

Kühne, Stefan; Mank, Svenja; Schulz, Stefan; Maaz, Kai  
**Sekundäranalysen mit Daten der amtlichen Statistik. Potenziale und Herausforderungen für das Bildungsmonitoring in Deutschland**

*Zeitschrift für Pädagogik 67 (2021) 6, S. 860-880*



Quellenangabe/ Reference:

Kühne, Stefan; Mank, Svenja; Schulz, Stefan; Maaz, Kai: Sekundäranalysen mit Daten der amtlichen Statistik. Potenziale und Herausforderungen für das Bildungsmonitoring in Deutschland - In: Zeitschrift für Pädagogik 67 (2021) 6, S. 860-880 - URN: urn:nbn:de:0111-pedocs-287889 - DOI: 10.25656/01:28788; 10.3262/ZP2106860

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-287889>

<https://doi.org/10.25656/01:28788>

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

# ZEITSCHRIFT FÜR PÄDAGOGIK

Heft 6

November/Dezember 2021

■ *Thementeil*

**Forschungsdaten in den Bildungswissenschaften – Archivierung und Nachnutzung**

■ *Allgemeiner Teil*

Soziale Topographien von Bildung. Sozialräumliche Kontexte der Bildungswege von jungen Erwachsenen mit Kinder- und Jugendhilfeeferfahrung

Vorbereiten auf den Ernst des Lebens.  
Eine Ethnografie zur pädagogischen Ordnung von Unterricht in Institutionen der Berufsvorbereitung

## Inhaltsverzeichnis

### *Thementeil: Forschungsdaten in den Bildungswissenschaften – Archivierung und Nachnutzung*

*Sabine Reh/Petra Stanat*

Forschungsdaten in den Bildungswissenschaften –  
Archivierung und Nachnutzung. Eine Einleitung ..... 827

*Malte Jansen/Claudia Neuendorf/Aleksander Kocaj*

Welche Potenziale bieten Sekundäranalysen für die Erhöhung  
von Forschungsqualität und Replizierbarkeit – Zur Rolle  
von Multiversumsanalysen und integrativen Datenanalysen  
für die Bestimmung der Robustheit und Generalisierbarkeit  
von Forschungsbefunden ..... 840

*Stefan Kühne/Svenja Mank/Stefan Schulz/Kai Maaz*

Sekundäranalysen mit Daten der amtlichen Statistik – Potenziale und  
Herausforderungen für das Bildungsmonitoring in Deutschland ..... 860

*Kathrin Zöller*

Zur zeithistorischen Nachnutzung sozialwissenschaftlicher Datenbestände  
am Beispiel der Sächsischen Längsschnittstudie zu Jugendlichen  
in Ostdeutschland ..... 881

*Tristan Bauder/Sinje Brinkmann/Christoph Leser/Mirja Silkenbeumer/  
Saskia Terstegen*

Qualität qualitativer Forschungsdaten im Kontext von Archivierung  
und nachhaltiger Nutzung – Sondierungen und Perspektiven für die qualitative  
Bildungsforschung ..... 899

*Deutscher Bildungsserver*

Linktipps zum Thema „Forschungsdaten in der Bildungsforschung“ ..... 915

## Allgemeiner Teil

*Stephan Sting*

Soziale Topographien von Bildung. Sozialräumliche Kontexte der Bildungswege von jungen Erwachsenen mit Kinder- und Jugendhilfeerfahrung ..... 923

*Marc Thielen/Antje Handelsmann*

Vorbereiten auf den Ernst des Lebens.  
Eine Ethnografie zur pädagogischen Ordnung von Unterricht in Institutionen der Berufsvorbereitung ..... 943

## Besprechungen

*Roland Merten*

Vorbemerkung zu den Rezensionen von Peter H. Ludwig:  
Grundbegriffe der Pädagogik. Definitionskriterien, kritische Analyse, Vorschlag eines Begriffssystems ..... 961

*Johannes Drerup*

Peter H. Ludwig: Grundbegriffe der Pädagogik. Definitionskriterien, kritische Analyse, Vorschlag eines Begriffssystems ..... 961

*Manfred Lüders*

Peter H. Ludwig: Grundbegriffe der Pädagogik. Definitionskriterien, kritische Analyse, Vorschlag eines Begriffssystems ..... 964

*Nicole Welter*

Julia Reuter/Markus Gamper/Christina Möller/Frerik Blome (Hrsg.):  
Vom Arbeiterkind zur Professur. Sozialer Aufstieg in der Wissenschaft ..... 968

*Klaus Zierer*

Erhard Wiersing: Hartmut von Hentig – Ein Essay zu Leben und Werk.  
Mit einem Geleitwort von Jörg W. Ziegenspeck ..... 972

## Dokumentation

Pädagogische Neuerscheinungen ..... 977

Impressum ..... U3

## Table of Contents

### *Topic: Research Data in Education – Infrastructures and Reuse*

*Sabine Reh/Petra Stanat*

Research Data in Education – Infrastructures and Reuse. An Introduction ..... 827

*Malte Jansen/Claudia Neuendorf/Aleksander Kocaj*

The Potential of Secondary Data Analysis for Increasing the Quality and Replicability of Educational Research Results: Using Multiverse Analysis and Integrative Data Analysis to Examine Robustness and Generalizability ..... 840

*Stefan Kühne/Svenja Mank/Stefan Schulz/Kai Maaz*

Secondary Analyses with Official Data. The Potential and Challenges for Monitoring Education in Germany ..... 860

*Kathrin Zöller*

Using Social Science Data on Eastern Germany in Upheaval. Historical Approaches and Suggestions ..... 881

*Tristan Bauder/Sinje Brinkmann/Christoph Leser/Mirja Silkenbeumer/Saskia Terstegen*

Quality of Qualitative Research Data in the Context of Archiving and Sustainable Use: Explorations and Perspectives for Qualitative Educational Research ..... 899

*Deutscher Bildungsserver*

Online Resources „Research Data in Education“ ..... 915

### *Articles*

*Stephan Sting*

Social Topographies of Education. Socio-spatial Contexts of the Educational Pathways Taken by Emerging Young Adults with Experience in Out-of-Home Care ..... 923

*Marc Thielen/Antje Handelsmann*

Preparing for the Seriousness of Life. An Ethnographical Study on the Pedagogical Order of Teaching in Institutions that Prepare for Vocational Training ..... 943

Book Reviews .....	961
New Books .....	977
Impressum .....	U3

*Stefan Kühne/Svenja Mank/Stefan Schulz/Kai Maaz*

## **Sekundäranalysen mit Daten der amtlichen Statistik**

*Potenziale und Herausforderungen für das Bildungsmonitoring  
in Deutschland*

**Zusammenfassung:** Im Vergleich zu Primärdaten aus wissenschaftlichen Erhebungen sind amtliche Statistiken mit Blick auf die erfassten Inhalts- und Merkmalsbereiche in der Regel deutlich stärker eingeschränkt. Ihre Vorteile liegen aber darin, dass sie ein flächendeckendes und in langen Zeitreihen vorhandenes Datenmaterial für sekundäranalytische Forschungszwecke bereitstellen. Sie können nicht nur als Bezugspunkt für wissenschaftliche Untersuchungen dienen (z. B. Stichprobenziehung) sondern auch für eigenständige Fragestellungen und Analysen genutzt werden, für die keine alternativen Forschungsdaten verfügbar sind (z. B. für kleinere Populationen oder kleinräumige Disparitäten). Der vorliegende Beitrag stellt in groben Zügen das verfügbare Spektrum und die Zugangsmöglichkeiten zu amtlichen Bildungsstatistiken vor. Mehr und mehr Statistiken werden inzwischen auf Individualdatenbasis erhoben, können also nicht nur für Struktur- oder Zeitreihenanalysen, sondern auch für Mikrodatenanalysen verwendet werden. Anhand ausgewählter Forschungsergebnisse wird der Mehrwert der unterschiedlichen amtlichen Datentypen und -zugänge aus Sicht des Bildungsmonitorings in Deutschland veranschaulicht.

