

Mehr Transparenz in der Bildungspolitik

Regina T. Riphahn und Ludger Wößmann

Eine rational gestaltete Politik nutzt alle zur Verfügung stehenden Informationsquellen, um zu bestmöglichen Entscheidungen zu gelangen. Dies geschieht derzeit in der deutschen Bildungspolitik nicht. Daher hat der Wissenschaftliche Beirat beim Bundesministerium für Wirtschaft und Energie (2016) ein Gutachten zum Thema "Mehr Transparenz in der Bildungspolitik" verfasst, das wir hier überblicksartig vorstellen möchten.¹

Die Bildung der Bevölkerung und die Fachkräftesicherung sind - besonders im demographischen Wandel - von zentraler Bedeutung, um langfristig Wirtschaftskraft und Wohlstand zu gewährleisten. Hierfür ist ein effektives Bildungswesen unverzichtbar. Die Basis für die Qualifikation von Fachkräften wird im allgemeinbildenden Schulsystem gelegt. Das deutsche Schulsystem ist föderal, d.h. auf der Ebene der Bundesländer, organisiert. Während die Heterogenität und Zersplitterung der bundeslandspezifischen Bildungsregularien häufig als Nachteil betrachtet werden, birgt die föderale Organisation auch Chancen: Sie erlaubt, in Bezug auf bildungspolitische Ziele und Mittel verschiedene Lösungen zu erproben. Länder können sich und ihre bildungspolitischen Strukturen und Maßnahmen vergleichen und voneinander lernen.

Mit der politischen Gestaltungsfreiheit ist stets die Verantwortung für die Ergebnisse verknüpft: Mindestens im Nachhinein müssen Wählerinnen und Wähler in die Lage versetzt werden, staatliches Handeln und seine Ergebnisse zu beurteilen. Die Bereitstellung von Information trägt hier zusätzlich dazu bei, Erfahrungen mit Modell- und Pilotvorhaben nutzbar zu machen: Der Vergleich zwischen den Bundesländern kann zur Ergebnisverbesserung bildungspolitischen Handelns beitragen. Die deutschen Bundesländer haben durch das föderale System - im Gegensatz zu Nationen mit einheitlichen Bildungssystemen - die Chance, Vorteile aus dem Erfahrungsaustausch und dem Vergleich zwischen Gebietskörperschaften zu ziehen

¹ Einzelne Passagen dieses Textes sind aus dem Gutachten des Wissenschaftlichen Beirats 'Mehr Transparenz in der Bildungspolitik' übernommen worden, der von Regina T. Riphahn (Federführung), Christoph Engel, Hans Gersbach, Eckhard Janeba und Ludger Wößmann vorbereitet wurde (Wissenschaftlicher Beirat, 2016).

und solche bildungspolitischen Lösungen zu wählen, die sich andernorts bereits als effektiv und effizient erwiesen haben.

Fehlt jedoch Transparenz in Bezug auf die Ergebnisse der Bildungspolitik, so hat dies vielfältige Nachteile. Erstens kann die Bildungspolitik nicht evidenzbasiert weiterentwickelt werden. Zweitens fehlt allen Beteiligten, also bspw. Eltern, Schülern, Arbeitgebern und Hochschulen, eine verlässliche Informationsbasis für Ausbildungs-, Einstellungs- und Studienzugangsentscheidungen. Ein transparenter Ausweis der Erfolge von Bildungspolitik und ihre Vergleichbarkeit über die Bundesländer hinweg könnte somit auf vielen Ebenen Nutzen stiften.

In der Arbeitsmarktpolitik hat die Bereitstellung von aussagekräftigen Daten zu eindeutiger Evidenz über die Funktionsweise einzelner arbeitsmarktpolitischer Maßnahmen beigetragen. Nachdem die Daten zu Arbeitsförderungsmaßnahmen der wissenschaftlichen Analyse zugänglich gemacht wurden, konnten teure und ineffektive Instrumente identifiziert und durch wirksamere Maßnahmen ersetzt werden (vgl. BMAS und IAB 2011). Die Arbeitsmarktpolitik wurde damit effektiver und kostengünstiger. In ähnlicher Form könnten Bildungspolitik und ihre Instrumente (z.B. G8 vs. G9, Ganztagschule, Gemeinschaftsschule, Klassengröße, Zentralabitur, Inklusion, Sprachförderung, Unterrichtsgestaltung) empirisch analysiert werden, damit ineffektive Maßnahmen als solche erkannt und durch effektivere ersetzt werden können.

