

Friederike Hertweck, Ingo E. Isphording, Sönke H. Matthewes, Kerstin Schneider, C. Katharina Spieß*

Bildungsdaten: Datenlücken durch ein Bildungsverlaufsregister schließen

Bildung ist von zentraler Bedeutung für die gesellschaftliche und wirtschaftliche Entwicklung und damit den Wohlstand Deutschlands. Bildung ist auch Voraussetzung für die Teilhabe und Entwicklung jedes einzelnen Menschen. In Deutschland ist spätestens seit dem PISA-Schock im Jahr 2001 bekannt und unstrittig, dass das deutsche Bildungssystem weiterentwickelt werden muss – auch heute nach der Coronapandemie wird vielfach über Reformnotwendigkeiten diskutiert. Für eine solche Weiterentwicklung sind Daten die unabdingbare Grundlage für gutes evidenzbasiertes politisches und administratives Handeln. So hat der Auftrag zum Bildungsmonitoring mit Art. 91b GG Satz 2¹ Eingang in das Grundgesetz gefunden. Allerdings sind gute Daten allein nicht ausreichend, sie müssen einem Erkenntnisgewinn dienen, der Konsequenzen für die Bildungspolitik nach sich zieht. Nur eine umfassende Datenbasis mit Zugang für die Wissenschaft ermöglicht es, die komplexen Zusammenhänge der Bildungssysteme zu verstehen und eine Grundlage für eine evidenzbasierte Bildungspolitik zu liefern.

Doch obwohl die große Bedeutung von Daten für gute Bildungspolitik seit vielen Jahren bekannt ist, steht der Bildungsforschung in Deutschland nur eine unzureichende Datenbasis zur Verfügung. So stellt z.B. die amtliche Statistik der meisten Bundesländer die Daten zur Soziodemografie der Schülerschaft oder Erfolgsindikatoren, wie Übergangsraten und Abschlüsse, nicht in Form von Individualdaten, sondern nur aggregiert bereit. Individuelle amtliche Bildungsverlaufsdaten gibt es bisher nur regional vereinzelt. Auch ist der Zugang der Wissenschaft zu vorhandenen amtlichen Mikrodaten vielfach mit hohen Hürden verbunden. Noch größer sind die Hürden, wenn mit Daten zu den Kompetenzen von Schüler:innen geforscht werden soll, auch wenn hier mit dem Zugang der Forschung zu den IQB-Bildungstrenddaten in der Vergangenheit Fortschritte erzielt wurden. Die VERA-Daten bzw. vergleichende Kom-

petenzerhebungen sind der Forschung aber nach wie vor nur in einigen Bundesländern zugänglich.

Diese Defizite sind für den Forschungsstandort Deutschland ein erheblicher Wettbewerbsnachteil und erschweren zielgenaue, effektive und effiziente Bildungspolitik. Für eine fundierte Bewertung von Wirkungen von Bildungspolitik sollten Bildungsergebnisse gemessen und zudem mit Ausgaben für Bildung in Verbindung gesetzt werden können. Nur so kann nach wissenschaftlichen Standards überprüft werden, ob mehr Geld für Bildung auch zu höherem Bildungserfolg führt und Mittel für Bildung wirksam eingesetzt werden (vgl. auch SWK, 2022).

Mittelvergabe und Evaluation von Maßnahmen erfordern aussagekräftige Datenbasis

Ein aktuelles Beispiel für ein solches Vorhaben ist das angekündigte „Startchancen-Programm“ der Bundesregierung, das mit einem Volumen von jeweils 2 Mrd. Euro über zehn Jahre insbesondere benachteiligte Schüler:innen unterstützen soll. Eine wissenschaftliche Begleitung und Eva-

Friederike Hertweck, Ph.D., ist Postdoc am RWI – Leibniz-Institut für Wirtschaftsforschung in Essen.

Dr. Ingo E. Isphording ist Forschungsdirektor des Forschungsinstituts zur Zukunft der Arbeit (IZA), Bonn.