**Schlagworte:** Bildungsstatistik, Schulstatistik, Amtliche Daten, Forschungsdaten, Bildungsmonitoring

### **1. Einleitung**

Die amtliche Statistik zählt zu den größten ‚Datenproduzenten‘ in Deutschland. Jahr für Jahr wird auf Basis verschiedener gesetzlicher Grundlagen eine Fülle standardisierter Daten über alle gesellschaftlichen Bereiche hinweg zusammen getragen und veröffentlicht – überwiegend, aber nicht nur, durch die Statistischen Ämter des Bundes und der Länder. Gleichwohl spielt die amtliche Statistik im Vergleich zu wissenschaftsgetragenen Primärerhebungen nur eine marginale Rolle in der empirischen Bildungsforschung. Dies dürfte unter anderem mit ihrer eingeschränkten Spezifität zusammenhängen, weil Bildungsstatistiken „nicht gezielt ad hoc auf spezielle Fragestellungen ausgerichtet sein können, wie dies zum Beispiel bei kurzfristig zusammengestellten Eckdaten oder eigens erhobenen Stichprobendaten für besondere Zwecke möglich ist“ (Köhler, 1980, S. 1262). Da sie überwiegend im Verwaltungsvollzug generiert, standardisiert erfasst und in der Regel erst 1 bis 2 Jahre nach Erhebung in aggregierter Aufbereitung bereitgestellt werden, erscheinen amtliche Daten auf den ersten Blick als Forschungs-

daten weniger interessant oder relevant. Welchen Wert Bildungsstatistiken auch für Forschungszwecke besitzen, wurde jedoch schon in den 1950er Jahren in einer Arbeit zum relativen Schulbesuch sichtbar (Carnap & Edding, 1962). Spätestens Georg Pichts einflussreiche Schrift über die sprichwörtlich gewordene „deutsche Bildungskatastrophe“ zeigte dann Mitte der 1960er Jahre, welchen eigenständigen Informationswert solche Daten zu Schüler:innen, Lehrpersonal, Hochschulabsolvent:innen oder Bildungsausgaben entfalten können (Picht, 1965).

Eine intensivierete Auseinandersetzung mit Bildungsstatistiken setzte gegen Ende des 20. Jahrhunderts ein, als moderne Formen des Bildungsmonitorings eingeführt wurden, um dem steigenden Informationsbedarf über die Leistungsfähigkeit des deutschen Bildungssystems in Bildungspolitik und -administration Rechnung zu tragen. Von der internationalen bis zur kommunalen Ebene haben sich inzwischen unterschiedliche Formate etabliert, die das Steuerungs- und Entscheidungshandeln der Akteur:innen stärker als zuvor auf eine empirisch gestützte Basis stellen sollen. Seither werden immer mehr amtliche Datensegmente mit zunehmend elaborierten wissenschaftlichen Analysemethoden für Monitoringinstrumente wie den nationalen Bildungsbericht erschlossen (vgl. zuletzt Autorengruppe Bildungsberichterstattung, 2020). Dass in den Bildungswissenschaften nur zögerlich zur Kenntnis genommen wird, welche Daten von Amtswegen erhoben werden und welche Analysen damit möglich sind, dürfte insofern weniger dem bürokratischen Ruf der amtlichen Statistik geschuldet sein, sondern eher an der fehlenden Transparenz über ihre Verfügbarkeit und Zugänglichkeit liegen.

Hier setzt der vorliegende Beitrag an, indem er einen grundlegenden Überblick über verfügbare Daten und Auswertungsmöglichkeiten der amtlichen Statistik bietet. Dafür wird zunächst über das Prinzip amtlicher Datenerhebungen zum Bildungswesen Auskunft gegeben (Datengenerierung). Anschließend werden die verschiedenen Möglichkeiten des Datenzugangs für wissenschaftliche Sekundäranalysen vorgestellt (Datenbereitstellung) und zugleich das jeweilige Informationspotenzial anhand eines ausgewählten Analysebeispiels aus dem Kontext des indikatorengestützten Bildungsmonitorings veranschaulicht (Datenauswertung). Der Fokus richtet sich dabei in erster Linie auf amtliche Daten zum Schulwesen, da die Forschung hier aufgrund der föderal geregelten Datenerfassung vor besonderen Herausforderungen steht.

## **2. Datengenerierung**

### *2.1 Grundzüge der amtlichen Bildungsstatistik*

Ohne den Anspruch auf eine allgemeingültige Definition zu erheben, lassen sich amtliche Statistiken als Datengrundlagen verstehen, die auf Bundes-, Landes- oder kommunaler Ebene durch offizielle Stellen nach gesetzlichen Regeln für das öffentliche Verwaltungshandeln erhoben werden und den Grundsätzen der fachlichen Unabhängigkeit, Neutralität und Objektivität verpflichtet sind. Bildungsstatistiken wurden im 19. und 20. Jahrhundert vor allem herangezogen, um staatlicherseits die Verwaltung und Res-

sourcenausstattung im Bildungswesen zu steuern. Zwar gab es bereits statistische Untersuchungen in der empirischen Sozialforschung, die mit den Arbeiten von Weber, Durkheim und Geiger ihren Höhepunkt hatten (Lüttinger & Schimpl-Neimanns, 1992). Jedoch kamen erst nach 1945 systematische Bestrebungen auf, eine möglichst einheitliche, überregionale Sozial- und – darin eingebundene – Bildungsstatistik aufzubauen. Hierfür waren auch steigende internationale (An-)Forderungen ausschlaggebend: Unter dem Schlagwort ‚social indicator movement‘ wurde die Sammlung und Analyse differenzierter, aktueller Daten zu sozialen Themen wie Bildung seit den 1970er Jahren maßgeblich von supranationalen Organisationen wie der OECD oder der Weltbank vorangetrieben (vgl. van Ackeren, 2003).

Durch diese und weitere Akteur:innen wurde und wird der wachsende und weiter harmonisierte Datenbestand zunehmend genutzt, um vergleichende Statistiken nationaler Bildungssysteme im globalen Zusammenhang mit Blick auf Innovationskraft, (ökonomische) Leistungsfähigkeit oder auch Teilhabe zu analysieren. Gleiches gilt für bildungsstatistische Analysen auf nationaler Ebene, die vor allem im Vergleich der Bundesländer etwa in der Arbeit und Programmatik deutscher Stiftungen an Bedeutung gewonnen haben. Die gesetzgeberischen Entscheidungen, welche statistischen Merkmale und Ausprägungen auf welche Weise erhoben bzw. (öffentlich) bereitgestellt werden, bleiben davon in einem auf Rechenschaftslegung und Transparenz ausgerichteten Bildungssystem nicht unberührt. Je häufiger auf Basis offizieller Daten negative Ergebnisse bzw. Schlagzeilen (re-)produziert werden und damit politischen Handlungsdruck erzeugen, desto vorsichtiger scheinen die bildungspolitischen Entscheidungen über die Erhebung, Bereitstellung und Veröffentlichung von Statistiken getroffen zu werden. Unterschiedliche Interessen wirken dabei in einem Geflecht aus wissenschaftlichen, politisch-administrativen sowie ökonomischen Akteur:innen zusammen (vgl. Forster, 2014; Hartong & Förchler, 2020).

Der beschriebene Entwicklungsprozess seit den 1950er Jahren lässt sich „als verstärkte Differenzierung der Statistiken und gesteigerte Orientierung an bildungspolitisch interessierenden Fragekomplexen“ (Köhler, 1980, S. 1248) charakterisieren. Dominierten lange Zeit aggregierte Rahmendaten über das Bildungswesen, so ging man zur Erfassung immer tiefer gegliederter Merkmalskombinationen über. Mit fortschreitender technologischer Entwicklung wurden auch alle bundesgesetzlich geregelten amtlichen Statistiken zu Bildungseinrichtungen, -teilnehmer:innen und -personal auf Individualdaten umgestellt. Einzig und allein in der föderal organisierten Schulstatistik liegen in einigen Bundesländern bis heute weiterhin nur Summendaten als zusammengefasste Meldung je Einzelschule vor. In Kindertageseinrichtungen, der betrieblichen Ausbildung und an Hochschulen werden demgegenüber inzwischen Einzeldaten für das gesamte Bundesgebiet erhoben. Dies geht mit neuen, modernen Dateninfrastrukturen und Mechanismen der Datenproduktion und -auswertung einher.