Im nächsten Abschnitt diskutieren wir die Theorie des föderalen Wettbewerbs von Gebietskörperschaften, wonach durch die Vergleichbarkeit der Ergebnisse in einzelnen Politikfeldern ein Wettbewerb ausgelöst werden kann. Dies lässt sich auf die Bundesländer und ihre Bildungspolitiken übertragen. Anschließend beschreiben wir die Ausgangslage in Deutschland im internationalen Vergleich. Auf dieser Basis werden die Vor- und Nachteile von mehr Transparenz diskutiert und ein Fazit gezogen.

Theorie des föderalen Wettbewerbs von Gebietskörperschaften

Die Bedeutung der Bereitstellung von Informationen lässt sich im Rahmen der Theorie des föderalen Wettbewerbs von Gebietskörperschaften verdeutlichen. Die Theorie geht davon aus, dass das Wahlvolk wegen mangelnder Information die Qualität der Entscheidungen der von ihm gewählten Politikerinnen und Politiker nicht wirksam überwachen kann. In dieser Situation können Wählerinnen und Wähler aus dem

Vergleich mit den Politikergebnissen benachbarter Gebietskörperschaften Rückschlüsse auf die Leistungen der Politik ziehen, so dass in der Politik ein Wettbewerb anhand von Vergleichsmaßstäben (*Yardstick Competition*) entsteht (vgl. Besley und Case 1995).

Wenn Information zu Politikergebnissen vorliegt, kann der Wettbewerb zwischen Gebietskörperschaften auch bei asymmetrisch verteilter Informationen zwischen Wahlvolk und Politik das Verhalten der Politik beeinflussen. Sofern Rückschlüsse auf die Leistung von Politikern in wichtigen Politikbereichen möglich sind, kann sich durch gute Ergebnisse in diesem Bereich die Chance auf eine Wiederwahl erhöhen. Die Bereitstellung von Information über Politikergebnisse gibt Politikern Anreize, ihre Reputation im betreffenden Feld zu verbessern. Letztlich verbessert die erhöhte Messbarkeit die Leistungen der Politiker. Es liegen viele empirische Studien vor, die diese positiven Effekte von Wettbewerb zwischen Gebietskörperschaften dokumentieren. Insgesamt kann institutioneller Wettbewerb zwischen Gebietskörperschaften die Effizienz der Bereitstellung von Dienstleistungen steigern.

In diesem Sinne kann *Yardstick Competition* einen institutionellen Wettbewerb intensivieren, wenn relevante und vergleichende Informationen über Politikergebnisse verfügbar sind, die politischen Entscheidungen dezentral getroffen werden und demokratische Politikverfahren vorherrschen. Je besser das Wahlvolk informiert ist, umso eher ist zu erwarten, dass sich Politiker für bessere Ergebnisse einsetzen.

Die Bereitstellung von Daten ist Voraussetzung dafür, dass Erfolge und Misserfolge von bildungspolitischen Maßnahmen sichtbar werden. In einem föderalen System ist der Nutzen solcher Daten besonders groß, da sie vielfältige Vergleichsmöglichkeiten und Anregungen liefern und Verbesserungen von bereits stattfindenden Vorgängen ermöglichen.

Informationsstand und Steuerung des deutschen Bildungswesens im internationalen Vergleich

Im internationalen Vergleich gehen Nationen unterschiedlich vor, um aus ihren Erfahrungen zu lernen und ihre Bildungssysteme zu verbessern:

In den skandinavischen Ländern können umfangreiche Registerdaten genutzt werden. Der Zugang zu diesen Daten und ihre Verknüpfung mit Informationen über Bildungseinrichtungen ermöglichen eine umfassende Erforschung der Wirkung

bildungspolitischer Maßnahmen. Diese Ergebnisse fließen in die bildungspolitische Beratung ein (vgl. Figlio et al. 2016).