Dr. Sönke H. Matthewes ist Postdoc an der Professur für VWL, insbesondere Wirtschaftspolitik der Universität Potsdam.

Prof. Dr. Kerstin Schneider ist Professorin an der Bergischen Universität Wuppertal und Vorsitzende des Wuppertaler Instituts für bildungsökonomische Forschung (WIB).

Univ.-Prof. Dr. C. Katharina Spieß ist Direktorin des Bundesinstituts für Bevölkerungsforschung (BiB) und Professorin an der Johannes Gutenberg-Universität Mainz.

¹ Art. 91b GG Satz 2: „(2) Bund und Länder können auf Grund von Vereinbarungen zur Feststellung der Leistungsfähigkeit des Bildungswesens im internationalen Vergleich und bei diesbezüglichen Berichten und Empfehlungen zusammenwirken.“

© Der/die Autor:in 2023. Open Access: Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht (creativecommons.org/licenses/by/4.0/deed.de).

Open Access wird durch die ZBW – Leibniz-Informationszentrum Wirtschaft gefördert.

* Dies ist eine Kurzversion der Stellungnahme der Unterarbeitsgruppe Bildung. Die Langfassung findet sich hier: www.socialpolitik.de/de/vfs_stellungnahmen_datens_bildung.

uation sind als Teil des Programms vorgesehen. Allerdings fehlt sowohl für die bedarfsgerechte Mittelvergabe als auch für eine Evaluation der Wirkung und der Effizienz von Bildungsprogrammen in Deutschland vielfach die notwendige Datenbasis. So benötigt die Bildungspolitik für eine zielgerichtete Mittelvergabe Informationen über Bedarfe von Bildungseinrichtungen. Dies gilt auch für die Unterstützung benachteiligter Schüler:innen, die das Startchancen-Programm erreichen möchte. Die Bedarfe können über verschiedene Merkmale definiert werden, darunter der soziale Hintergrund der Familien oder auch nachweisbare Defizite im Kompetenzerwerb der Bildungsteilnehmer:innen. Diese relevanten Kennzahlen liegen in der Regel jedoch nicht auf ausreichend kleinräumiger Ebene, wie z.B. auf der Ebene der Schule, vor – und sind erst recht nicht auf Ebene der Bildungsteilnehmer:innen vorhanden.

Unklar, ob und wie eingesetzte Mittel wirken

Soll Bildungspolitik nicht nur bedarfsgerecht, sondern auch wirkungsorientiert sein, muss der Mitteleinsatz zudem in Hinblick auf die erlangten Kompetenzen und Bildungsabschlüsse der Bildungsteilnehmer:innen evaluiert werden. Für eine solche Kausalanalyse von Bildungswirkungen fehlt aber die flächendeckende Datenbasis insbesondere zu Kompetenzen und individuellen Bildungskarrieren, Bildungsübergängen und -abschlüssen. Eine über das Bildungssystem hinausgehende Analyse, etwa über langfristige Wirkungen auf berufliche Laufbahnen oder die Gesundheit, ist aufgrund fehlender verknüpfter Datenquellen erst recht nicht möglich.

Auch wegen fehlender Daten beschränkt sich Begleitforschung der Bildungssteuerung weitestgehend auf eine inputbasierte Steuerung über Verwendungskonzepte der Bildungseinrichtungen. Ob nach dem Mittelfluss die damit finanzierten Konzepte ihre angedachte Wirkung kosteneffizient entfalten, wird in der Regel nicht überprüft. Damit ist weder die Effektivität noch die Effizienz der Mittelvergabe sichergestellt. So ist die Wirkung von Bildungsprogrammen auf Bildungsverläufe in Deutschland weitgehend unbekannt und kann mit den verfügbaren Bildungsverlaufsdaten auch nicht erfasst werden.