Für die Bildungsstatistik im engeren Sinne ist charakteristisch, dass es sich um regelmäßige Meldestatistiken der Bildungseinrichtungen bzw. -anbieter handelt. Gegenstand dieser prozessgenerierten Daten aus dem Verwaltungsgang sind ausschließlich Sachverhalte und Abläufe in Einrichtungen des formalen Bildungssystems. Differen-

ziert nach den Sektoren und Zuständigkeiten innerhalb des Bildungswesens lässt sich so auch die amtliche Bildungsstatistik weiter bereichsspezifisch aufschlüsseln:

- Angaben zur Kindertagesbetreuung in Kindertageseinrichtungen und öffentlich geförderter Tagespflege werden auf bundeseinheitlicher Grundlage des Sozialgesetzbuchs durch die ‚Kinder- und Jugendhilfestatistik‘ der Statistischen Ämter des Bundes und der Länder erhoben.
- Angaben zum allgemeinbildenden und zum beruflichen Schulwesen werden in der ‚Schulstatistik‘ auf der Grundlage der Landesgesetze und -verordnungen dezentral durch die Statistischen Landesämter, teilweise auch durch die zuständigen Landesministerien selbst erfasst.
- Angaben über das Ausbildungsgeschehen außerhalb beruflicher Schulen (Schulstatistik) liefert die ‚Berufsbildungsstatistik‘ der Statistischen Ämter auf Basis des Berufsbildungsgesetzes sowie die im Sozialgesetzbuch geregelte ‚Ausbildungsstellenmarktstatistik‘ der Bundesagentur für Arbeit.
- Angaben zur akademischen Bildung werden als ‚Hochschulstatistik‘ durch die Statistischen Ämter des Bundes und der Länder auf der Grundlage des Hochschulstatistikgesetzes erhoben.
- Der Weiterbildungsbereich wird durch keine amtliche Institutionenstatistik umfassend abgedeckt. Lediglich zu den Volkshochschulen sowie zu den von der Bundesagentur für Arbeit finanzierten beruflichen Fort- und Weiterbildungsmaßnahmen liegen bundesweit statistische Informationen vor.

## 2.2 Erhebungsmodalitäten der Bildungsstatistik

Die genannten Erhebungen werden zwar noch durch andere, bereichs- und themenübergreifende amtliche Daten wie etwa die Personalstandstatistik, (Bildungs-)Finanzstatistik oder qualifikationsspezifische Bevölkerungs- und Arbeitsmarktstatistiken ergänzt, sie bilden aber gleichwohl den Kern der Bildungsstatistik. Anders als forschungsgenerierte Befragungsdaten werden sie nicht bei den einzelnen Bildungsteilnehmer:innen und Lehrenden persönlich, sondern in Form von Vollerhebungen mit Auskunftspflicht durch mindestens jährliche Abfrage bei den Einrichtungen erfasst. Das Datenmaterial zeichnet sich damit einerseits durch seine flächendeckende Verfügbarkeit in Zeitreihe aus, was Trendbeschreibungen ermöglicht und auch präzise die Situation kleinerer Populationen bzw. Regionen abbildet. Andererseits liefert der amtlich erfasste Merkmalskanon nur solche Informationen, die vom Gesetzgeber für das Verwaltungshandeln als erforderlich angesehen werden. Angaben zur sozialen Einbindung in andere Lebensbereiche – z. B. Bezüge zu den Sozialisationsprozessen in Familie, Beruf oder Freizeit – fehlen hingegen weitgehend.

Für die einzelnen Bildungsbereiche erfolgen die Datenerhebung (Erhebungstermin, -einheiten, -inhalte etc.) und Datenweitergabe entsprechend der jeweiligen Gesetze auf unterschiedliche Art und Weise. Schon für die landesgesetzlich geregelten Schulsta-

tistiken unterscheiden sich die Erhebungsmodalitäten so sehr zwischen den Ländern, dass eine systematische Recherche und Dokumentation eine eigene Forschungsarbeit erfordern würde. Teilweise werden die Daten noch als Summendaten auf Klassen- und Schulebene, teilweise als Individualdaten der Schüler:innen und Lehrkräfte erfasst, in Hessen und Hamburg sogar mit datenschutzkonformer Personenkennung über die Zeit. Auch der Merkmalskanon und die Merkmalsausprägungen, also welche Informationen in welcher Kategorisierung erfasst werden, differieren erheblich, trotz gemeinsamer Beschlüsse der Kultusministerkonferenz (KMK) über die Einführung eines Kerndatensatzes für schulstatistische Individualdaten (vgl. Böttcher & Kühne, 2017; Mundelius, 2019). Es bleibt abzuwarten, ob der neuerliche KMK-Beschluss aus dem Jahr 2020 nun in den Ländern dazu führt, ihre „Zusammenarbeit bei der ständigen Weiterentwicklung der Bildungsstatistiken“ zu verstärken sowie „abgestimmte Verfahren“ und „einheitliche Festlegungen zu den Erhebungsmerkmalen und ihrer Definitionen“ zu sichern (KMK, 2020).

Will man sich einen Überblick verschaffen, welche Beobachtungsgegenstände und Merkmale gegenwärtig überhaupt in den Ländern vorliegen, reicht ein Blick in die zugehörigen statistischen Veröffentlichungen nicht aus. Sie stellen nämlich nur den kleinsten gemeinsamen Nenner aller Landesstatistiken dar. Das vollständige Analysepotenzial lässt sich nur durch eine Sichtung der jeweiligen Gesetze erschließen. Auch bei den übrigen, bundesgesetzlich geregelten Bildungsstatistiken empfiehlt sich immer zu prüfen, wie die Gegenstände und Erhebungsmodi der Statistiken konkret normiert und dokumentiert sind, um ein Verständnis dafür zu entwickeln, wie die Daten ausgewertet und gedeutet werden können.

Neben der Uneinheitlichkeit gesetzlicher Vorgaben erwachsen weitere Herausforderungen der Generierung und Weitergabe schulstatistischer Daten aus der fortschreitenden Digitalisierung der Dateninfrastruktur. So ist schon jetzt von erheblichen Diskrepanzen in der Ressourcenausstattung zwischen einzelnen Schulen und Schulträgern auszugehen, sowohl mit Blick auf die IT-Infrastruktur als auch hinsichtlich der erforderlichen Kompetenzen verantwortlicher Personen im Umgang mit digitalen Daten (vgl. Williamson, 2016). Diese Ressourcenunterschiede – etwa in der Verfügbarkeit und Handhabung hochwertiger Schulverwaltungssoftware – erklären bis zu einem gewissen Grad auch die unterschiedliche Geschwindigkeit, mit der die Umstellung auf schulstatistische Individualdaten erfolgen konnte oder eben nicht. Durch zunehmende Einflussnahme von privaten intermediären Akteur:innen, welche Serviceangebote für die Datengenerierung bereitstellen sowie Datenstandards vorantreiben (vgl. Hartong & Förschler, 2020), könnten derartige Differenzen künftig sogar noch verstärkt werden. Wie im privaten Leben entstehen mit der Digitalisierung so z. B. größere Risiken, künftig als Bildungseinrichtung, -teilnehmer:in oder Lehrperson jenseits der statistischen Zweckbestimmung durch Algorithmen von kommerziell ausgerichteten Agenturen oder Vertragspartner:innen ungewollt oder sogar unwissentlich kategorisiert, klassifiziert oder sortiert zu werden. Obgleich gewinnorientierte Unternehmen in Deutschland für die Erhebung und Weitergabe schulstatistischer Daten bislang von geringerer Relevanz sind als andernorts, beginnt sich – befürwortet von staatlichen Akteur:innen wie der

KMK und dem BMBF – durch den Ausbau des EdTech-Markts insbesondere im Bereich der Digitalisierung von Unterricht ein zunehmender Einfluss abzuzeichnen (Hartong & Förchler, 2020).