Im Vereinigten Königreich wurde im letzten Jahrzehnt ein Datensatz aufgebaut, der die extern getesteten Leistungen aller Schülerinnen und Schüler auf Schlüsselstufen mit zahlreichen Hintergrundinformationen verbindet. Dabei ist eine Beobachtung individueller Leistungsentwicklungen über die Zeit sowie eine Verknüpfung mit Schul- und Regionaldaten möglich. So konnten bereits zahlreiche Modellversuche und Bildungsreformen evaluiert und die gewonnenen Erkenntnisse in die Politikberatung eingespeist werden (vgl. Machin 2014).

In den USA wurden in vielen Bundesstaaten Datensätze aufgebaut, die die jährliche Leistungsentwicklung aller Schülerinnen und Schüler in öffentlichen Schulen individuell abbilden. Dies erlaubt umfangreiche Forschung zu den Effekten von Bildungsmaßnahmen.

Im Gegensatz dazu behindern die Kultusbehörden in Deutschland seit vielen Jahren die systematische Erforschung von bundeslandspezifischen Bildungsinitiativen und die Bereitstellung von Vergleichsgrößen, indem sie den Zugang zu relevanten Informationen verweigern: Vorhandene Information wird nicht bekannt gemacht und die Erhebung neuer Daten wird eingeschränkt. Wir benennen drei Beispiele.

(a) In den ersten drei Wellen des PISA-Tests wurde eine (große) Stichprobe (PISA-E, als deutsche Ergänzung des Standardprogramms) gezogen, die für jedes Bundesland repräsentativ war. Dies ermöglichte einen direkten Ländervergleich und hat zahlreiche politische Impulse ausgelöst (vgl. z.B. Wößmann 2007). Dieser Vergleich der Bundesländer ist in PISA seither nicht mehr möglich. Seit 2008/09 wurde ein auf den neu entwickelten nationalen Bildungsstandards basierender Ländervergleichstest eingeführt, der in der Grundschule alle fünf Jahre und in der Sekundarstufe I alle drei Jahre durchgeführt werden soll, im letzteren Fall allerdings mit abwechselnder fachlicher Ausrichtung. Der Bruch in den bisherigen Erhebungen und die zukünftig nur im Abstand von fünf bzw. sechs Jahren möglichen Vergleiche werden eine zielgerichtete Analyse möglicher Wirkungen länderspezifischer Bildungsmaßnahmen behindern.

(b) Man kann die Bildungspolitiken von Bundesländern nur vergleichen, wenn befragte Personen einem Bundesland zugeordnet sind. Dies ist im internationalen PISA-

Datensatz unmöglich, in den alten PISA-E-Datensätzen (mittlerweile) unter unnötig erschwerten Bedingungen teilweise möglich und im Nationalen Bildungspanel (National Educational Panel Study, NEPS) erst seit kurzem überhaupt und nur eingeschränkt möglich. Dies ist bemerkenswert, da die Landeszugehörigkeit eines Schülers in den Rohdaten natürlich enthalten ist. Die Information wird aus dem Datensatz bewusst entfernt, um Ländervergleiche unmöglich zu machen. Auf der einseitigen Leitlinie, wie die Verwendung von Länderkennungen im NEPS-Datensatz nun möglich ist, findet sich nicht weniger als dreimal der Satz: „Eine Identifikation einzelner Bundesländer in Ergebnisdarstellungen ist nicht gestattet.“

Die Bedeutung des Zugangs zu Bundeslandkennungen wird in der Stellungnahme des Wissenschaftsrats zum NEPS aus dem Jahr 2013 eingängig verdeutlicht: „Für wissenschaftliche Fragestellungen und die Entwicklungen von komplexen Modellen zur Datenanalyse sind Informationen zu länderspezifischen Besonderheiten jedoch essentiell. (...) Für evidenzbasierte Forschung und Politikberatung sind empirische Bildungsdaten einschließlich regionaler Strukturdaten unbedingt erforderlich.“ (Wissenschaftsrat 2013, S.17).

