Leistungen in unterschiedlichen Fächern (Mathematik, Englisch, Biologie, Physik). Bei den Leistungsmessungen fand eine enge Anlehnung an die Schulleistungsstudien der IEA und an das in den USA durchgeführte *National Assessment of Educational Progress* (NAEP; Jones & Olkin, 2004) statt.

Im Jahr 1995 begann die *Lernausgangslagenuntersuchung* (LAU; Lehmann & Peek, 1997) in Hamburg. Erstmals wurde eine komplette Jahrgangskohorte (Schülerinnen und Schüler der 5. Jahrgangsstufe) in verschiedenen Fächern getestet und dann über die gesamte schulische Zeit wissenschaftlich begleitet. Analog zu BIJU kamen in LAU wiederholt Fachleistungstests zum Einsatz, um Leistungsverläufe nachzeichnen und erklären zu können.

Wenig später begann in Landau in Kooperation zwischen der Universität und dem zepf das Projekt „Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext“ (MARKUS; Helmke & Jäger, 2001). MARKUS stellte die erste flächendeckende Untersuchung der Ergebnisse und Bedingungen eines Unterrichtsfaches, hier der Mathematik, dar.

BIJU, LAU und MARKUS können gemeinsam mit TIMSS und RLS von ihrer Anlage her als Prototypen der großen *Large-scale Assessments* bezeichnet werden. An LAU beispielsweise beteiligten sich mehr als 12 000 Schülerinnen und Schüler. Aber erst den intensiven öffentlichen Reaktionen auf die TIMSS-Befunde ist es zu verdanken, dass sich dieses kostenintensive Untersuchungsparadigma forschungspolitisch etablieren konnte und im Nationalen Bildungspanel (NEPS; vgl. u. a. Blossfeld, von Maurice & Schneider, 2009) seinen vorläufigen Höhepunkt erreicht hat.

4. Forschungsförderung nach der empirischen Wende

Die empirische Wende Anfang der 1990er-Jahre, gefolgt von der Publikation der PISA-2000-Befunde (Baumert et al., 2001), hatte neben weiteren bildungspolitischen Konsequenzen (z. B. Einführung von Bildungsstandards) erhebliche Folgen für die Forschungsförderung in Deutschland. Rückblickend sind die heute existierenden Strukturen der universitären empirischen Bildungsforschung maßgeblich durch die Förderinitiativen der DFG zu Beginn des 21. Jahrhunderts entstanden. Inzwischen wurden diese Initiativen weitgehend durch das große „Rahmenprogramm Bildungsforschung“ des Bundesministeriums für Bildung und Forschung (BMBF, 2007) abgelöst bzw. weitergeführt. Dem BMBF und den 16 Ländern ist es weiterhin zu verdanken, dass Strukturen für ein langfristiges Bildungsmonitoring in Deutschland geschaffen wurden und mit dem Pakt für Forschung und Innovation erhebliche zusätzliche Mittel in die außeruniversitäre Forschung fließen, die dort zur Weiterentwicklung der Bildungsforschung beitragen.

4.1 DFG-Förderung als Katalysator für die Expansion der empirischen Bildungsforschung

Anfang des neuen Jahrtausends entstand die Initiative zur Einrichtung von *Forschergruppen Empirische Bildungsforschung* der DFG (2002). Das unbefriedigende Abschneiden Deutschlands in PISA 2000 (Baumert et al., 2001) gab Anlass zu Fragen nach den Ursachen und zu deren empirischer Erforschung. Die DFG sah in der empirischen Bildungsforschung einen wichtigen Ansprechpartner für die Bildungspolitik bei der Bereitstellung von Systemwissen. Ziel der Initiative war es, universitäre Strukturen aufzubauen, in denen disziplinübergreifend Kernprobleme des Bildungssystems erforscht werden sollten, um letztlich dieses Wissen in Handlungswissen für Politik zu übertragen. Dieser Initiative zur langfristigen Gründung von universitären Zentren der empirischen Bildungsforschung waren in den 1990er-Jahren die DFG-Unterstützung der fachdidaktischen Forschung, hier vor allem in den mathematisch-naturwissenschaftlichen Fachdidaktiken, sowie das Schwerpunktprogramm *Lehr-Lern-Prozesse in der kaufmännischen Erstausbildung* vorausgegangen. Auf das BiQua-Schwerpunktprogramm wurde oben schon hingewiesen. Parallel zu der Einrichtung von Forschergruppen nahm das DFG-Graduiertenkolleg „Bildungsgangforschung“ an der Universität Hamburg seine Arbeit im Jahre 2002 auf.

Die Initiative *Forschergruppen Empirische Bildungsforschung* hat letztlich zu positiven Begutachtungen von drei interdisziplinären Zentren geführt: *Naturwissenschaftlicher Unterricht* (NWU) an der Universität Duisburg-Essen; *Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter* (BIKS) an der Universität Bamberg; und *Analyse und Förderung effektiver Lehr-Lernprozesse* an der Universität Tübingen.

Für die Standorte Bamberg und Tübingen lässt sich rückblickend festhalten, dass die Etablierung der DFG-Forschergruppen dort zum nachhaltigen Aufbau von leistungsstarken Zentren der empirischen Bildungsforschung beigetragen hat. Mitglieder der BIKS-Gruppe in Bamberg waren später an der Planung und Realisierung des Nationalen Bildungspanels (Blossfeld et al., 2009), das in einem Multi-Kohorten-Design Bildungsprozesse über die Lebensspanne empirisch untersucht, beteiligt. Mit Beginn des Jahres 2014 wurde diese Struktur in die Leibniz-Gemeinschaft und damit letztlich auch in die dauerhafte Bund-Länder-Förderung aufgenommen.

In Tübingen war die Forschergruppe von Beginn an ein kooperatives Vorhaben des Leibniz-Instituts für Wissensmedien (IWM) und der Universität Tübingen. Im Jahre 2009 bildeten die von der DFG aufgebauten Strukturen die Grundlage für die Gründung des WissenschaftsCampus „Bildung in Informationsumwelten“ der aus Mitteln der Leibniz-Gemeinschaft, des Landes Baden-Württemberg und der Universität Bamberg finanziert wird. Neben der Psychologie, der Erziehungswissenschaft und der Soziologie sind dort auch die Informatik und die Medizin in das Forschungsnetzwerk integriert. Mithilfe dieses Netzwerks ist es dem Standort Tübingen im Jahr 2012 schließlich gelungen, im Rahmen der DFG-Exzellenz-Initiative die interdisziplinäre Graduiertenschule *Learning, Educational Achievement, and Life Course Development* (LEAD) einzuwerben.

Am Standort Duisburg-Essen sind im Rahmen der Forschergruppe viele empirisch arbeitende Nachwuchswissenschaftlerinnen und -wissenschaftler ausgebildet worden, die heute Fachdidaktik-Professuren an deutschen Universitäten besetzen.