### **3. Datenbereitstellung und Datenauswertung am Beispiel schulstatistischer Analysen**

Die mit amtlichen Daten verbundenen Forschungspotenziale unterscheiden sich nicht nur aufgrund ihrer jeweiligen Generierung, sondern auch mit Blick auf die Bereitstellung des erhobenen Datenmaterials. Über die periodischen Veröffentlichungen der statistischen Ämter des Bundes und der Länder frei verfügbar und damit für jede:n Interessierte:n leicht zugänglich sind (nur) ausgewählte Aggregatdaten. Auf kleinräumiger Ebene stellen die statistischen Landesämter über eine gemeinsame Bildungsdatenbank ebenfalls in regelmäßigem Turnus öffentlich Daten bereit. Statistische Summendaten reichen jedoch oftmals nicht, um Merkmale erfolgreicher Bildungsbiographien oder Risikofaktoren für einen gelingenden Bildungserwerb zu beschreiben und daraus Interventionen abzuleiten (Artelt, Bug, Kleinert, Maaz & Runge, 2019; Böttcher & Kühne, 2017). Um an differenzierte Merkmalskombinationen oder Daten auf Individualebene zu gelangen, können Forschende bei den statistischen Ämtern oder den Statistikreferaten der ministeriellen Verwaltung in Bund und Ländern die Bereitstellung von Daten beantragen oder externe Sonderauswertungen in Auftrag geben. Ferner besteht die Möglichkeit, über die Forschungsdatenzentren des Bundes und der Länder auf standardisierte Datenprodukte zuzugreifen oder die Bereitstellung von Mikrodaten außerhalb des Standardportfolios der amtlichen Veröffentlichungen zu beantragen. Mit Fokus auf das allgemeinbildende Schulwesen werden diese verschiedenen Wege des Datenzugangs im Folgenden erläutert und die damit verbundenen Analysepotenziale anhand ausgewählter Beispiele veranschaulicht. Dabei wird zwischen regelmäßigen Veröffentlichungen (3.1), laufend aktualisierten Bildungsdatenbanken (3.2), Sonderaufbereitungen auf Anfrage (3.3) und der Beantragung von Mikrodaten zur eigenen Bearbeitung (3.4) unterschieden.

#### *3.1 Veröffentlichungen der amtlichen Statistik*

Für die Schulstatistik werden die von den Schulträgern übermittelten, komplexen Datensätze in den statistischen Landesämtern zusammengeführt und in aggregierter Form auszugsweise über Publikationsreihen veröffentlicht. „In aller Regel handelt es sich um knappe tabellarische Zusammenstellungen von Struktur- und Basisdaten, d. h. wenigen zentralen Merkmalskombinationen zu den einzelnen schulstatistischen Datensegmenten“ (Kühne, 2015, S. 116). Parallel dazu erfolgt die Übermittlung einer festgelegten Schnittmenge der Landesdaten an das Statistische Bundesamt und das Sekretariat der KMK, welche die Datensätze wiederum nach einheitlichen Kriterien aggregieren, in

weiteren statistischen Publikationen veröffentlichen sowie in andere Berichterstattungssysteme (z. B. Integrierte Ausbildungsberichterstattung) integrieren.

In der frei verfügbaren Fachserie 11 des Statistischen Bundesamtes wird jedes Jahr in verschiedenen Reihen unter anderem zu allgemeinbildenden und beruflichen Schulen eine Gesamtübersicht mit wesentlichen Eckzahlen zu Schulen, Klassen, Schüler:innen, Absolvent:innen/Abgänger:innen, Lehrkräften und Unterrichtsstunden präsentiert. Dass die auf diesem Wege veröffentlichten Datentabellen oftmals unter komplexen, technischen Überschriften lange Zahlenreihen zu scheinbar profanen Beobachtungsgegenständen aneinanderreihen, kann den Blick für den eigentlichen Informationswert verstellen. Denn die Merkmalskombinationen auf Bundes- und Landesebene ermöglichen deutlich mehr als eine globale Beschreibung von Bestandsgrößen. In nur einer Tabelle können sämtliche Daten enthalten sein, die man braucht, um z. B. alle gängigen Übergangs-, Wechsel- und Verbleibsquoten im Schulwesen ermitteln zu können.<sup>1</sup> Darüber hinaus eignen sich die Daten auch zur Berechnung komplexerer Kennziffern, wie die im Folgenden exemplarisch dargestellten quasi-längsschnittlichen Durchgangsquoten zeigen.

### **Durchgangsquoten und -wahrscheinlichkeiten als Beispiel für die Ableitung von Kennziffern aus Aggregatdaten**

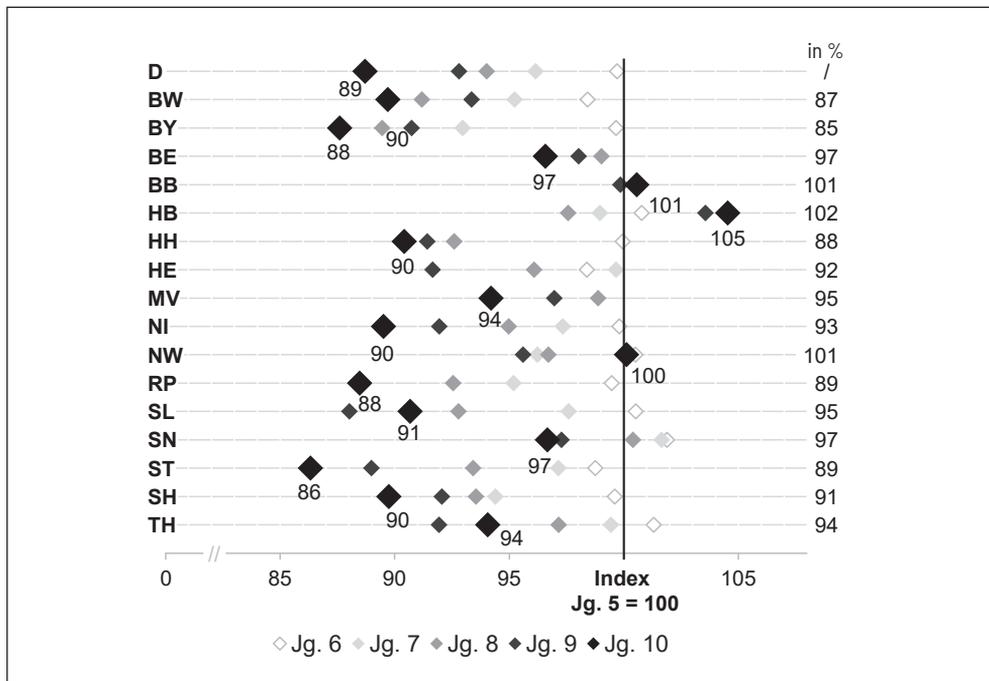
Der Übergang auf die weiterführenden Schulen des Sekundarbereichs ist für die individuelle Entwicklung von Schüler:innen ein einschneidendes Ereignis. Um zu untersuchen, inwiefern diese Übergangentscheidung am Ende der Grundschulzeit über die weitere Schullaufbahn von Kindern und Jugendlichen entscheidet, reichen die üblichen globalen Wechselquoten je Schuljahr nicht aus. Daher wurde im Rahmen der nationalen Bildungsberichterstattung 2020 die amtliche Schulstatistik mehrerer Jahre verknüpft, um Übergänge über die Jahrgangsstufen hinweg in einer kohortenspezifischen Betrachtung zu untersuchen und die Verbleibchancen von Schüler:innen an verschiedenen Schularten in den einzelnen Ländern zu ermitteln. Um einen länderübergreifenden Vergleich zu ermöglichen, wird hier exemplarisch auf die Gymnasien fokussiert. Ausgehend von Fünftklässler:innen im Schuljahr 2013/14 lässt sich mittels Durchgangsquotenverfahren die kohortenspezifische Entwicklung der Schüler:innenzahlen nachzeichnen. Dazu wird der Schüler:innenbestand des Ausgangsjahres in Relation zur Schüler:innenzahl in der nächsthöheren Jahrgangsstufe des darauffolgenden Schuljahres gesetzt. Die quantitative Veränderung bildet demnach den Saldo aus Zu- und Abgängen ab.

Im Ergebnis zeigen sich für die betrachtete Kohorte an Gymnasien im bundesdeutschen Durchschnitt Schüler:innenzahlverluste von 11 % (vgl. Autorengruppe Bildungsberichterstattung, 2020, S. 112). In nahezu allen Ländern ist im Verlauf der Jahrgangsstufen 5 bis 10 ein sukzessiver Rückgang der Gymnasialbeteiligung zu verzeichnen, welcher sich vor allem auf das Wechselverhalten der Jugendlichen zurückführen lässt

1 Tabelle 3.7 der Fachserie 11, Reihe 1. [https://www.destatis.de/DE/Service/Bibliothek/\\_publikationen-fachserienliste-11.html](https://www.destatis.de/DE/Service/Bibliothek/_publikationen-fachserienliste-11.html) [22.07.22].