(c) Im Jahr 2003 hat die Kultusministerkonferenz die verbindliche Einführung eines Schülerkerndatensatzes in den Bundesländern beschlossen, auch „um wichtige Informationslücken zu wesentlichen nationalen und internationalen Fragestellungen zu schließen (...). Deutschland kann bereits seit einigen Jahren internationale Datenanforderungen nicht vollständig erfüllen, wie es von internationalen Organisationen wie der OECD regelmäßig festgestellt wird (...)“ (KMK 2011). Dennoch liegt bis heute der so beschlossene Datensatz weder zu Verwaltungs- noch zu Forschungszwecken vor.

Durch die systematische Verweigerung des Zugangs zu bestehenden Bildungsinformationen und die Beschränkungen bei der Erhebung neuer Daten durch die Länder ist es schwierig, die Qualität des Bildungssystems zu beurteilen, und es ist unmöglich, die eigentlichen Vorteile im Wettbewerb des Föderalismus zu nutzen. Effizienzgewinne aus *Yardstick Competition* zwischen den Bundesländern bleiben unrealisiert. Der Wettbewerb der Politik um Wählerstimmen erfolgt ohne solide Informationsgrundlage.

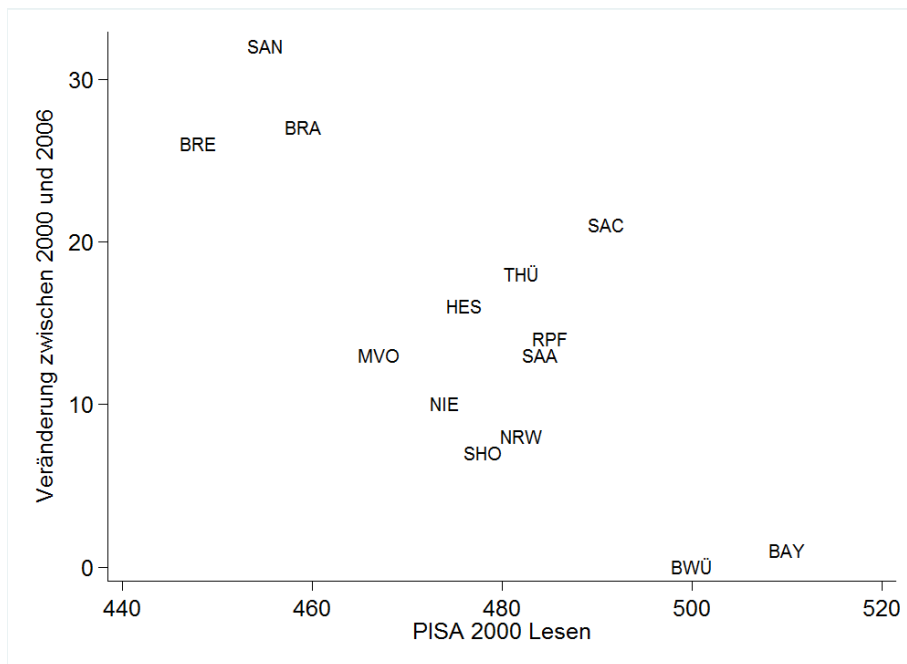
Gleichzeitig entstehen auf diese Weise erhebliche Nachteile für die Evaluation spezifischer Instrumente wie bspw. Sprachförderungsmaßnahmen, Inklusion oder

Sekundarschuldauer. Erkenntnisse über die Wirksamkeit dieser Instrumente, ihre Bedeutung für unterschiedliche Schülergruppen und Vergleiche relevanter Kosten-Nutzen-Relationen können nicht gewonnen werden. Während andere Nationen der Wissenschaft und Politikberatung detaillierte Schülerdaten angereichert durch regionale Informationen bereitstellen, wird vergleichbare Information in Deutschland von Seiten der Kultusbehörden nicht zur Verfügung gestellt, obwohl sie zum Teil bereits vorhanden ist.