Den Abschluss der DFG-Bemühungen, die empirische Bildungsforschung zu stärken, bildete die Einrichtung des Schwerpunktprogramms *Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen* im Jahre 2007, das vor allem der Entwicklung, theoretischen Fundierung und empirischen Überprüfung von Kompetenzmodellen in unterschiedlichen Domänen dienen soll (Fleischer, Leutner & Klieme, 2012).¹ Inwieweit dieses Programm in der Lage sein wird, langfristige oder gar dauerhafte inhaltliche oder strukturelle Stärkungen der empirischen Bildungsforschung zu bewirken, ist eine offene Frage. Zusammenfassend lässt sich aber schlussfolgern, dass es der DFG vor allem mit den Forschergruppen gelungen ist, forschungsintensive Strukturen in Deutschland zu schaffen. Die Standorte Bamberg und Tübingen zählen zu den national und international erfolgreichsten Zentren empirischer Bildungsforschung.

Ein wichtiger Nebeneffekt der DFG-Initiativen war, dass viele Länder vergleichbare Initiativen gestartet haben, mit denen die Bildungsforschung weiter gestärkt wurde. Das vielleicht prominenteste Beispiel hierfür ist das IDEA-Zentrum für Individuelle Entwicklung und Lernförderung in Frankfurt, das mit erheblichen Mitteln des Landes Hessen gegründet wurde. Nicht unerwähnt bleiben soll in diesem Zusammenhang auch das Promotionsprogramm „Fachdidaktische Lehr- und Lernforschung – Didaktische Rekonstruktion“, das in den Jahren 2001 bis 2008 vom Land Niedersachsen gefördert wurde.

4.2 *Programmförderung des BMBF: Rahmenprogramm „Empirische Bildungsforschung“*

Mit dem Auslaufen der DFG-Initiativen zur Förderung empirischer Bildungsforschung hat das BMBF mit großem Finanzvolumen die Drittmittelfinanzierung des Feldes übernommen und fördert seit 2007 mit dem *Rahmenprogramm Empirische Bildungsforschung* universitäre und außeruniversitäre Forschungsvorhaben. In diesem Zusammenhang muss angemerkt werden, dass der Bund gemeinsam mit den Ländern jenseits dieser Initiativen seit jeher eine Förderung der Bildungsforschung in außeruniversitären Einrichtungen, beispielsweise solchen der Leibniz-Gemeinschaft (s. u.), betreibt.

Insgesamt verfolgt das BMBF mit dem Rahmenprogramm ähnliche Ziele wie zuvor die DFG: „Das Rahmenprogramm dient dazu, die empirische Bildungsforschung in Deutschland insgesamt strukturell und damit langfristig zu stärken, ihre Internationali-

1 Seit 2011 existiert noch eine weitere DFG-Forschergruppe zu „Mechanismen der Elitebildung im deutschen Bildungssystem“. In dieser Gruppe werden Prozesse der Konstruktion und Herstellung von Exzellenz in Bildungsinstitutionen von der Vorschule bis zur Hochschule und deren Bedeutung für die Bildungsadressaten und Professionellen untersucht.

tät und Interdisziplinarität weiter zu fördern und zugleich Erkenntnisse für die Reformprozesse im Bildungsbereich zu generieren“ (BMBF, 2007, S. 5). Die strukturelle Sicherung und Verstärkung der empirischen Bildungsforschung soll vor allem auch durch eine qualifizierte Förderung des wissenschaftlichen Nachwuchses erreicht werden.

Eine Bewertung der Ergebnisse der noch laufenden Forschungsförderungsinitiative ist schwierig, die oben bereits genannte Überführung des Nationalen Bildungspanels in die Leibniz-Gemeinschaft kann allerdings als erster großer Erfolg gewertet werden.

4.3 Finanzierung des Bildungsmonitorings durch den Bund und die Länder

Mit den oben erwähnten Konstanzer Beschlüssen der KMK von 1997 wurde die Beteiligung Deutschlands an internationalen *Large-scale Assessments* (zunächst an PISA) beschlossen. Über das internationale Benchmarking hinaus sollte PISA genutzt werden, um ein nationales Bildungsmonitoring anhand von Leistungsvergleichen zwischen den 16 Ländern der Bundesrepublik Deutschland zu etablieren. PISA 2000 wurde dann von einem wissenschaftlichen Konsortium unter der Ägide des Max-Planck-Instituts für Bildungsforschung durchgeführt (vgl. Baumert et al., 2001, 2002). In solchen Konsortialstrukturen wurde PISA bis 2009 fortgeführt. Im Jahre 2001 wurde das Programm um die Internationale Grundschul-Lese-Untersuchung (IGLU; Bos et al., 2003) erweitert. IGLU wird ebenfalls von Konsortien durchgeführt und vom Bund und von den Ländern finanziert.

Mit der Einführung der KMK-Bildungsstandards für die Primarstufe und die Sekundarstufe I in den Jahren 2003 und 2004 wurde Ende 2004 das durch die 16 Länder finanzierte *Institut zur Qualitätsentwicklung im Bildungswesen* (IQB) an der Humboldt-Universität zu Berlin gegründet, das seit 2009 die vorher in PISA durchgeführten Ländervergleiche (nationales Bildungsmonitoring) übernommen hat (vgl. Köller, Knigge & Tesch, 2010; Stanat, Pant, Böhme & Richter, 2012). Das IQB hat neben seinem Auftrag zum Bildungsmonitoring auch den Anspruch, grundlagenbasierte Forschung im Bereich des Assessments durchzuführen und die standardbasierte Unterrichtsentwicklung voranzubringen. Seit der erfolgreichen externen Evaluation ist das IQB mit seiner Länderfinanzierung als wissenschaftliche Infrastruktur auf Dauer gestellt und wird langfristig das nationale Bildungsmonitoring in Deutschland durchführen. Als zweite Infrastruktur für das Bildungsmonitoring wurde das Bund-Länderfinanzierte *Zentrum für Internationale Bildungsvergleichsstudien* (ZIB) in München gegründet. Das ZIB stellt ein wissenschaftliches Netzwerk aus der School of Education der Technischen Universität München, dem IPN und dem Deutschen Institut für Internationale Pädagogische Forschung (DIPF) dar. Die Projektförderung ist zunächst auf sechs Jahre befristet (2011 bis 2016). Auftrag ist die nationale Durchführung und Berichterstattung von PISA sowie Forschung im Bereich der pädagogisch-psychologischen Diagnostik und Methoden. Zusätzlich zum ZIB stellt das BMBF jährlich Forschungsmittel in Höhe von einer Million Euro für universitäre Forschung in Anlehnung an die *Large-scale Assessments* zur Verfügung.

Schließlich haben Bund und Länder die Nationale Bildungsberichterstattung am DIPF institutionalisiert (vgl. Autorengruppe Bildungsberichterstattung, 2012). Der Bildungsbericht erscheint alle zwei Jahre und berichtet indikatorengestützt über die Entwicklung des Bildungssystems. Die Stärke der Bildungsberichterstattung liegt darin, dass dabei Bildungsprozesse über die gesamte Lebensspanne in den Fokus genommen werden.

Bund und Länder haben so Strukturen für eine Dauerbeobachtung des Bildungssystems geschaffen, das wissenschaftsbasiert ist und neben seinen wissenschaftlichen Dienstleistungen Grundlagenforschung betreiben soll. Dieses ist umso bemerkenswerter, als die Bildungspolitik Infrastrukturen finanziert, die in den vergangenen Jahren eher schlechte Nachrichten verbreitet haben. Der Dialog zwischen Bildungsforschung und Politik ist sehr stabil, obwohl die großen Schulleistungsstudien in regelmäßigen Abständen der Politik im Wesentlichen ihre Versäumnisse vorhalten. Allein in dem anhaltenden Dialog ist ein großer Ertrag der letzten 20 Jahre zu sehen.