(vgl. Autorengruppe Bildungsberichterstattung, 2020, S. 111). Frühzeitige Abgänge vom Gymnasium können nur in Bremen, Brandenburg und Nordrhein-Westfalen durch entsprechende Zugänge kompensiert werden.

Durch Multiplikation der Durchgangsquoten lässt sich sodann näherungsweise die Wahrscheinlichkeit ermitteln, mit der Schüler:innen, die nach der Grundschule auf ein Gymnasium übergehen, dieses bis zum Ende der 10. Jahrgangsstufe vollständig durchlaufen (Durchgangswahrscheinlichkeit). Um temporäre Schwankungen der Schüler:innenzahl (z. B. aufgrund der Geburtenentwicklung oder regionaler Zu-/Abwanderung) zu reduzieren, wird dabei die mittlere Schüler:innenzahl aus drei aufeinanderfolgenden Erhebungsjahren herangezogen (vgl. Autorengruppe Bildungsberichterstattung, 2020). Dabei kann es zu Wahrscheinlichkeiten von 100 % und mehr kommen, wenn im Verlauf des Sekundarbereichs I mehr Zu- als Abgänge zu verzeichnen sind. Es zeigt sich zum einen, dass die Verbleibchancen am Gymnasium erwartungsgemäß in Baden-Württemberg (87 %) und Bayern (85 %) am geringsten sind (Abb. 1). Zum anderen weist auch



\* Durchgangsquotenverfahren, vgl. Autorengruppen Bildungsberichterstattung, 2020.

\*\* Für die Länder BE, BB und MV ist die Jg. 7 im Schuljahr 2015/16 der Ausgangspunkt; Werte für Deutschland schließen diese Länder nicht ein.

Quelle: eigene Darstellung

Abb. 1: Veränderung der Schüler:innenzahlen im Verlauf der Jahrgangsstufen 5 bis 10 (2013/14 bis 2018/19)\* sowie Durchgangswahrscheinlichkeit an Gymnasien nach Ländern (Index Jg. 5 = 100)\*\*

Hamburg mit 88% einen geringen Wert auf, der vor allem auf die in diesem Bundesland obligatorischen Schulartwechsel vom Gymnasium bei unzureichenden Leistungen am Ende der Jahrgangsstufe 6 zurückgeführt werden kann (vgl. § 13 Absatz 2 APO-GrundStGy). Die höchsten Werte erreichen Bremen, Brandenburg und Nordrhein-Westfalen mit jeweils knapp über 100%. Während die Durchgangswahrscheinlichkeiten in Bremen und Brandenburg tatsächlich auf höhere Verbleibchancen am Gymnasium im Verlauf des Sekundarbereichs I hindeuten, geht der hohe Wert für Nordrhein-Westfalen maßgeblich auf Zugänge zum Gymnasium in der 10. Jahrgangsstufe zurück, die die vorzeitigen Abgänge (-4,4 Prozentpunkte zwischen Jahrgangsstufe 5 und 9) ausgleichen. Ein größerer Teil der Schülerschaft entscheidet sich hier also erst am Ende des Sekundarbereichs I für die Fortsetzung der allgemeinbildenden Schullaufbahn am Gymnasium.

Wie zuvor dargelegt, geht die landesgesetzliche Regelung der Schulstatistik mit einer großen Vielfalt an Erhebungsmodalitäten einher, sodass länderübergreifende Auswertungen mit Einschränkungen verbunden sind. Bei den beschriebenen Analysen können beispielsweise für das Land Hessen die Durchgangsquoten nur bis zur 9. Jahrgangsstufe und die Durchgangswahrscheinlichkeit nur für das Erreichen ebendieser ermittelt werden (Abb. 1). Der Grund hierfür ist, dass die Schüler:innen der Eingangsphase an G8-Gymnasien (10. Jahrgangsstufe) mit jenen im Sekundarbereich II der G9-Gymnasien zusammen ausgewiesen werden. Schwerer noch wiegt das grundlegende Defizit der Schulstatistik, dass die Kinder und Jugendlichen nicht nach dem besuchten Bildungsgang (Haupt-, Realschul- und Gymnasialbildungsgang) erfasst werden. Innerhalb der zunehmend verbreiteten Schularten mit zwei oder drei Bildungsgängen lassen sich Wechsel zwischen den Schulzweigen bzw. Lerngruppen daher in keinem Land statistisch abbilden. Das statistische Bild der Auf- und Abwärtsmobilität im deutschen Schulwesen bleibt damit bislang unvollständig.

### 3.2 *Bildungsdatenbanken*

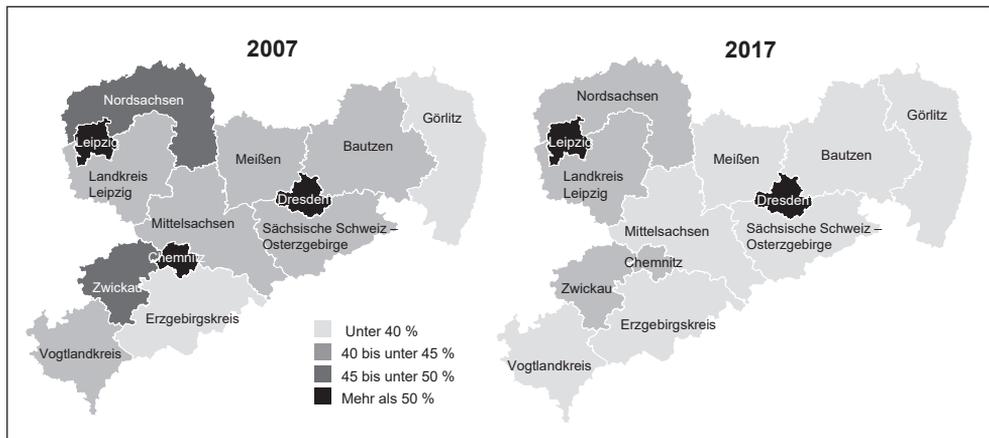
Neben den regelmäßigen Publikationsreihen, die getrennt für einzelne Erhebungsjahre und Bildungsbereiche Struktur- und Basisdaten zur Verfügung stellen, wird ein Teil der Bildungsstatistiken auch in jahres- und bereichsübergreifenden Datenbanken zusammengeführt und laufend aktualisiert. Neben der Hauptdatenbank des Statistischen Bundesamtes (Genesis Online) seien als Beispiel die internationale Bildungsdatenbank OECD.Stat oder das Datensystem DAZUBI des Bundesinstitutes für Berufsbildung angeführt. Die meisten Datenbanken bieten den Nutzer:innen anstelle statisch tabellierter Merkmalskombinationen interaktive Auswahlmöglichkeiten und eignen sich insbesondere bei Forschungsvorhaben, die auf speziell zugeschnittene oder differenziertere Informationen angewiesen sind. Ein besonderer Fokus soll nachfolgend auf die fachlich und regional tiefer gegliederten Datenaufbereitungen gelegt werden. Die von den statistischen Landesämtern erfassten, dezentral vorgehaltenen Bildungsdaten auf Landes- und kommunaler Ebene können etwa über eine gemeinsame Daten-

bank (Kommunale Bildungsdatenbank) kostenfrei abgerufen werden. Für verschiedene Merkmalskombinationen stehen somit auf kleinräumiger Ebene jährlich aktualisierte Bestandsdaten und abgeleitete Kennziffern aus den diversen Bildungsstatistiken zur Verfügung. Darüber hinaus hält die Datenbank Informationen zu gesellschaftlichen Rahmenbedingungen aus dem von den Statistischen Ämtern des Bundes und der Länder verantworteten Regio-Stat-Katalog vor. Über die Kommunale Bildungsdatenbank werden also standardisierte und repräsentative Regionaldaten öffentlich zugänglich gemacht, die es ermöglichen, Stand und Entwicklungsperspektiven in den einzelnen Bildungsbereichen auf kleinräumiger Ebene aufzuschlüsseln, wie das folgende Beispiel veranschaulicht.