Stichprobenbasierte Umfragedaten sind wichtig, aber nicht ausreichend für differenziertere Analysen

Bildungsverläufe sind in Deutschland bisher nur auf der Basis von Umfragedaten, wie dem Nationalen Bildungspanel (NEPS), möglich, welche teilweise auch mit den Daten der Sozialversicherung des IAB verknüpft werden können. Diese Verknüpfung von Bildungs- mit registerbasierten Arbeitsmarktdaten ermöglicht die Messung langfristiger Bildungseffekte über die berufliche Bildung hinaus und

erschließt neue Forschungsmöglichkeiten. Umfragedaten alleine sind jedoch nicht ausreichend für eine effektive und evidenzbasierte Bildungspolitik. Hier ist ein breit definiertes Bildungsverlaufsregister über alle Bildungsabschnitte erforderlich und die geeignete Grundlage.

Bildungsverlaufsregister kann Politik fundierte Entscheidungsgrundlagen liefern

Bildungsverlaufsregister sind in vielen europäischen Ländern bereits Realität. So stellen die skandinavischen Länder, die Niederlande, Österreich, die Schweiz und das Vereinigte Königreich umfassende und flächendeckende Bildungsverlaufsdaten der Bildungsforschung und -steuerung zur Verfügung. Aus dem Vergleich der Infrastruktur europäischer Nachbarländer lassen sich Vorschläge für die institutionelle Ausgestaltung der entsprechenden Infrastruktur in Deutschland ableiten, etwa über die Verantwortung für Datenhaltung und -bereitstellung, den Umfang der Daten, deren Verknüpfbarkeit mit weiteren Datenquellen und über die Bedeutung von Kompetenzmessungen.

Auch in Deutschland wird über ein Bildungsverlaufsregister seit geraumer Zeit diskutiert und erste Schritte sind unternommen worden (RatSWD, 2022; Giar et al., 2023). Ein Bildungsverlaufsregister für Deutschland ist von der Bildungspolitik selbst schon als wichtig erkannt worden. Bereits 2003 wurde von der Kultusministerkonferenz (KMK, 2020) der Kerndatensatz für den Schulbereich beschlossen, der die schulische Laufbahn von Schüler:innen abbilden soll. Dieses Vorhaben ist allerdings noch nicht umgesetzt. Ein Bildungsverlaufsregister bedeutet ausdrücklich nicht, dass Schüler:innen „gläsern“ werden. Stattdessen sollen der Politik für ihre Entscheidungen empirische Befunde basierend auf exzellenter Forschung geliefert werden. Und diese Forschung setzt gute Daten voraus. Datenschutz muss und kann dabei auf höchstem Niveau gewährleistet werden.

In einem Bildungsverlaufsregister sollen Merkmale einer Bildungskarriere in einzelnen nicht miteinander verbundenen Bildungsstatistiken von einem Datentreuhänder pseudonymisiert für Zwecke der Forschung und Steuerung verfügbar gemacht werden. Dabei ist wichtig, dass die Daten nur in dem Maße, wie sie zur Beantwortung einer Fragestellung notwendig sind, vom Datentreuhänder über eine unveränderliche Identifikationsnummer verknüpft und in anonymisierter Form und über sichere Zugangswege an Forschende weitergegeben werden. Die Ergebnisse der Forschung werden in aggregierter Form publiziert. Ein Rückschluss auf einzelne Personen soll und darf nicht möglich sein.

Das Bildungsverlaufsregister sollte sämtliche Bildungsabschnitte umfassen, von der frühkindlichen Bildung in Familie und Kindertagesbetreuung über Grundschule

und Sekundarschule, Berufsbildung und Hochschullaufbahn. Auch lebenslanges Lernen über betriebliche oder außerbetriebliche Weiterbildung sollte abgebildet werden können. Grundlegend sollten soziodemografische Merkmale der Kinder/Schüler:innen sowie bei minderjährigen Bildungsteilnehmer:innen der Eltern erfasst werden. Dazu gehören Geschlecht, Alter, Zuwanderungsgeschichte sowie Sozialleistungsbezug und Bildungsniveau der Eltern. Der individuelle Bildungserfolg muss über Übergänge (auf Schulen, Berufsschulen und Hochschulen) erlangte Abschlüsse, Studienerfolg bzw. -abbruch und die Erreichung von Bildungszielen erfasst sein sowie über weitere Leistungsmaße, wie standardisierte Testergebnisse oder Noten. Insbesondere die Erfassung von Kompetenzmaßen hat eine wichtige Bedeutung.