5. Außeruniversitäre Bildungsforschung

Die von Roth eingeleitete realistische Wende der Erziehungswissenschaft manifestierte sich in den 1960er-Jahren unter anderem in Neugründungen von Bund-Länder-finanzierten außeruniversitären Forschungseinrichtungen. Die empirische Bildungsforschung wurde mit der Gründung des MPI für Bifo (1963) und dem IPN (1966) deutlich gestärkt, nachdem bereits im Jahre 1951 das DIPF seine Arbeit aufgenommen hatte. Alle drei Institute wurden und werden gemeinsam vom Bund und von den Ländern finanziert (jeweils 50 Prozent). DIPF und IPN waren zunächst sogenannte Blaue-Liste-Institute, später wurden sie (nach positiver Evaluation durch den Wissenschaftsrat) Mitglieder der Leibniz-Gemeinschaft. MPI für Bifo, DIPF und IPN verfolgen überregionale Forschungsvorhaben, die sich u. a. von universitärer Forschung dadurch abgrenzen, dass sie langfristiger ausgerichtet sind und ein Finanzvolumen erreichen, das typischerweise an Universitäten nicht realisierbar ist.

Für alle drei Institute war von Beginn an die Interdisziplinarität konstituierend: Psychologie, Soziologie, Erziehungswissenschaft, Bildungsökonomie, Fachdidaktik und Bildungsrecht koexistierten. Und Auftrag der drei Institute war und ist es explizit, Bildungsforschung interdisziplinär durchzuführen. Rückblickend gilt für alle drei Einrichtungen, dass die großen Impulse, die zunächst von ihnen im Bereich der Bildungsforschung ausgegangen waren, Ende der 1980er-Jahre beinahe zum Erliegen gekommen waren und erst mit der Berufung von Jürgen Baumert 1991 an das IPN, 1996 an das MPI für Bifo und der durch ihn forcierten Etablierung der großen Schulleistungsstudien (s. o.) eine neue Dynamik in die außeruniversitäre empirische Bildungsforschung kam, die bis heute am DIPF und IPN anhält, am MPI für Bifo nach Baumerts Emeritierung und der bereits früheren Aufgabe der Bildungssoziologie weitgehend zum Erliegen gekommen ist. Stattdessen konzentriert man sich dort auf neurowissenschaftliche, kognitionspsychologische und kulturhistorische Schwerpunkte, deren Nähe zur empirischen Bildungsforschung nicht mehr sichtbar ist.

Vor allem die Leibniz-Gemeinschaft hat es in den letzten Jahren verstanden, die empirische Bildungsforschung an ihren Einrichtungen auszubauen. Neben den eher erziehungswissenschaftlichen Einrichtungen wie IPN und DIPF sind Strukturen entstanden, in denen aus unterschiedlichen Perspektiven Bildungsforschung betrieben wird. Die großen ökonomischen Forschungsinstitute der Leibniz-Gemeinschaft (Zentrum für Europäische Wirtschaftsforschung, ZEW; Deutsches Institut für Wirtschaftsforschung, DIW; Leibniz-Institut für Wirtschaftsforschung, IFO) haben leistungsstarke bildungsökonomische Arbeitsgruppen etabliert. Das eher soziologisch ausgerichtete Wissenschaftszentrum Berlin für Sozialforschung (WZB) führt große empirische bildungssoziologische Forschungsprojekte durch. Das Leibniz-Institut für Wissensmedien (IWM) untersucht aus einer psychologischen Perspektive Wissenserwerbsprozesse mithilfe des Web 2.0. Am Deutschen Institut für Erwachsenenbildung/Leibniz-Zentrum für lebenslanges Lernen (DIE) stehen Fragen des Wissens- und Kompetenzerwerbs im Erwachsenenalter im Vordergrund des Interesses. Die genannten Einrichtungen haben sich im Jahre 2013 gemeinsam mit weiteren Leibniz-Instituten und universitären Partnern zu einem Forschungsverbund „Bildungspotenziale“ zusammengetan, der mit Beginn des Jahres 2014 um das Leibniz-Institut für Bildungsverläufe (LIfBi) erweitert wird, welches das Nationale Bildungspanel in Bamberg fortführen wird.

Mit dem Forschungsverbund ist eine Struktur entstanden, die es erlaubt, unterschiedliche disziplinäre Perspektiven zusammenzuführen, immer in der Hoffnung, mehr interdisziplinäre Forschungsarbeiten zu realisieren. Solch eine breit aufgestellte Struktur kann nicht nur die internationale Visibilität der deutschen Bildungsforschung stärken, sondern stellt für die Politik einen wichtigen Ansprechpartner bei bildungspolitischen Fragen dar.

6. Trägt die Expansion der empirischen Bildungsforschung Früchte?

In den vergangenen 20 Jahren sind sehr große Summen öffentlicher Forschungsgelder in die empirische Bildungsforschung geflossen. Der daraus resultierende strukturelle Ausbau, von dem vor allem die empirisch arbeitende Erziehungswissenschaft profitiert hat, wurde in diesem Aufsatz ausführlich beschrieben. Der stabile Dialog zwischen Bildungsforschung und Politik wurde oben auch schon hervorgehoben. Wissenschaftlicher Fortschritt lebt allerdings nicht allein von Strukturen und einem guten Verhältnis zwischen Politikern und Wissenschaftlern. Erkenntnisfortschritte ergeben sich nicht automatisch durch die Bereitstellung immer größerer Zuwendungsbeträge. Mancher Kritiker der empirischen Bildungsforschung suggeriert sogar, dass in der empirischen Bildungsforschung viel Geld ohne großen Erkenntnisgewinn verbrannt wird (u. a. Gruschka, 2013). Ziel dieses Abschnitts ist es daher, das Erreichte zu beschreiben und zu würdigen.

6.1 Etablierung eines Beobachtungssystems im Bildungswesen

Infolge der Stärkung der empirischen Bildungsforschung ist in Deutschland ein im internationalen Vergleich relativ schlankes und effektives System der Dauerbeobachtung des Bildungssystems aufgebaut worden. Innerhalb von zehn Jahren wurden mit dem IQB, dem ZIB und der Bildungsberichterstattung am DIPF funktionierende Bildungsmonitoringstrukturen geschaffen, die eine weitgehende Umsetzung der KMK-Gesamtstrategie der Qualitätssicherung (KMK, 2006) erlauben. Hinzu kommen noch die wissenschaftlichen Konsortien unter der Ägide der Technischen Universität Dortmund, welche die internationalen Studien IGLU und TIMSS im Grundschulbereich realisieren. Die zeitliche Taktung dieses Systems der Dauerbeobachtung zeigt die Tabelle 1.

IGLU wiederholt sich alle fünf Jahre, TIMSS alle vier, PISA alle drei und die Bildungsberichterstattung alle zwei Jahre. Da sich die nationalen Ländervergleiche an PISA und IGLU anlehnen, finden sie alle drei bzw. fünf Jahre statt. Zählt man die Zellen in der Tabelle durch, so sind lediglich 12 von 42 besetzt, man wird kaum argumentieren können, dass sich damit ein überladenes System konstituiert hätte. Im Übrigen steht Deutschland mit solch einem Dauerbeobachtungssystem des Bildungswesens nicht allein, sondern stellt sich in eine Reihe mit vielen anderen Industrienationen, in denen erkannt wurde, dass das Wohl jedes einzelnen, aber auch das der gesamten Gesellschaft entscheidend von den Bildungsständen der Bürgerinnen und Bürger abhängt.