Evidenz zur Auswirkung der Bereitstellung von Information im Bildungssystem

Die Leistungsentwicklung der Schülerinnen und Schüler in deutschen Bundesländern in den Jahren, in denen die PISA-E-Daten einen Bundesländervergleich ermöglichten (2000, 2003 und 2006), legt nahe, dass die Verfügbarkeit vergleichender Leistungsinformationen das Verhalten der Bundesländer in der Tat beeinflussen kann. Gerade Bundesländer, die bei den Tests 2000 schlecht abgeschnitten haben (Bremen, Sachsen-Anhalt, Brandenburg), hatten sich im letzten PISA-E-Bundesländervergleich 2006 am meisten verbessert, während sich das Leistungsniveau der Spitzenreiter nicht signifikant verändert hat (vgl. Abbildung). Aufgrund der Abschaffung des PISA-E-Tests auf Bundeslandebene lässt sich die weitere Entwicklung seither leider nicht beurteilen. Aber das Bild der Jahre 2000-2006 deckt sich mit der Interpretation, dass der durch die erstmalige Veröffentlichung der PISA-E-Ergebnisse von 2000 ausgelöste Wettbewerbsdruck aus *Yardstick Competition* zwischen den Bundesländern zu deutlichen Leistungsverbesserungen geführt hat.

Abbildung: Leistungsniveau und Leistungsveränderung der Schülerinnen und Schüler nach Bundesländern im PISA-E-Test, 2000-2006



Hinweise: Durchschnittliche Leistungen der Bundesländer im PISA-E-Lesetest 2000 gemessen in PISA Kompetenzpunkten sowie Leistungsveränderung in diesem Test bis 2006. Quelle: Wößmann (2015) anhand der PISA-E-Daten.

Auch die in zahlreichen empirischen Studien gezeigte leistungssteigernde Rolle zentraler Abschlussprüfungen untermauert, dass die Bereitstellung von Schülerleistungsmaßen zu einer erfolgreicherer Steuerung des Bildungssystems beitragen kann. Sowohl im internationalen Vergleich als auch im Vergleich der deutschen Bundesländer ist wiederholt belegt worden, dass Schülerinnen und Schüler in Schulsystemen mit zentralen Abschlussprüfungen deutlich besser abschneiden als in Schulsystemen ohne zentrale Abschlussprüfungen. Lüdemann (2011) untersucht die Effekte der Einführung zentraler Abschlussprüfungen in vielen Bundesländern seit Beginn der 2000er Jahre. Anhand der PISA-E-Daten von 2000 bis 2006 findet sie, dass die Einführung zentraler Abschlussprüfungen in der Haupt- und Realschule mit einem signifikanten Anstieg der Schülerleistungen einherging.

Anhand von Daten aus Spanien legen Anghel et al. (2015) nahe, dass die Bereitstellung von Schülerleistungsmaßen die Leistungsanreize für Lehrer erhöht: Die Einführung eines Tests mit veröffentlichten Ergebnissen in einer Region hat die Leistungen der Schüler in dieser Region gegenüber denjenigen von Schülern in Regionen ohne diesen Test signifikant erhöht. Anhand vergleichbarer Ergebnisse für

Deutschland argumentieren Jürges et al. (2005), dass die Existenz von zentralen Abschlussprüfungen die Leistungen von Lehrern verbessert.

Anwendbarkeit der Theorie des föderalen Wettbewerbs auf das Bildungswesen

Aus der Theorie der *Yardstick Competition* lässt sich ableiten, dass die Bereitstellung von Informationen über die Ergebnisse bildungspolitischer Maßnahmen die Leistungsfähigkeit der Bildungssysteme der Bundesländer steigern kann. Erstens entstehen neue Anreize für die Entscheidungsträger, ihre Bildungspolitik effektiv und effizient zu gestalten. Zweitens ermöglicht die Bereitstellung von Informationen eine Wirkungsforschung in Bezug auf bildungspolitische Instrumente. Deren Ergebnisse können dazu beitragen, die von den Ländern getragene Bildungspolitik zu verbessern.