Bildungsdaten müssen verknüpfbar sein

Für viele Fragestellungen von hoher gesellschaftlicher Relevanz ist zudem eine Verknüpfbarkeit der Daten über das Bildungsverlaufsregister hinaus erforderlich. Denn Bildungserfolg misst sich nicht nur in Kompetenzmaßen und Bildungsabschlüssen. Eine gute Bildungspolitik kann Arbeitsmarkt-, Wirtschafts-, Sozial- und Familienpolitik zugleich sein und bessere Bildung und geringere Bildungsungleichheit können starke Auswirkungen auf Armut, Gesundheit, Kriminalität und vieles mehr haben. Die Bestimmung der monetären und nicht-monetären Bildungsrenditen verlangt nach einer Verknüpfung mit Datenregistern anderer Gesellschaftsbereiche, etwa mit Arbeitsmarkt- und Gesundheitsdaten. Ebenso sinnvoll ist die Möglichkeit einer Verknüpfung mit weiteren umfragebasierten Datensätzen, um Lücken der Registerdaten gezielt für spezielle Forschungsfragen schließen zu können.

In einzelnen Bildungsabschnitten sind Datenquellen vorhanden – große Lücken bei Kompetenzmessungen

Generell werden in Deutschland bildungsbezogene Daten oft nur in Stichproben und als Querschnitte erhoben. Sie liegen häufig in nicht verknüpfbaren Datensilos und um sie wissenschaftlich nutzen zu können, sind hohe bürokratische und administrative Hürden zu überwinden. In den einzelnen Bildungsabschnitten sind jedoch bereits Datenquellen vorhanden, aus denen ein zukünftiges Bildungsverlaufsregister gespeist werden könnte. So sind Daten zur frühen Bildung in der öffentlich finanzierten Kindertagesbetreuung (Spieß, 2009), Sprachstanderhebungen und Schuleingangsuntersuchungen im Übergang zwischen vorschulischer und schulischer Bildung bereits vorhanden, doch sowohl Datenerhebungsmethoden als auch Datenzugang unterscheiden sich teilweise zwischen den Bundesländern. Eine Schülerindividualstatistik für allgemeinbildende sowie berufsbildende Schulen gibt es bis-

lang nicht in allen Bundesländern. Und auch wenn – insbesondere in einem mehrgliedrigen Schulsystem – Übergänge zwischen Bildungsabschnitten besonders wichtig sind, werden Übergangsraten zwischen Bildungsabschnitten und Schulformen nur in aggregierter Form von den Statistischen Landesämtern zur Verfügung gestellt. Es ist mit diesen Daten nicht möglich, Bildungsverläufe zu analysieren. Die Möglichkeit, individuelle Bildungskarrieren in der schulischen Bildung zu analysieren, kann erst durch den von der KMK bereits 2003 beschlossenen, aber bisher nicht in allen Bundesländern umgesetzten, Kerndatensatz geschaffen werden. Er kann mit über die Länder hinweg harmonisierten individuellen Verlaufsdaten als Basis für ein zukünftiges Bildungsverlaufsregister dienen.

Eine große Lücke der bestehenden amtlichen Statistik sind die fehlenden Kompetenzmessungen. Bildungserfolg lässt sich bisher ausschließlich aus Bildungsübergängen und erreichten Abschlüssen erfassen. Für eine effektive Steuerung wären allerdings standardisierte, regelmäßige und flächendeckende Kompetenzmessungen unerlässlich. Dazu könnten die bereits bestehenden aber für die Forschung bisher kaum nutzbaren Vergleichsarbeiten (VERA) weiterentwickelt werden. Durch eine Verknüpfung von flächendeckenden Kompetenzmessungen mit den Individualdaten eines Kerndatensatzes ließe sich eine der wichtigsten Datenlücke der amtlichen Schulstatistik ohne großen Mehraufwand schließen.