Mit der Etablierung des Bildungsmonitorings war bei vielen die Hoffnung verbunden, aus den Studien Steuerungswissen für die Bildungspolitik gewinnen zu können. Studien wie PISA, TIMSS oder IGLU haben aber in erster Linie deskriptiven Charakter – dies gilt auch für die Ländervergleiche des IQB – und geben jeweils zu festen Zeitpunkten Auskunft über den Zustand des Bildungssystems. Sie liefern aber keine Blaupausen zur Verbesserung des Systems. Forschungs- und Entwicklungsarbeiten, die zur

	Jahr						
	2014	2015	2016	2017	2018	2019	2020
IGLU			◆				
TIMSS		◆				◆	
PISA		◆			◆		
LV GS			◆				
LV Sek.		◆			◆		
NBB	◆		◆		◆		◆

IGLU: Internationale Grundschul-Lese-Untersuchung; TIMSS: Trends in International Mathematics and Science Study; PISA: Programme for International Students Assessment; LV GS: Ländervergleich in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe; LV Sek.: Ländervergleich in den Fächern Deutsch, Englisch, Biologie, Physik, Chemie und Mathematik am Ende der 9. Jahrgangsstufe

Tab. 1: Das System des Bildungsmonitorings in der Bundesrepublik Deutschland

Frage „What works“ Auskunft geben können, sehen von ihrer Anlage anders aus; Beispiele für solche Vorhaben mögen die SINUS-Programme sein (zur wissenschaftlichen Begleitung dieser Programme vgl. Fischer, Kobarg, Dalehefte & Trepke, 2013).

6.2 Aufbau von Forschungsinfrastrukturen

Das Bildungsmonitoring generiert große Datensätze mit Leistungs- und Fragebogenmaßen, die in der Regel ein Jahr nach der Berichterstattung für Sekundäranalysen zur Verfügung gestellt werden. Im IQB in Berlin ist bereits 2007 ein Forschungsdatenzentrum (FDZ) eingerichtet worden, in dem die Datensätze der internationalen (PISA, TIMSS, IGLU) und nationalen Schulleistungsstudien (Ländervergleiche in der Grundschule und Sekundarstufe I) archiviert, dokumentiert und für Wissenschaftlerinnen und Wissenschaftler unterschiedlicher Disziplinen bereitgestellt werden. Die Idee, Daten zu generieren und sie dann der *Scientific Community* zur Verfügung zu stellen, wurde im SOEP bereits 1983 umgesetzt. Das 2009 begonnene NEPS stellt ebenfalls solch eine Forschungsinfrastruktur dar, die den Auftrag hat, Daten zu generieren, die von den unterschiedlichen Disziplinen der empirischen Bildungsforschung für statistische Analysen und damit verbunden für die Beantwortung wissenschaftlicher Fragestellungen genutzt werden können. Das große wissenschaftliche Potenzial solcher Datensätze, die über viele Jahre ausgewertet werden können, rechtfertigt sicherlich die mit ihnen verbundenen Kosten. Für das SOEP gilt beispielsweise, dass die Daten mittlerweile weltweit von Wissenschaftlerinnen und Wissenschaftlern genutzt werden. Die ersten vom NEPS generierten Daten werden aktuell im Rahmen eines DFG-Schwerpunktprogramms analysiert, sind aber auch für Forscherinnen und Forscher außerhalb dieses Programms verfügbar.

6.3 Erkenntnisgewinn

Wissenschaft legitimiert sich durch die Generierung von Wissen, das es erlaubt, Phänomene zu beschreiben, zu erklären und vorherzusagen. Die Frage, die sich die empirische Bildungsforschung gefallen lassen muss, ist die, ob wir heute mehr als vor 20 Jahren wissen. Exemplarisch sollen daher im Folgenden Erkenntnisse der letzten 20 Jahre berichtet werden.

Unterrichtsforschung

Dominierte bis in die 1990er-Jahre eine eher psychologische Unterrichtsforschung (z. B. Weinert et al., 1989), welche die Domäne, in der Wissen generiert wird, weitgehend ignorierte, so hat sich in den vergangenen Jahren die Einsicht durchgesetzt, dass empirische Unterrichtsforschung als enge Kooperation zwischen Fachdidaktik, Erziehungswissenschaft und Psychologie durchgeführt werden sollte. Erfolgreiche Projekte hierzu sind die DESI-Studie (Beck & Klieme, 2007) für die Fächer Deutsch und Englisch oder

das COACTIV-Projekt (Kunter et al., 2013) für das Fach Mathematik. Letzteres ist deutlich umfangreicher dokumentiert und zeigt, dass es drei Kerndimensionen gelingenden Mathematikunterrichts sind, die aufseiten der Schülerinnen und Schüler lern- und motivationsfördernd sind: Klassenführung, kognitive Aktivierung und konstruktive Unterstützung (s. hierzu auch Seidel & Shavelson, 2007). Alle drei Dimensionen hängen deutlich mit dem Professionswissen von Lehrkräften zusammen (Kunter et al., 2013). Zu den großen Fortschritten der letzten Jahre zählt, dass es gelungen ist, reliable und valide Instrumente zur Messung des Professionswissens zu entwickeln (vgl. dazu auch Blömeke, Kaiser & Lehmann, 2010) und dieses Professionswissen dann in Zusammenhang mit Unterrichtsvariablen zu bringen. Als besonders fruchtbar hat sich dabei die Analyse von Unterrichtsvideos erwiesen (vgl. Alonzo, Kobarg & Seidel, 2012). Einschränkend muss konzediert werden, dass sich das Gros der Arbeiten auf die Mathematik und die Naturwissenschaften bezieht.

Ebenfalls auf der Basis der Unterrichtsforschung sind bemerkenswerte Materialien und Programme für die schulische Praxis entstanden. Auf SINUS wurde bereits eingegangen. Im Bereich des kompetenzorientierten Unterrichtens sind Ideen für die Unterrichtspraxis publiziert worden, mit deren Hilfe es eher gelingt, Schülerinnen und Schüler kognitiv zu aktivieren und damit nachhaltiges Lernen zu ermöglichen (vgl. z. B. Blum, Drüke-Noe, Hartung & Köller, 2006; Bremerich-Vos, Granzer, Behrens & Köller, 2009; Keller, 2013).

Schließlich muss in diesem Kontext auf die Gründung von *Schools of Education* eingegangen werden, die vielfach als eigene Fakultäten gegründet wurden (z. B. in Berlin, Bochum, München und Wuppertal). An diesen Schools haben sich Modelle der Lehrerbildung etabliert, die eng an aktuelle Befunde der Professions- und Unterrichtsforschung anknüpfen. Besonders prominent ist hier ohne Frage die *School of Education* der Technischen Universität München, die sich der Aus- und Weiterbildung von Lehrkräften an Gymnasien in den MINT-Fächern (Mathematik, Informatik, Naturwissenschaften, Technik) sowie der Aus- und Weiterbildung von Lehrkräften an beruflichen Schulen im gewerblich-technischen Bereich widmet.