### **Übergangsquoten im kommunalen Vergleich als Beispiel für die Analyse kleinräumiger Disparitäten**

In ihrer Gesamtschau von Einzelbefunden hat die Autorengruppe Bildungsberichterstattung im Jahr 2018 den Trend zu höherer Bildung als eine wesentliche Entwicklungslinie in Deutschland identifiziert. Für den Schulbereich lieferten unter anderem schulstatistische Daten zum Übergang von der Grundschule in die weiterführenden Schulen des Sekundarbereichs dafür ein Indiz. In der Entwicklung dieser Übergangsquoten zeigten sich dabei bereits im Ländervergleich teils erhebliche Unterschiede (vgl. Autorengruppe Bildungsberichterstattung, 2018, S. 94). Beispielsweise sank in Sachsen zwischen 2007 und 2017 der Anteil der Schüler:innen, die im Anschluss an die Grundschule den direkten Weg zum Abitur am Gymnasium einschlagen, von 46 auf 43% (vgl. Kühne, Maaz, Mank, Ordemann & Schulz, 2019). Diese von den meisten anderen Ländern abweichende Trendentwicklung war unter anderem darauf zurückzuführen, dass den beruflichen Qualifikationswegen in Sachsen ein hoher Stellenwert eingeräumt wird, indem seitens der Bildungspolitik und der Kammern die weiterführende Schulart ohne Abituroption (Oberschule), die duale Ausbildung oder berufliche Wege zur Studienberechtigung gegenüber der Gymnasiallaufbahn bewusst gestärkt werden.

Die über die Kommunale Bildungsdatenbank verfügbaren Daten zeigten allerdings innerhalb Sachsens unterschiedliche Entwicklungen auf kleinräumiger Ebene. In Auswertungen für den sächsischen Landesbildungsbericht schien die Gymnasiallaufbahn vor allem in den Ballungszentren Dresden und Leipzig von höherer Bedeutung zu sein; zum einen ging 2017 in beiden Städten jeweils mehr als die Hälfte der Schüler:innen nach der Grundschule in ein Gymnasium über, wohingegen es in den meisten anderen Kreisen nur jedes dritte Kind war. Zum anderen blieben die Übergangsquoten in Dresden und Leipzig auch im zeitlichen Verlauf auf hohem Niveau, während die Mehrzahl der Landkreise sowie die Stadt Chemnitz erhebliche Rückgänge bei den Gymnasialübergängen zu verzeichnen hatten (vgl. Kühne et al., 2019, S. 14; Abb. 2). Im Unterschied zu dem auf Bundesebene und in vielen anderen Ländern anhaltenden, langjährigen Trend zum Gymnasialbesuch zeichneten die kleinräumigen Analysen also ein weit differenzierteres Bild. In einer Gesamtschau mit weiteren regional aufgeschlüsselten Kennziffern der frühen und schulischen Bildung ergaben sich schließlich unterschiedliche kommunale Profile der Bildungsentwicklung, aus denen jeweils spezifische Her-



Quelle: Kühne et al., 2019, S. 15

Abb. 2: *Gymnasialbeteiligung in Sachsen 2007 und 2017 nach Kreisen (in % aller Fünftklässler:innen, die zuvor die Grundschule besuchten)*

ausforderungen und Entwicklungsperspektiven für die Entscheidungsträger:innen vor Ort abgeleitet werden konnten (vgl. Kühne et al., 2019, S. 32–35).

Bei Datenauswertungen dieser Art ist zu berücksichtigen, dass es aufgrund des eingangs aufgezeigten Mangels an Standardisierung der Zulieferungen und landesspezifischen Analysemethoden zu Unstimmigkeiten zwischen den schulstatistischen Veröffentlichungen auf Bundesebene und den landesspezifisch bereitgestellten Daten kommen kann. So weichen beispielsweise die Quoten der Schulartwechsel für Bayern im länderübergreifenden Vergleich deutlich von den Werten in der landesspezifisch aufbereiteten Kennziffer ab, die über die Kommunale Bildungsdatenbank bereitgestellt wird. Dies unterstreicht nochmals die Notwendigkeit, vorab die zugrundeliegenden Erhebungskonstrukte und Operationalisierungen eingehend zu prüfen und weitere, oft in detaillierten Fußnoten enthaltene Hinweise in die Interpretation der Daten einfließen zu lassen. Die vertiefte Auswertung solcher amtlichen und öffentlich zugänglichen Landes- und Kommunaldaten eröffnet dann für die Bildungsforschung großes Potenzial, steuerungsrelevante Erkenntnisse zu generieren.

### 3.3 (Externe) Sonderaufbereitungen

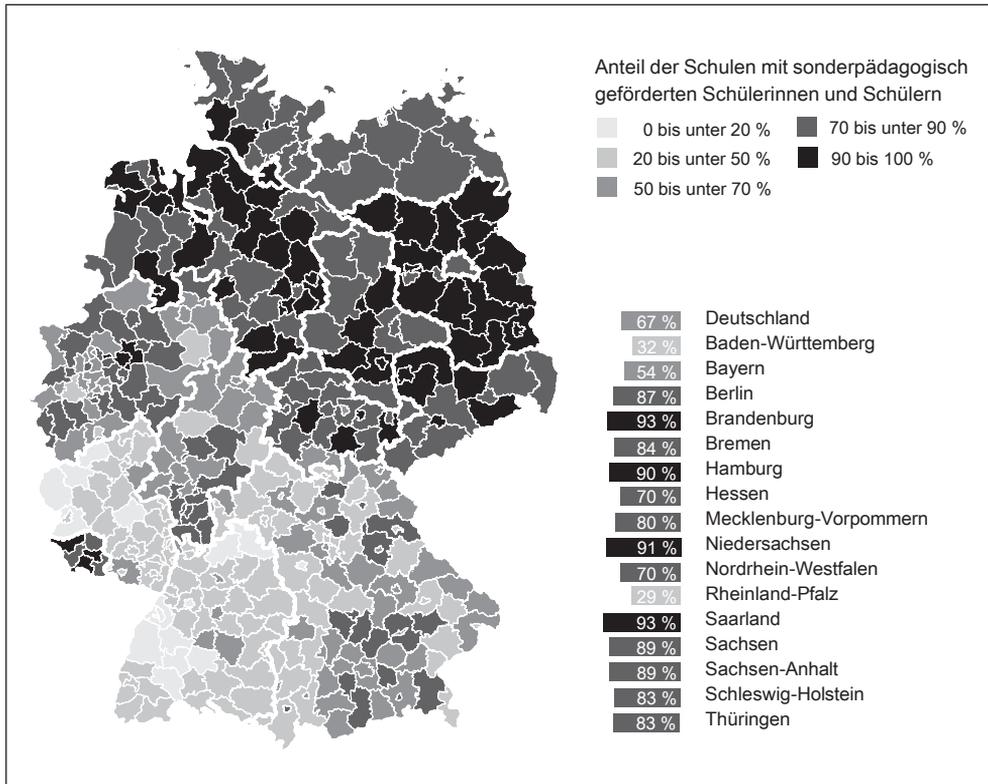
Um an differenzierte Merkmalskombinationen zu gelangen, die nicht im Standardportfolio amtlicher Veröffentlichungen enthalten sind, steht grundsätzlich allen Forschenden die Möglichkeit offen, diese bei den Statistischen Landesämtern sowie ministeriellen Statistikstellen (z. B. Sekretariat der KMK) abzufragen und auszuwerten. Je nach Fragestellung und ersten Überlegungen zur möglichen Operationalisierung sollte dafür anhand verfügbarer Dokumentationen oder Gesetzestexte zur infrage kommenden

Statistik zunächst die grundsätzliche Eignung der Daten (verfügbare Merkmale etc.) geprüft werden. Anschließend empfiehlt es sich, mit den Ansprechpartner:innen der statistischen Ämter bzw. der entsprechenden Statistikstellen in den Austausch zu treten und dabei den angedachten Verwendungszweck sowie die benötigten Merkmale und Erhebungszeitpunkte für eine Datenanfrage so konkret wie möglich zu formulieren (z. B. anhand von Merkmalsübersichten, Referenztabellen/-vorlagen oder Berechnungsformeln). Die Realisierung einer solchen externen Datenanfrage ist dabei auch davon abhängig, welcher zeitliche und finanzielle Aufwand im Einzelfall entsteht. Derartige Sonderauswertungen sind typischerweise erforderlich, wenn in den statistischen Veröffentlichungen oder Datenbanken durch die Standardisierung und Aggregation der erhobenen Daten Informationen verloren gehen, die für die jeweiligen Analysen benötigt werden – z. B. bei sehr kleinräumigen Analysevorhaben oder Operationalisierungen, die einen Rückgriff auf die Einzelangaben zu Schulen, Klassen oder sogar Schüler:innen erfordern (sofern diese in dem jeweiligen Land zur Verfügung stehen).