Gesamterhebungen im Hochschulbereich liegen über die Statistik der Studierenden und die Statistik der Prüfungen vor, allerdings nur im Querschnitt und mit wenigen erfassten Merkmalen. Mit der Novelle des Hochschulstatistikgesetzes von 2016 wurde erstmals eine Studienverlaufsstatistik eingeführt. Sie steht der Wissenschaft aktuell noch nicht für Auswertungen zur Verfügung, sollte aber in das Bildungsverlaufsregister integriert werden. Anpassungsbedarf gibt es zudem bei den Löschungsvorschriften, denn die Pseudonyme und die Zusammenführungen müssen nach 18 Jahren gelöscht werden, nach einer Promotion bereits nach vier Jahren. Aus Sicht der Wissenschaft wird dadurch enormes Analysepotenzial verschenkt und eine Anpassung der gesetzlichen Regelungen im Hochschulstatistikgesetz wäre wünschenswert. Bisher nicht abgebildet in der amtlichen Statistik ist der Übergang zwischen Schule und Hochschule. Dieser könnte im Bildungsverlaufsregister durch Daten der Plattform Hochschulstart, über die viele Hochschulen ihr Verfahren zur Studienplatzvergabe abwickeln, ergänzt werden.²

Auch die Berufsbildung im dualen System sollte Teil eines Bildungsverlaufsregisters sein. Hier bietet die amtliche Sta-

2 Plattform Hochschulstart, <https://hochschulstart.de/startseite/>.

tistik mit der Statistik der beruflichen Schulen zur Schulstatistik vergleichbare Querschnittsdaten. Der betriebliche Teil der Ausbildung wird über die Berufsbildungsstatistik abgedeckt, die die Grundgesamtheit aller in Deutschland erfassten Ausbildungsverträge umfasst. Weitere stichproben- und umfragebasierte Datensätze zur betrieblichen Ausbildung sollten perspektivisch mit den Registern verknüpfbar sein. Für Bildungsinvestitionen außerhalb der sekundären und tertiären Bildungssysteme gibt es nach derzeitigem Kenntnisstand keine administrative Datenquelle, die in ein Bildungsverlaufsregister eingebunden werden könnte.

Empfehlungen aus Sicht der Wissenschaft

Die Umsetzung des Bildungsverlaufsregisters in Deutschland konkretisiert sich und der politische Wille ist (mehrheitlich) da. Die folgenden Empfehlungen sind relevant:

Umfang und Abdeckung der Bildungsverläufe: Das Bildungsverlaufsregister darf nicht zu eng definiert sein, um Analysepotenziale nicht unnötig zu vergeben. Um datenschutzrechtlichen Einwänden Rechnung zu tragen, sollten Individualdaten zu einzelnen Bildungsabschnitten in einzelnen abschnittsbezogenen Registern gespeichert werden. Außerdem sollten alle Bildungsbereiche von der frühkindlichen Bildung bis mindestens zum Berufsbildungs- und/oder Hochschulabschluss Eingang in ein Bildungsverlaufsregister finden.

Transparenter Datenzugang für die Wissenschaft: Der Zugang zu Bildungsdaten sollte für die Wissenschaft klar und transparent geregelt werden. Anfallende Nutzungsgebühren dürfen nicht dazu führen, dass Forschende, insbesondere Nachwuchswissenschaftler:innen, ihre Forschung nicht durchführen können. Der Datenzugang könnte z.B. über das geplante Forschungsdatengesetz geregelt werden (RatSWD, 2023).