Disparitäten im Bildungssystem

Soziale und migrationsbedingte Disparitäten im Bildungssystem beschäftigen die Bildungsforschung seit einigen Jahrzehnten (zu sozialen Disparitäten s. Picht, 1964). Mit den großen Datensätzen der letzten 15 Jahre ist es gelungen, nicht nur Disparitäten empirisch immer wieder zu belegen, sondern auch zu erklären, woher sie rühren. Die wohl einschlägigste Studie dürfte die von Maaz, Baumert, Gresch und McElvany (2010) sein, die aufgedeckt hat, in welchem Ausmaß primäre und sekundäre Disparitäten beim Übergang von der Grundschule in die Sekundarstufe I entstehen. Die Befunde lassen sich dahingehend interpretieren, dass primäre soziale Disparitäten (benachteiligte Kinder weisen schwächere Schulleistungen auf und erhalten daher keine Gymnasialempfehlung) stärker als sekundäre (trotz gleicher Leistungen erhalten sozial benachteiligte Kinder keine Gymnasialempfehlung) sind. Die sekundären Disparitäten scheinen nicht allein dadurch zu entstehen, dass Lehrkräfte sozial benachteiligte Schülerinnen

und Schüler seltener für das Gymnasium empfehlen, sondern sozial benachteiligte Eltern auch selbst bei einer Gymnasialempfehlung für ihr Kind diese Option nicht immer einlösen.

Die Arbeiten von Krolak-Schwerdt (z. B. Krolak-Schwerdt, Böhmer & Gräsel, 2012) zeigen im Übrigen sehr eindrucksvoll auf, wie Informationsverarbeitungsprozesse aufseiten der Lehrkräfte bei der Beurteilung ihrer Schülerinnen und Schüler ablaufen. Gerade bei so wichtigen Fragen wie der Übertrittsempfehlung scheinen Stereotype gegenüber Migrantenkindern oder sozial benachteiligten Kindern zugunsten leistungsbezogener Informationen in den Hintergrund zu treten.

Im Bereich des Kompetenzerwerbs konnten die großen Schulleistungsstudien zeigen, dass die migrationsbedingten Disparitäten weitgehend verschwinden, wenn für Sprachpraxis (wird zuhause Deutsch gesprochen?) und kognitive Grundfähigkeiten kontrolliert wird. Wenn Vorwissensindikatoren vorliegen, so erklären diese in der Regel weitgehend die Disparitäten. Das gilt auch für die Effekte der Klassenzusammensetzung. Der Effekt der sozialen Zusammensetzung auf die Leistungsentwicklung von Schülerinnen und Schülern geht typischerweise gegen null, wenn auf Individual- und Klassenebene das Vorwissen kontrolliert wird (vgl. Dumont, Neumann, Maaz & Trautwein, 2013).

Dass die Größe der Disparitäten erheblich durch das Herkunftsland moderiert wird, zeigen die großen Schulleistungsstudien. Die Disparitäten sind bei den aus Osteuropa Zugewanderten deutlich geringer als bei türkischen Kindern und Jugendlichen (vgl. u. a. Köller, Knigge et al., 2010).

Pädagogische Diagnostik und Methodenforschung

Mit der Gründung des IQB ist eine Professionalisierungswelle im Bereich der Aufgabenentwicklung ausgelöst worden, die zu einem großen Pool fachdidaktisch und empirisch abgesicherter Testitems für unterschiedliche Fächer und Jahrgangsstufen geführt hat (vgl. z. B. Granzer et al., 2009). Hierbei ist bemerkenswert, dass die Aufgabenentwicklung nicht allein in der Hand der Wissenschaft lag und liegt, sondern in enger Kooperation zwischen Lehrkräften, Fachdidaktikern und Psychometrikern geschieht. Das DFG-Schwerpunktprogramm *Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen* hat neben den Arbeiten des IQB erheblich unser Wissen über Kompetenzstrukturen und deren reliable und valide Messung erweitert.

Im Bereich der Methodenforschung sind vor allem die Arbeiten zur adäquaten Modellierung von Kontext und Kompositionseffekten (Lütke et al., 2008; Lütke, Marsh, Robitzsch & Trautwein, 2011) hervorzuheben, die unser Wissen erweitert haben, wie die Daten zum Unterricht und zur Zusammensetzung von Klassen, die in den großen Schulleistungsstudien gewonnen werden, adäquat ausgewertet werden.

Zur Rolle von Institutionen für die kognitive und motivationale Entwicklung

Die längsschnittlichen Studien wie BIJU oder LAU haben unser Wissen über Entwicklungsverläufe unter den institutionellen Bedingungen von Schule erheblich erweitert.

Wir wissen heute, wie groß der Ertrag eines Schuljahres ist ($d = .60$ bis $.90$ in der Primarstufe; $d = .10$ bis $.50$ in der Sekundarstufe I; im Überblick Köller & Baumert, 2012). Wir wissen weiterhin, dass schulische Effekte auf Entwicklungsprozesse im kognitiven Bereich deutlich stärker als im motivationalen, emotionalen oder sozialen Bereich sind (z. B. Baumert & Köller, 1998). Dass Schule nicht nur auf domänenspezifische Kompetenzverläufe wirkt, sondern auch auf kognitive Grundfähigkeiten, konnten Becker, Lüdtke, Trautwein, Köller und Baumert (2012) auf der Basis der BIJU-Studie eindrucksvoll belegen. Dort zeigte sich ein deutlicher Effekt der Schulform auf die Entwicklung der kognitiven Grundfähigkeiten. Schülerinnen und Schüler an Gymnasien gewannen im Laufe von vier Schuljahren gegenüber ihren Kameraden an anderen Schulen rund 7.5 IQ-Punkte mehr ($d = .48$). Schulformen erscheinen so als differenzielle Entwicklungsmilieus.

6.4 Internationalisierung der empirischen Bildungsforschung

Wissenschaft spielt sich heute in einer globalisierten Welt ab, deren gemeinsame Sprache Englisch ist. Diese Erkenntnis hat sich in der naturwissenschaftlichen Forschung seit Langem durchgesetzt. Gleiches gilt für die Psychologie, Ökonomie und große Teile der Soziologie. Das Ziel aller Disziplinen, die empirische Bildungsforschung betreiben, sollte dementsprechend die Partizipation an der internationalen Kommunikation zur Erkenntnismehrung sein. Wer sich dieser Kommunikation verweigert, wird kaum etwas zur Weiterentwicklung der Wissenschaft beitragen können; und wer in Deutschland glaubt, dass Publikationen allein in der Landessprache Wirkungen in der Wissenschaft entfalten, wird feststellen, dass die Zahl der Rezipienten der eigenen Arbeiten eher klein bleibt. Man mag darüber denken, wie man will, es bleibt aber die Botschaft, dass wissenschaftliche Forschung in einer globalisierten Welt kein nationales Unternehmen sein kann. Wissenschaftlerinnen und Wissenschaftler, die sich der internationalen Diskussion verschließen, laufen im Übrigen Gefahr, Standards guter wissenschaftlicher Arbeit nicht zu erreichen.