### **Regionale Verteilung inklusiver Schulangebote als Beispiel für die externe Abfrage und Aufbereitung landesspezifischer Daten**

Einen wesentlichen Impuls für eine gleichberechtigte Teilhabe und den Abbau von Barrieren im Schulwesen lieferte die 2009 von Bund und Ländern ratifizierte UN-Konvention für die Rechte von Menschen mit Behinderungen, in deren Folge im letzten Jahrzehnt verstärkte Inklusionsbemühungen im Schulwesen zu beobachten sind. Im Zusammenhang mit Fragestellungen zu Heterogenität im schulischen Kontext nimmt Inklusion auch in der Schulforschung mittlerweile einen zentralen Stellenwert ein, nicht zuletzt da die Feststellung eines sonderpädagogischen Förderbedarfs für die Bildungsteilhabe von Kindern und Jugendlichen und ihre individuelle Entwicklung weitreichende Folgen haben kann.

Auf Grundlage öffentlich zugänglicher Daten können die Inklusionsbemühungen, z. B. anhand von Förder- und Inklusionsquoten (zur Interpretation dieser Kennziffern siehe Autorengruppe Bildungsberichterstattung, 2020, S. 116–117), in ihrer zeitlichen Entwicklung nachgezeichnet werden, allerdings sind weitere Vergleiche nur auf Länderebene möglich. Im nationalen Bildungsbericht 2018 richtete sich ein vertiefender Blick auf die regionale Infrastruktur inklusiver Schulangebote. Dies erforderte kleinräumige Daten zu allgemeinbildenden Schulen, die mindestens eine:n sonderpädagogisch geförderte:n Schüler:in unterrichten, welche mittels einer durch die KMK koordinierten Länderumfrage zur Verfügung gestellt werden konnten. Im Ergebnis zeigten sich auf kleinräumiger Ebene beträchtliche Differenzen (Abb. 3): Während sich für die nördlichen und östlichen Länder ein nahezu flächendeckendes Angebot an Regelschulen abzeichnete, die Kinder und Jugendliche mit sonderpädagogischer Förderung unterrichten, wiesen insbesondere Bayern, Hessen und Nordrhein-Westfalen teils erhebliche Unterschiede zwischen den Kreisen auf (vgl. Autorengruppe Bildungsberichterstattung, 2018). Die Sonderaufbereitung der Landesstatistiken lieferte erste Hinweise auf die Umsetzungsstrategien der aktuellen Inklusionsbestrebungen, erlaubte jedoch weder Rückschlüsse darauf, ob die Schüler:innen gemeinsam in Regelklassen unterrich-



\* Ohne Förderschulen

1) Für Hessen werden Daten auf Ebene der staatlichen Schulämter herangezogen und für die zugrundeliegenden Kreise dargestellt.

2) Ohne Schüler:innen, die ohne förmlich festgestellten Förderbedarf sonderpädagogisch gefördert werden.

3) Ohne Schüler:innen, die im Rahmen sogenannter sonderpädagogischer Dienste gefördert werden.

Quelle: Autorengruppe Bildungsberichterstattung, 2018, S. 106

Abb. 3: Anteil der allgemeinbildenden Schulen\*, die Schüler:innen mit sonderpädagogischer Förderung unterrichten, 2016/17 nach Kreisen und Ländern (in %)

tet werden, noch auf Qualitätsmerkmale der (inkluisiven) Unterrichtsbedingungen. Um Aussagen über die Lerngruppen und -arrangements in den Schulen und Klassen zu treffen, bedürfte es einer Auswertung von Mikrodaten zur sonderpädagogischen Förderung auf der Ebene einzelner Klassen und Schulen. Diese würde aber den vertretbaren Rahmen einer externen Sonderaufbereitung sprengen und wäre eher über andere Zugangswege realisierbar, die nachfolgend vorgestellt werden.

### 3.4 Mikrodatenzugriff über die Forschungsdatenzentren

Die Prozesse für eine bessere Nutzbarmachung von Bildungsdaten für die (Bildungs-) Forschung wurden in den letzten beiden Jahrzehnten vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) maßgeblich vorangetrieben (vgl. Artelt et al., 2019). Seit 2002 steht Forschenden aus unabhängigen wissenschaftlichen Einrichtungen mit den Forschungsdatenzentren (FDZ) der Statistischen Landesämter sowie des Statistischen Bundesamtes ein institutionalisierter Zugang zu (Mikro-)Daten der amtlichen Statistik zur Verfügung. Um die Forschungsdateninfrastruktur für die Wissenschaft zu verbessern, führen die FDZ die (dezentral vorgehaltenen) Mikrodaten der Länder an einer zentralen Stelle zusammen und bereiten sie auf. Informationen zur Datensatzstruktur sowie zur Erhebung und Qualität der Daten werden dabei durch Metadatendokumentationen nachvollziehbar gemacht (Zühlke & Christians, 2005).

Neben standardisiert vorliegenden Produkten werden Daten in den FDZ entsprechend des Untersuchungsinteresses und der Anonymisierungserfordernisse individuell angepasst und zur Verfügung gestellt. Forschende können darüber hinaus die Nutzung von Daten, die nicht im Standardportfolio der FDZ vorhanden sind, beantragen. Nach Prüfung der formalen Zugangsvoraussetzungen und der Wahl der geeigneten Auswertungsstrategie erfolgt der Zugang zu den amtlichen Mikrodaten über einen der nachstehend beschriebenen Wege (vgl. Zühlke & Christians, 2005):

Der Zugangsweg der *On-Site-Nutzung* eignet sich für die Analyse von nicht oder nur gering anonymisierten Datensätzen, die nicht die Räume der amtlichen Statistik verlassen dürfen:

- *Gastwissenschaftsarbeitsplätze*: In geschützten Räumen der amtlichen Statistik können Forschende auf faktisch oder formal anonyme Mikrodaten zugreifen und sich hierzu beraten lassen. In allen FDZ stehen ihnen Gastwissenschaftsarbeitsplätze zur Verfügung, an denen sie Datensätze einsehen und analysieren, jedoch nicht kopieren können. Nach erfolgter Prüfung der Analyseergebnisse auf Wahrung der statistischen Geheimhaltung durch die FDZ werden ihnen diese zur Verfügung gestellt.
- *Kontrollierte Datenfernverarbeitung*: Bei Daten, die aufgrund gesetzlicher Vorgaben nur durch die Datenhalter eingesehen werden dürfen, erhalten die Forschenden keinen direkten Zugriff auf die Mikrodaten, sondern auf Strukturdatensätze, die im Aufbau den Originaldaten entsprechen. Dies ermöglicht es ihnen, Auswertungssyntaxen zu erstellen, die in den FDZ mit dem Originaldatensatz ausgeführt werden. Auch hier werden die Ergebnisse vor der Weitergabe noch auf Geheimhaltung geprüft.

Durch die Aggregation oder Entfernung einzelner Merkmale können faktisch oder absolut anonymisierte Daten für die Nutzung außerhalb der Räume der amtlichen Statistik *off-site* angeboten werden:

- *Scientific Use Files (SUF)*: Forschende können gängige Statistiken in Form von standardisierten und faktisch anonymisierten Datensätzen auswerten.

- *Public Use Files (PUF)*: Auf absolut anonymisierte Mikrodatensätze, die für einige gängige Statistiken erstellt werden, können Forschende auf den Internetseiten der FDZ zugreifen. Auswerten können sie jedoch nur ausgewählte Merkmale ohne die Möglichkeit, kleinräumige Analysen vorzunehmen.
- *Campus Files (CF)*: Für den Einsatz in der akademischen Lehre können speziell konzipierte, ebenfalls absolut anonymisierte Datensätze von der Internetseite der FDZ kostenfrei heruntergeladen werden.