Gleichwohl lassen sich auch Argumente gegen das Funktionieren von *Yardstick Competition* und gegen die Bereitstellung von Informationen zur Leistungsfähigkeit von Bildungssystemen vorbringen. Es kann Szenarien geben, in denen *Yardstick Competition* negative Auswirkungen hat. So ist es möglich, dass sich alle Länder bei schlechten Resultaten einpendeln, die dann im Vergleich nicht mehr als schlecht empfunden werden. Da keine Sanktionen durch den Wähler zu befürchten sind, kann es auch zu einer Anpassung nach unten kommen. Solche Szenarien sind allerdings für den Fall der deutschen Bildungspolitik so lange nicht plausibel, wie sich die deutsche Bildungspolitik an internationalen Kompetenzmessungen beteiligt, denn diese würden die Schwächen auf breiter Front aufdecken.

Von Seiten der Länder wird argumentiert, dass gerade im Bildungsbereich fehlerhafte, ideologiebefrachtete Rankings und qualitativ schlechte Bildungsforschung die Bevölkerung irreführen können. Selbstverständlich ist es möglich, dass Wissenschaft fehlerhafte Ergebnisse produziert. Es trifft ebenfalls zu, dass Ranglisten gelegentlich problematisch sind und möglicherweise in der Öffentlichkeit (zu) viel Aufmerksamkeit bekommen. Dass aber wie im Status Quo der Verzicht auf Informationsbereitstellung die beste Antwort auf dieses Problem darstellt, ist nicht überzeugend: Zum einen sollten für politische Entscheidungen nicht einfach Rangordnungen, sondern absolute Kompetenzergebnisse und inhaltliche Argumente ausschlaggebend sein. Solche Informationen können nur auf Basis empirischer Untersuchungen erstellt werden. Zum anderen gilt im Bereich der Bildungspolitik ebenso wie in anderen Lebensbereichen (z.B. der Medizin, der Familienpolitik oder der Arbeitsmarktpolitik), dass ein

transparentes Wissenschaftssystem über die Zeit gute von schlechter Forschung unterscheidet.

Der Gefahr einer zu großen Fixierung auf einzelne, besonders sichtbare Kennziffern kann dadurch begegnet werden, dass nicht nur diese Kennziffern, sondern ein breiter Datensatz öffentlich zugänglich gemacht wird. Dann wird es möglich, verschiedene Kennzahlen zu generieren. Wenn politische Akteure verhindern wollen, dass bestimmte Informationen ein zu hohes Gewicht bekommen, setzt *Yardstick Competition* einen Anreiz, zusätzliche Daten zu erheben, die andere Dimensionen des Bildungserfolgs sichtbar machen. Dann kann sich im Wettbewerb der Meinungen zeigen, auf welche Kennziffern Wähler, Arbeitgeber und die Politik Wert legen.

Ein weiterer Kritikpunkt an Vergleichserhebungen besteht darin, dass sie zu einer testorientierten Lehre beitragen können. Hiermit verknüpft ist das Argument, dass die Pluralität und Differenzierung des Bildungssystems auf der Strecke bleiben könnten und sich der Blick auf das verengt, was messbar ist. Allerdings spricht eine Reihe von Gründen dagegen, dass eine Bereitstellung von Vergleichserhebungen sich insgesamt negativ auswirken würde. Erstens besteht eine wichtige Aufgabe des Schulsystems in der Tat darin, grundlegende und objektiv vergleichbare Fähigkeiten zu vermitteln. Zweitens muss die Einführung von Vergleichserhebungen vor dem Hintergrund der jetzigen Situation bewertet werden: Derzeit werden Prüfungen zumeist dezentral von einzelnen Lehrerinnen und Lehrern durchgeführt, was für viele Schülerinnen und Schüler eine Verzerrung weg von den zentralen Zielen des Bildungssystems bedeuten kann. Es ist geradezu das Ziel von Prüfungssystemen, dass sie den Unterricht auf diejenigen Lerninhalte ausrichten, die von der Gesellschaft insgesamt als erstrebenswert angesehen werden. Drittens kann empirisch in den seltensten Fällen eine negative Korrelation zwischen verschiedenen Kompetenzbereichen festgestellt werden. Tendenziell gehen bessere Ergebnisse in einem kognitiven Bereich auch mit besseren Werten in anderen Bereichen einher. Empirische Studien zeigen, dass bildungspolitische Maßnahmen wie die Nutzung von Vergleichsprüfungen, die mit besseren Leistungen in den schulischen Fächern einhergehen, nicht auf Kosten anderer Bildungsziele gehen (Wößmann et al. 2009). Insofern ist nicht zu erwarten, dass die Bereitstellung von Vergleichsinformationen zu schlechteren Ergebnissen in testfernen Bereichen führt.