Die Internationalisierung hat sich je nach Teildisziplin der empirischen Bildungsforschung unterschiedlich schnell vollzogen. Insbesondere für die erziehungswissenschaftliche und fachdidaktische Forschung gilt, dass sie in erheblichen Teilen nach wie vor primär den deutschen Markt bedient. Dies ist auch dem Umstand geschuldet, dass mit deutschsprachigen Publikationen der Durchdringungsgrad in die Praxis höher ist. Allerdings gibt es hier Hinweise auf die zunehmende Internationalisierung. Im Rahmen der großen Schulleistungsstudien sind ohne Frage internationale Erfolge erzielt worden, Forscherinnen und Forscher aus den nationalen PISA-Konsortien haben es geschafft, international Einfluss auf die Weiterentwicklungen in der Anlage und in den Instrumenten von PISA zu nehmen.

Die Visibilität auf internationalen Konferenzen und in internationalen Zeitschriften hat ebenfalls zugenommen. Beispielsweise lag der Anteil der Beiträge mit deutscher Beteiligung in der Zeitschrift *Learning and Instruction* im Jahr 2013 bei rund

25 Prozent, so hoch wie nie zuvor. Eine Analyse der 25 am häufigsten zitierten Artikel aus *Learning and Instruction* (Stand: September 2013) weist aus, dass immerhin sechs dieser *top-cited papers* unter deutscher Beteiligung zustande gekommen sind. Bemerkenswert ist weiterhin die von Jones et al. (2010) vorgestellte Produktivitätsanalyse der Pädagogischen Psychologie. Die Autoren werteten die wichtigsten englischsprachigen pädagogisch-psychologischen Zeitschriften im Zeitraum zwischen 2003 und 2008 aus. In der Liste der weltweit zehn produktivsten Forscher tauchten immerhin drei Deutsche auf.

Wertet man den *Social Science Citation Index* (SSCI) aus, so lässt sich dort als Maß, das zwischen der Zahl der Publikationen und der Zitationsquote der Papiere relativiert, der Hirsch-Index (Hirsch, 2005) analysieren. Liegt der Hirsch-Index einer Wissenschaftlerin/eines Wissenschaftlers bei 10, so gilt für zehn der von ihr/ihm publizierten Artikel, dass jeder mindestens zehnmal zitiert wurde, die übrigen seltener. Liegt der Hirsch-Index bei 0, so kann dies zwei Gründe haben: (1) Eine Wissenschaftlerin/ein Wissenschaftler hat noch keine Arbeit in einer SSCI-Zeitschrift publiziert. (2) Ihre oder seine in einer SSCI-Zeitschrift publizierte Arbeit wurde noch nicht von anderen Autoren, die in SSCI-Zeitschriften publizieren, zitiert. Mittlerweile erreicht eine kleine Gruppe von empirisch arbeitenden deutschen Erziehungswissenschaftlern, deren Arbeiten in den letzten 20 Jahren entstanden sind, Hirsch-Indizes, die zwischen 20 und 30 liegen, d. h. in dieser Gruppe sind (a) viele international publizierte Arbeiten entstanden, die (b) auch die Schwelle zur Visibilität überschreiten und weltweit gelesen und zitiert werden. Nicht unerwähnt bleiben darf, dass die große Mehrzahl deutscher empirisch arbeitender Erziehungswissenschaftler gar nicht oder mit Hirsch-Indizes unter 5 im SSCI auftaucht.

7. Probleme und offene Forschungsflanken

Bei allen Erfolgen der letzten 20 Jahre bleiben viele Fragen bzw. Herausforderungen, denen sich die empirische Bildungsforschung stellen muss. Diese betreffen politische, strukturelle und inhaltliche Bereiche. Im politischen Bereich ist es zu einem gewissen Schulterschluss zwischen empirischer Bildungsforschung und Bildungspolitik gekommen, gleichzeitig hat die Bildungsforschung lernen müssen, dass die Politik die Befunde selten so auslegt, wie sie sind, sondern so, wie sie in die politische Agenda passen. Somit besteht die Gefahr, dass die Bildungsforschung nicht-intendiert zum Steigbügelhalter politischer Kräfte wird, die unter dem scheinbaren Schutz wissenschaftlicher Erkenntnis politische Interessen durchsetzen. Die empirische Bildungsforschung mag aus diesem Dilemma am einfachsten herauskommen, wenn sie ihre Ergebnisse primär in wissenschaftlichen Zeitschriften dem wissenschaftlichen Diskurs aussetzt und nicht in erster Linie danach strebt, die Feuilletons der großen Tageszeitungen zu füllen.

Im strukturellen Bereich ist die Frage unbeantwortet geblieben, worin die breite Förderung von Nachwuchsprogrammen für Doktoranden und Postdoktoranden münden soll. Mit großem finanziellen Aufwand werden Generationen von Nachwuchs-

wissenschaftlerinnen und -wissenschaftlern produziert, für die es perspektivisch keine Anschlussstellen gibt, weder als Senior-Researcher in außeruniversitären Forschungseinrichtungen noch als Professorinnen/Professoren im universitären Bereich. Hier werden die Vertreter der empirischen Bildungsforschung Konzepte entwickeln müssen, die alternative berufliche Einfädelungen für den Nachwuchs jenseits der Forschung möglich machen. Weiterhin lassen die großen Zahlen an Qualifikationsarbeiten die Frage aufkommen, ob und in welcher Form ihre wissenschaftliche Qualität gesichert werden kann.

Das viele Geld für die empirische Bildungsforschung wirft im Übrigen nicht nur die Frage nach der Qualität der Qualifikationsarbeiten, sondern auch nach der generellen Qualität der beantragten Forschungsprojekte auf. Bei knappen Forschungsgeldern setzen sich in der Regel die besten Projektanträge im Wettbewerb durch. Bei großen Geldmengen ist zumindest zu befürchten, dass auch solche Projekte gefördert werden könnten, die vergleichsweise geringe Erträge versprechen. Möglicherweise ist eine Konsolidierung der Forschungsförderung für die empirische Bildungsforschung zielführender als eine weitere Ausdehnung der Fördergelder.

Inhaltlich besteht ein Problem in der besonderen Schwerpunktsetzung bei der Kompetenzmessung. Es gab sicherlich gute Gründe, das IQB einzurichten und so eine Infrastruktur zu schaffen, die für die Testentwicklung im Rahmen des Nationalen Assessments zuständig ist (s. o.). Daneben sind viele weitere BMBF- und DFG-Projekte gefördert worden, die über die Testentwicklung mit knapper Validierung kaum hinauskommen. So wichtig die pädagogisch-psychologische Diagnostik auch ist, sie löst keine theoretischen und praktischen Probleme im Kontext Schule. Bestenfalls liefert sie Werkzeuge, um substanzielle Forschung zu betreiben. Die stärkere Hinwendung von methodischen zu substanziellen Themen ist demnach ein Kerndesiderat für die zukünftige Forschung. Exemplarisch soll hier auf vier solcher zukünftigen Felder eingegangen werden: (1) die fachdidaktische Forschung, (2) Forschung im vorschulischen Bereich, (3) Implementationsforschung und (4) Hochschulforschung.