Verknüpfbarkeit: Die reiche Datenlandschaft in Deutschland leidet darunter, dass einzelne Datensätze unverbundene Datensilos bleiben. Das trifft auch auf Bildungsdaten zu. Ein Bildungsverlaufsregister würde bislang unverbundene Bildungsdaten verschiedener Quellen verknüpfen. Dies erfordert die Einführung von Personen-IDs und einen Datentreuhänder, der Daten verknüpfen und für Forschungszwecke

verfügbar machen kann. Das Bildungsverlaufsregister darf nicht selbst zum Datensilo werden. Die Analyse langfristiger Effekte von Bildung kann erst durch die Verknüpfung des Bildungsverlaufsregisters, z. B. mit Erwerbsverläufen, Steuerdaten, der Kriminalitätsstatistik oder Gesundheitsdaten, erfolgen. Auch die Verknüpfung mit etablierten Survey-Studien in Deutschland sollte ermöglicht werden.

Datenschutz: Den Anforderungen der DSGVO und der informellen Selbstbestimmung nach GG Art. 2 Abs. 1 muss und kann ohne Einschränkungen entsprochen werden. Bei der Ausgestaltung des Zugangs zum Bildungsverlaufsregister für die unabhängige Wissenschaft muss das Forschungsprivileg für den Zugang der Wissenschaft zu sensiblen Daten Berücksichtigung finden. Zudem muss es eine einheitliche Auslegung der datenschutzrechtlichen Bestimmungen geben, denn es sind sowohl Landes- als auch Bundesstatistiken betroffen. Außerdem darf es keine Löschungspflicht für Daten geben, wenn nicht sichergestellt ist, dass pseudonymisierte Daten für Bildungsverläufe verfügbar sind.

Literatur

- Giar, K., F. Hohlstein, M. Wipke und A. Scharnagl (2023), Konzeption eines Statistischen Bildungsverlaufsregisters in Deutschland – Entwicklungen bis 2023 und Ausgestaltungsoptionen, Statistisches Bundesamt, WISTA 3, 51–62.
- KMK – Kultusministerkonferenz (2020), Ländervereinbarung über die gemeinsame Grundstruktur des Schulwesens und die gesamtstaatliche Verantwortung der Länder in zentralen bildungspolitischen Fragen, https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2020/2020_10_15-Laendervereinbarung.pdf (1. November 2023).
- RatSWD – Rat für Sozial- und Wirtschaftsdaten (2022), Aufbau eines Bildungsverlaufsregisters: Datenschutzkonform und benutzerfreundlich: Empfehlungen des Rates für Sozial- und Wirtschaftsdaten für den Aufbau eines Bildungsverlaufsregisters, <https://www.konsortswd.de/wp-content/uploads/Positionspapier-RatSWD-Aufbau-eines-Bildungsverlaufsregisters.pdf> (1. November 2023).
- RatSWD – Rat für Sozial- und Wirtschaftsdaten (2023), Forschungsdatengesetz: Was zentral ist, https://www.konsortswd.de/wp-content/uploads/RatSWD_Positionspapier_Forschungsdatengesetz_Was_zentral_ist.pdf (1. November 2023).
- Spieß, C. K. (2009), Improvements and Future Challenges for the Research Infrastructure in the Field of “Preschool Education”, *Working Paper of the Council for Social and Economic Data*, 60, https://www.ratswd.de/download/workingpapers2009/60_09.pdf (1. November 2023).
- SWK – Ständige Wissenschaftliche Kommission (2022), Entwicklung von Leitlinien für das Monitoring und die Evaluation von Förderprogrammen im Bildungsbereich, https://www.kmk.org/fileadmin/Dateien/pdf/KMK/SWK/2022/SWK-2022-Impulspapier_Monitoring.pdf (1. November 2023).

Title: Education Data: Closing Data Gaps Through a Longitudinal Education Register

Abstract: Evidence-based education policy needs comprehensive data. Education economists of the Verein für Socialpolitik call for the implementation of a longitudinal education register in Germany to understand the complex interrelationships of education systems. The education register should cover all levels of formal education and include data on student performance in order to close a serious data gap in Germany. The statement also calls for regular access to this data for independent research. Once implemented, the education register should enable linking with other data sources, for example from the labour market, tax and health records, in order to be able to analyse the long-term effects of education. By establishing a longitudinal education register, Germany would finally catch up with established standards of other European countries, creating a resource where education register data not only exists, but are available for researchers and can be linked with other data.