Die Gesellschaft unterstützt derzeit eine Vielfalt von Bildungsmöglichkeiten, ohne deren Wert bemessen zu können. Mit der Bereitstellung von Informationen wird es möglich, die Vorteile der verschiedenen Ausbildungsformen – auch wenn sie nur Teilgruppen der Bevölkerung betreffen – nachzuweisen. Somit kann die Ermöglichung von Analysen die Pluralität des Angebotes unterstützen. Die Länder sollten ihre bildungspolitische Expertise nutzen und zur Entwicklung eines breiten Indikatorensatzes beitragen. Die Bereitstellung von Informationen kann als Lernprozess interpretiert werden, in dem schlechte durch bessere Maße ersetzt werden.

Fazit

Vor diesem Hintergrund spricht sich der Wissenschaftliche Beirat beim Bundesministerium für Wirtschaft und Energie (2016) dafür aus, dass Informationen über die Leistungsfähigkeit der Bildungssysteme insbesondere auf Bundeslandebene systematisch bereitgestellt werden. Die Bundesländer sollten sich einer regelmäßigen Überprüfung der Leistungsfähigkeit ihres Schulwesens öffnen. Die bestehenden Systeme zur Messung schulischer Leistungen müssen so ausgestaltet werden, dass aussagekräftige Daten zur Verfügung stehen, die einen Vergleich über die Bundesländer und über die Zeit hinweg zulassen. Die Messverfahren sollten reichhaltig genug sein, um bewerten zu können, wie gut konkurrierende Bildungsziele erreicht werden.

Im Detail empfiehlt der Beirat die Wiederaufnahme des PISA-E Formates, die Bereitstellung von Schülerkerndatensätzen im Rahmen der Datenschutzvorgaben sowie die uneingeschränkte Verwendbarkeit von Bundeslandindikatoren in den verfügbaren Datensätzen inklusive den IQB-Ländervergleichen und dem Nationalen Bildungspanel (NEPS). In der zukünftigen Ausgestaltung zentraler Erhebungen des NEPS sollte die Repräsentativität der Daten auf der Ebene der (zumindest größten) Bundesländer berücksichtigt werden, damit es möglich wird, Politikmaßnahmen der Länder zu evaluieren.

Im Bereich der Bildungspolitik ist Informationsbereitstellung die Voraussetzung dafür, dass die Stärken des Föderalismus zum Tragen kommen. Die Bereitstellung von Informationen kann als dynamischer Lernprozess interpretiert werden, in dem demokratisch legitimierte Bildungsziele durch effektive und effiziente Instrumente

verfolgt werden. Derzeit werden in den Ländern unkoordiniert Reformen durchgeführt, und Modellversuche bleiben empirisch unausgewertet. Die Bildungspolitik könnte von einer Ausrichtung auf den Grundgedanken der Evidenzbasierung profitieren: Im Sinne einer größeren Transparenz sollten Reformen bildungspolitischer Maßnahmen standardmäßig evaluiert werden.