Die fachdidaktische Forschung hat seit der ersten DFG-Initiative für die naturwissenschaftlichen Didaktiken Anfang der 1990er-Jahre einen deutlichen Aufschwung genommen, der oben berichtet wurde. Bis heute ist sie aber auf die Mathematik und die drei Naturwissenschaften Biologie, Chemie und Physik beschränkt geblieben. Sieht man einmal von der DESI-Studie (Beck & Klieme, 2007) ab, so sind die geisteswissenschaftlichen Didaktiken weitgehend unsichtbar geblieben. Kooperationsvorhaben mit Erziehungswissenschaftlern oder Psychologen, die einen empirischen Zugang zur Unterrichtsforschung erleichtern könnten, fehlen, und drittmittelgeförderte Projekte in größerem Umfang sind seltene Ereignisse. Hier wird es ohne Frage in den nächsten Jahren darauf ankommen, Förderinitiativen zu starten, um unser Wissen über gelingende Unterrichtsprozesse in den geisteswissenschaftlichen Fächern zu erweitern.

Ein zweiter Bereich, der mit der Einrichtung der BIKS-Forschergruppe und dem Nationalen Bildungspanel in Angriff genommen wurde, aber weiter ausgebaut werden kann, ist die Forschung zu Effekten vorschulischer Erziehungs- und Bildungsangebote auf die soziale und kognitive Entwicklung von Kindern. Dass soziale Disparitäten

im kognitiven und sozialen Leistungsstand bereits lange vor institutionellen Bildungsangeboten (beispielsweise in Kindertageseinrichtungen) auftreten, konnte u. a. in der U.S.-amerikanischen *Early Childhood Longitudinal Study – Birth Cohort* (ECLS-B) eindrucksvoll nachgewiesen werden. Bereits mit neun Monaten traten signifikante Unterschiede in Abhängigkeit vom familiären Hintergrund auf, die im Alter von zwei Jahren noch deutlich zugenommen hatten (Halle et al., 2009). Ähnliche Befunde ergaben sich in der in Großbritannien durchgeführten *Millenium Cohort Study* (MCS; Hansen & Joshi, 2007). Wie, d. h. mit welchen Interventionen man solchen früh auftretenden Disparitäten begegnen kann, ist in Deutschland bisher kaum untersucht, auch die Frage nach der langfristigen Wirksamkeit vieler existierender Programme ist nicht beantwortet, es scheint aber so zu sein, dass Einzelmaßnahmen kaum Nachhaltigkeit garantieren: „Nicht selten findet man, dass ‚erfolgreiche‘ Frühförderung zwar unmittelbar nach Abschluss der Fördermaßnahmen zu nachweislich positiven Effekten geführt hat, aber einige Monate nach Beendigung der Förderung sich der alte Entwicklungsrückstand – oder gar ein noch größerer – einstellt“ (Hasselhorn, 2010, S. 170).

Im Übrigen beschränkt sich das Manko der fehlenden Untersuchungen zur Wirksamkeit von Programmen nicht auf den vorschulischen Bereich. Vielmehr mangelt es generell an entsprechenden Studien in allen Altersgruppen. Die Frage: „What works?“, d. h. welche Interventionsmaßnahmen lassen sich in ökologisch validen Settings etablieren, und welche Person- und Interventionsmerkmale erhöhen die Chance der erfolgreichen Implementation, ist vielfach unbeantwortet.

Schließlich sei auf die Hochschulforschung hingewiesen. Zwar existiert das Hochschulinformationssystem (HIS) seit über 40 Jahren, hinsichtlich der Wissenserwerbsprozesse im Kontext der Hochschulforschung wissen wir in Deutschland allerdings sehr wenig, nicht zuletzt auch deshalb, weil Instrumente zur Erfassung der im Studium erworbenen Kompetenzen bzw. Wissensbestände fehlen. Das BMBF-Förderprogramm „Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor“ (KoKoHs) ist ein erster Versuch, hier Abhilfe zu schaffen. Erste längsschnittliche Untersuchungen zu Studienverläufen existieren inzwischen, so das „Panel zum Lehramtsstudium“ (PaLea; Bauer et al., 2010), aber auch dort fehlen fachspezifische Leistungstests, um den universitären Wissenserwerb besser zu verstehen.

8. Resümee

Die Expansion der empirischen Bildungsforschung hat in den letzten 20 Jahren zu erheblichen Veränderungen im Wissenschafts- und Hochschulsystem geführt. Die empirische Bildungsforschung wurde mit riesigen Fördersummen der DFG und des BMBF an Universitäten gestärkt, teilweise mit erheblichen Kosten für die nicht empirisch ausgerichteten Subdisziplinen der Erziehungswissenschaft. Mit der Expansion verbunden waren und sind die großen nationalen und internationalen Schulleistungsstudien (Large-scale Assessments), die nicht nur wichtige Informationen über den Status und die Entwicklung unseres Bildungssystems liefern, sondern auch Ausdruck des Dialo-

ges zwischen Wissenschaft und Politik sind. Passend zur Philosophie der *Large-scale Assessments* ist auch der gegenwärtige Trend, große Forschungsverbünde zu etablieren. Die Wissenschaftspolitik verspricht sich durch solche „Forschungstanker“ größere wissenschaftliche Erfolge, für die empirische Bildungsforschung ist es allerdings eine offene Frage, ob sich diese Philosophie bewähren wird (s. hierzu die sehr kritischen Anmerkungen bei Gruschka, 2013). Aber selbst wenn man skeptisch ist, sind die Erfolge der empirischen Bildungsforschung seit Beginn der 1990er-Jahre bemerkenswert. Die große Zahl theoretisch wie praktisch relevanter Ergebnisse der letzten 20 Jahre ebenso wie die verstärkte Internationalisierung der empirischen Bildungsforschung lassen hoffen, dass sich die Expansion am Ende als Erfolgsgeschichte herausstellen wird.

Literatur

- Alonzo, A. C., Kobarg, M., & Seidel, T. (2012). Pedagogical content knowledge as reflected in teacher-student interactions: Analysis of two video cases. *Journal of Research in Science Teaching*, 49, 1211–1239.
- Autorengruppe Bildungsberichterstattung (2012). *Bildung in Deutschland 2012. Ein indikatoren-gestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*. Bielefeld: Bertelsmann.
- Bauer, J., Drechsel, B., Retelsdorf, J., Kauper [Sporer], T., Rösler, L., Prenzel, M., & Möller, J. (2010). Panel zum Lehramtsstudium – PaLea: Entwicklungsverläufe zukünftiger Lehrkräfte im Kontext der Reform der Lehrerbildung. *Beiträge zur Hochschulforschung*, 32, 34–55.
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J., & Weiß, M. (Hrsg.) (2002). *PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich*. Opladen: Leske + Budrich.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., & Weiß, M. (Hrsg.) (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Baumert, J., & Köller, O. (1998). Nationale und internationale Schulleistungsstudien: Was können sie leisten, wo sind ihre Grenzen. *Pädagogik*, 50, 12–18.
- Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller, O., & Neubrand, J. (1997). *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich*. Opladen: Leske + Budrich.
- Beck, B., & Klieme, E. (Hrsg.). *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistung International)*. Weinheim/Basel: Beltz.
- Becker, M., Lüdtke, O., Trautwein, U., Köller, O., & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology*, 104, 682–699.
- BLK (1997) = Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (Hrsg.) (1997). *Gutachten zur Vorbereitung des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“* (Materialien zur Bildungsplanung und zur Forschungsförderung, Heft 60). Bonn: BLK.
- Blömeke, S., Kaiser, G., & Lehmann, R. (Hrsg.) (2010). *TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich*. Münster: Waxmann.
- Blossfeld, H.-P., von Maurice, J., & Schneider, T. (2009). Das Nationale Bildungspanel: Eine neue Datenbasis für die empirische Bildungsforschung. *Zeitschrift für Soziologie*, 38, 331–338.