Ob, in welchem Ausmaß und wo die Ziele der Bildungspolitik in Deutschland heute erreicht werden, wissen wir nicht. Es liegt im gemeinsamen Interesse, das deutsche Bildungssystem in seiner Vielfalt bestmöglich zu gestalten. Eine notwendige Voraussetzung dafür sind regelmäßige Berichterstattung, die Bereitstellung und Analyse von Daten und das Zulassen und Fördern von Transparenz.

Literatur

- Anghel, Brindusa, Antonio Cabrales und Ismael Sanz, 2015, Publicizing the Results of Standardized External Tests: Does It Have an Effect on School Outcomes?, *IZA Journal of European Labor Studies* 4(7), 1-20.
- Besley, Timothy und Anne Case, 1995, Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition, *American Economic Review* 85(1), 25-45.
- BMAS (Bundesministerium für Arbeit und Soziales) und IAB (Institut für Arbeitsmarkt- und Berufsforschung), 2011, Sachstandsbericht der Evaluation der Instrumente. Berlin: BMAS.
- Figlio, David, Krzysztof Karbownik und Kjell G. Salvanes, 2016, Education Research and Administrative Data, in: Eric A. Hanushek, Stephen Machin, Ludger Wößmann (Hrsg.), *Handbook of the Economics of Education*, Vol. 5. Amsterdam: North Holland, S. 75-138.
- Jürges, Hendrik, Wolfram F. Richter und Kerstin Schneider, 2005, Teacher Quality and Incentives: Theoretical and Empirical Effects of Standards on Teacher Quality, *FinanzArchiv* 61(3), 298-326.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland), 2011, FAQ's – Frequently Asked Questions zum Kerndatensatz und zur Datengewinnungsstrategie (www.kmk.org/fileadmin/pdf/Statistik/FAQ_KDS.pdf letzter Zugriff: 24.07.2015).
- Lüdemann, Elke, 2011, Intended and Unintended Short-run Effects of the Introduction of Central Exit Exams, in: Elke Lüdemann, *Schooling and the Formation of Cognitive and Non-cognitive Outcomes*, Chapter 4. ifo Beiträge zur Wirtschaftsforschung 39. München: ifo Institut.
- Machin, Stephen, 2014, Economics of Education Research and Its Role in the Making of Education Policy, *Fiscal Studies* 35(1), 1-18.

- Wissenschaftlicher Beirat beim Bundesministerium für Wirtschaft und Energie, 2016, Mehr Transparenz in der Bildungspolitik. Berlin: BMWi.
- Wissenschaftsrat, 2013, Stellungnahme zum Nationalen Bildungspanel (NEPS), Drs. 2999-13. Würzburg 26.04.2013 (www.wissenschaftsrat.de/download/archiv/2999-13.pdf letzter Zugriff: 24.07.2015).
- Wößmann, Ludger, 2007, Letzte Chance für gute Schulen: Die 12 großen Irrtümer und was wir wirklich ändern müssen. München: ZS Verlag Zabert Sandmann.
- Wößmann, Ludger, 2015, Warum PISA-Leistungen ein wichtiger Indikator der Schülerkompetenzen im internationalen Vergleich sind, in: Friedhelm Pfeiffer (Hrsg.), Bildungsindikatoren im internationalen Vergleich: Bestandsaufnahme und kritische Würdigung. Mannheim: Zentrum für Europäische Wirtschaftsforschung, S. 48-55.
- Wößmann, Ludger, Elke Lüdemann, Gabriela Schütz und Martin R. West, 2009, School Accountability, Autonomy and Choice around the World. Cheltenham: Edward Elgar.

Regina T. Riphahn ist Professorin für Statistik und empirische Wirtschaftsforschung an der Friedrich-Alexander-Universität Erlangen-Nürnberg und Vorsitzende des Rates für Sozial- und Wirtschaftsdaten. Ludger Wößmann ist Leiter des ifo Zentrums für Bildungsökonomik am Münchner ifo Institut und Professor für Volkswirtschaftslehre an der Ludwig-Maximilians-Universität München. Beide sind Mitglied im Wissenschaftlichen Beirat beim Bundesministerium für Wirtschaft und Energie.