- Blum, W., Drücke-Noe, K., Hartung, R., & Köller, O. (Hrsg.) (2006). *Bildungsstandards Mathematik konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsideen und Fortbildungsmöglichkeiten*. Berlin: Cornelsen Scriptor.
- BMBF (2007) = Bundesministerium für Bildung und Forschung (2007). *Rahmenprogramm zur Förderung der empirischen Bildungsforschung*. Bonn: BMBF.
- 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.
- Bremerich-Vos, A., Granzer, D., Behrens, U., & Köller, O. (2009). *Bildungsstandards für die Grundschule: Deutsch konkret: Aufgabenbeispiele – Unterrichts Anregungen – Fortbildungsideen*. Berlin: Cornelsen Scriptor.
- Deutscher Bildungsrat (1974). *Aspekte für die Planung der Bildungsforschung. Empfehlungen der Bildungskommission* (Bd. 1). Stuttgart: Klett.
- DFG (2002) = Deutsche Forschungsgemeinschaft (2002). Stellungnahme zur strukturellen Stärkung der empirischen Bildungsforschung. Ausschreibung von Forschungsgruppen in der Empirischen Bildungsforschung. *Zeitschrift für Pädagogik*, 48(5), 786–798.
- Ditton, H. (2011). Entwicklungslinien der empirischen Bildungsforschung. Vom Bildungsrat zu aktuellen Themen. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Strukturen und Methoden* (S. 29–42). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dumont, H., Neumann, M., Maaz, K., & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen: Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60, 163–183.
- Elley, W. B. (1992). *How in the World Do Students Read? IEA Study of Reading Literacy*. The Hague: IEA.
- Fend, H. (1991). Schule und Persönlichkeit: Eine Bilanz der Konstanzer Forschungen zur „Sozialisation in Bildungsinstitutionen“. In R. Pekrun & H. Fend (Hrsg.), *Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung* (S. 9–32). Stuttgart: Enke.
- Fischer, C., Kobarg, M., Dalehefte, I. M., & Trepke, F. (2013). Ein Unterrichtsentwicklungsprogramm wissenschaftlich begleiten. *Psychologie in Erziehung und Unterricht*, 60, 26–31.
- Fleischer, J., Leutner, D., & Klieme, E. (2012). Modellierung von Kompetenzen im Bereich der Bildung. Eine psychologische Perspektive. *Psychologische Rundschau*, 63, 1–2.
- Granzer, D., Köller, O., Bremerich-Vos, A., et al. (2009). *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule*. Weinheim/Basel: Beltz.
- Gräsel, C. (2011). Was ist empirische Bildungsforschung? In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Strukturen und Methoden* (S. 13–27). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gruschka, A. (2013). Empirische Bildungsforschung am Ausgang ihrer Epoche? *Profil*, 6/2013, 1–6.
- Halle, T., Forry, N., Hair, E., Perper, K., Wandner, L., Wessel, J., & Vick, J. (2009). *Disparities in early language development: Lessons from the Early Childhood Longitudinal Study – Birth Cohort (ECLS-B)*. Washington: Child Trends.
- Hansen, K., & Joshi, H. (2007). *Millennium Cohort Study. Second survey. A user's guide to initial findings*. London: Institute of Education, Centre for Longitudinal Studies.
- Hasselhorn, M. (2010). Möglichkeiten und Grenzen der Frühförderung aus entwicklungspsychologischer Sicht. *Zeitschrift für Pädagogik*, 56(2), 168–177.
- Heckhausen, H. (1980). *Motivation und Handeln*. Berlin: Springer.
- Helmke, A., & Jäger, R. S. (2001). Die Evaluationsstudie MARKUS. *Empirische Pädagogik*, 15(4), 507–512.

- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569–16572.
- Jones, S. J., Fong, C. J., Torres, L. G., Yoo, J. H., Decker, M. L., & Robinson, D. H. (2010). Productivity in educational psychology journals from 2003 to 2008. *Contemporary Educational Psychology*, 35, 11–16.
- Jones, L. V., & Olkin, I. (2004). *The Nation's Report Card: Evolution and Perspectives*. Bloomington: Phi Delta Kappa Educational Foundation.
- Keller, S. (2013). *Kompetenzorientierter Englischunterricht*. Berlin: Cornelsen.
- KMK (2006) = Kultusministerkonferenz (Hrsg.) (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: Luchterhand.
- Köller, O., & Baumert, J. (2012). Schulische Leistungen und ihre Messung. In W. Schneider & U. Lindenberger (Hrsg.), *Entwicklungspsychologie* (7. Aufl., S. 645–661). Weinheim/Basel: Beltz/PVU.
- Köller, O., Baumert, J., Cortina, K. S., & Trautwein, U. (2010). Bildungsverläufe und psychosoziale Entwicklung im Jugendalter und jungen Erwachsenenalter. In C. Spiel, B. Schober, P. Wagner & R. Reimann (Hrsg.), *Bildungspsychologie* (S. 245–252). Göttingen: Hogrefe.
- Köller, O., Knigge, M., & Tesch, B. (Hrsg.) (2010). *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann.
- Kroh, M. (2013). *SOEP 2011 – Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2011)* (SOEP Survey Papers 140: Series C). Berlin: DIW/SOEP.
- Krolak-Schwerdt, S., Böhmer, M., & Gräsel, C. (2012). Leistungsbeurteilungen von Schulkindern: Welche Rolle spielen Ziele und Expertise der Lehrkraft? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44, 111–122.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105, 805–820.
- Lehmann, R. H., & Peek, R. (1997). *Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen*. Hamburg: Hamburger Schulbehörde.
- Lüdtke, O., Marsh, H. W., Robitzsch, A., & Trautwein, U. (2011). A 2 x 2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error correction models. *Psychological Methods*, 16, 444–467.
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. O. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13, 203–229.
- Maaz, K., Baumert, J., Gresch, C., & McElvany, N. (2010). *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten*. Bonn: BMBF.
- Mayer, K. U. (Hrsg.) (1990). Lebensverläufe und sozialer Wandel. *Sonderheft Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 31.
- Picht, G. (1964). *Die deutsche Bildungskatastrophe – Analyse und Dokumentation*. Freiburg: Walter.
- Popper, K. R. (1963). *Conjectures and Refutations*. London: Routledge.
- Prenzel, M., & Allolio-Näcke, L. (2006). *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms*. Münster: Waxmann.
- Roth, H. (1963). Die realistische Wendung in der Pädagogischen Forschung. *Die Deutsche Schule*, 55, 109–119.
- Seidel, T., & Shavelson, R. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis research. *Review of Educational Research*, 77, 454–499.

Stanat, P., Pant, H. A., Böhme, K., & Richter, D. (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.

Weinert, F. E., Schrader, F.-W., & Helmke, A. (1989). Quality of instruction and achievement outcomes. *International Journal of Educational Research*, 13, 895–914.

Anschrift des Autors

Prof. Dr. Olaf Köller, Leibniz-Institut für die Pädagogik der Naturwissenschaften
und Mathematik, Olshausenstraße 62, 24118 Kiel, Deutschland

E-Mail: koeller@ipn.uni-kiel.de