Direkte Zugriffsmöglichkeiten auf statistische Mikrodaten sind vor allem dann von Vorteil, wenn Fragestellungen multidimensionale Merkmalsverknüpfungen auf verschiedenen Beobachtungsebenen voraussetzen, aber Merkmalskatalog und -ausprägungen mit Blick auf die Fallzahlen nicht hinreichend antizipiert werden können und ein exploratives Vorgehen erfordern. Als Beispiel lassen sich hier Analysen zur räumlichen Verteilung und differenziellen Zusammensetzung von Lerngruppen auf Einzelschulebene anführen.

### **Schulische Segregation als Beispiel für die Analyse von Mikrodaten der Schüler:innen und Schulen**

In der Schulforschung erfahren Fragestellungen zur Komposition von Lerngruppen und deren Einfluss auf die individuelle, nicht nur leistungsbezogene sondern auch sozio-emotionale Entwicklung von Schüler:innen zunehmend Aufmerksamkeit. Diesem Gegenstand nähern sich einige Forschungsarbeiten aus einem heterogenitäts- bzw. diversitätsorientierten Blickwinkel (vgl. z. B. Rjosk, Lütke, Richter & Eccles, 2017; Edele et al., 2020), während ihn andere unter dem Gesichtspunkt der Segregation untersuchen (vgl. z. B. Makles, Schneider & Terlinden, 2019). Ungeachtet der Analyseperspektive sind solche Fragestellungen auch bildungspolitisch relevant, da sie mit Fragen der Chancengerechtigkeit und sozialer Teilhabe in Verbindung stehen.

Auf Grundlage der Individualstatistik des Landes Bremen untersuchten Makles et al. (2019) den Einfluss von Schulwahl auf die Segregation<sup>2</sup> an Grundschulen im Stadtstaat. Amtliche Daten sind hierfür besonders geeignet, da die Analysen zum einen Individualdaten der Schüler:innen mit regionalem Bezug erfordern. Zum anderen ist mindestens eine repräsentative Stichprobe, besser gar eine Vollerhebung notwendig, um auf dieser kleinräumigen Ebene zu aussagekräftigen Ergebnissen zu gelangen. Die Daten ermöglichen es zudem, die schulische Segregation in ihrer zeitlichen Entwicklung nachzuzeichnen (Makles et al., 2019, S. 177).

In der Bremer Individualstatistik liegen u. a. Informationen zum Wohnort der Schüler:innen vor. Diese können zum einen um soziodemografische Kennziffern wie den Anteil an Arbeitslosen oder Alleinerziehenden zur Beschreibung des sozialen Milieus ergänzt werden, welche wiederum die Berechnung eines Sozialindex ermöglichen. Zum

---

2 Segregation wurde für die Merkmale Migrationshinweis und Erstsprache des Kindes sowie auf Basis eines Sozialindex berechnet (für die Definition eines Migrationshinweises siehe Makles et al., 2019, S. 185).

anderen lassen sich durch Ergänzung um Informationen zum Schulbezirk die zuständigen Einzugschulen der Schüler:innen bestimmen. So war es möglich, die soziale und migrationsbezogene Zusammensetzung der Schülerschaft am Wohn- und Schulstandort zu bestimmen. Die Autor:innen konnten damit nicht nur das Ausmaß an Entmischung an Bremer Grundschulen berechnen, welches auf die residenzielle Segregation von Stadtbezirken zurückzuführen ist, sondern sie untersuchten darüber hinaus, inwieweit eine verstärkte Segregation durch die Schulwahl der Eltern erklärt werden kann und sicherten diesen Effekt zufallskritisch ab. Wenngleich sich für die Segregation nach Migrationshinweis und Erstsprache der Schüler:innen zu (nahezu) allen Zeitpunkten ein signifikanter Effekt der Schulwahl zeigte, war dieser als sehr gering einzuschätzen. Die Befunde weisen darauf hin, dass hinsichtlich der Segregation von Schulen nach wie vor der größte Einfluss auf die Komposition der Schülerschaft von der residenziellen Segregation ausgeht (Makles et al., 2019, S. 192).

#### **4. Desiderate und Entwicklungsperspektiven der amtlichen Bildungsstatistiken und ihrer Nutzung**

Obwohl sich in den letzten 20 Jahren die Verfügbarkeit und Zugänglichkeit amtlicher Daten erheblich verbessert haben, kann für die Bildungsforschung nach wie vor eine gewisse „Distanz zur Bildungsstatistik“ (Weishaupt, 2019, o. S.) konstatiert werden. Intensiv genutzt werden Bildungsstatistiken bislang vor allem in regelmäßigen Berichten sowie Einzelstudien im weiteren Kontext des Bildungsmonitorings. Je nach Akteur:innen- bzw. Netzwerkkonstellation folgen solche Monitoringinstrumente jedoch oftmals keiner rein wissenschaftlichen Forschungsagenda, sondern einer eigenen Programmatik. Wissenschaftler:innen treten in diesen Handlungskontexten im Vergleich zu staatlichen (Behörden, Ministerien etc.) oder privaten Akteur:innen (z. B. Stiftungen, Unternehmen, Nichtregierungsorganisationen) selten als Impulsgebende in Erscheinung, die vom Forschungsstand geleitete Erkenntnisinteressen verfolgen oder wissenschaftlich elaborierte Problemanalysen und -kontextualisierungen durchführen. Eine intensivierte Nutzung amtlicher Daten aus dem Wissenschaftssystem heraus würde zum einen zeigen, welche Erkenntnispotenziale mit Sekundäranalysen amtlicher Statistiken verbunden sind, und zum anderen auch die noch bestehenden Defizite transparenter machen. Dies könnte weiterführende Impulse dafür geben, verbesserte Dateninfrastrukturen für die Forschung zu implementieren und dazu beitragen, die Deutungshoheit über bildungsstatistische Problembeschreibungen im Bildungssystem nicht allein Akteur:innen zu überlassen, die fachlich oder politisch eine eigene Agenda verfolgen. Neben der empirischen pädagogischen Forschung ist aber auch die Allgemeine Erziehungswissenschaft gefordert, die fortschreitende Digitalisierung und Datafizierung hinsichtlich der Folgen für die Steuerung von Bildungsprozessen und die Konstruktion sozialer Wirklichkeit zu reflektieren. Mit den neuen Dateninfrastrukturen und Mechanismen der Datenproduktion und -auswertung entstand etwa der Forschungsbereich sogenannter Critical Data Studies, die „die Entstehung und Veränderung von Governanceräumen,

Regelungsstrukturen und damit einhergehenden (wahrgenommenen) Handlungsoptionen durch digitale Daten selbst ins Zentrum der Betrachtung rücken“ (Hartong & Fischer, 2020, S. 420). Je nach Merkmalskranz sind z. B. elektronisch gespeicherte Informationen über einzelne Schüler:innen, Lehrkräfte und ihren Unterricht hoch sensible Daten, deren (statistische) Nutzung klar von anderen Zweckbestimmungen abgegrenzt sein muss, die dem Datenschutz oder der informationellen Selbstbestimmung zuwider laufen könnten. Bildungspolitik und -administration stehen hier vor der Herausforderung, entsprechende rechtliche, organisatorische und fachliche Voraussetzungen kontinuierlich dem technischen Fortschritt anzupassen, in geeignete Verfahren der Datenübermittlung, -speicherung und Datenweitergabe zu überführen und nicht zuletzt die Angemessenheit und Qualität von Hard- und Softwarelösungen für die laufende Pflege und Betreuung der Dateninfrastrukturen zu prüfen, sicherzustellen und transparent zu machen.

Welche Daten verfügbar und auf welchem Wege sie zugänglich sind, ist heute weitaus schwerer zu überblicken als noch vor 20 oder 30 Jahren. Früher beschränkte sich das amtliche Datenangebot auf hochaggregierte Informationen in gedruckten Veröffentlichungen der Ämter und anderer Statistikstellen. Mit der fortschreitenden Digitalisierung können die Daten inzwischen in differenzierterer Aufbereitung und auf vielfältigen Wegen erschlossen werden (vgl. Tab. 1). Sie bieten so für Forschende wissenschaftlich unabhängiger Einrichtungen das Potenzial, neue Forschungsfragen mit geringem Ressourcenaufwand zu untersuchen. Je nachdem, auf welche Detailtiefe und Flexibilität die Wissenschaftler:innen angewiesen sind, können mit den unterschiedlichen Zugangswegen kleinere oder größere Hürden verbunden sein (Antragsverfahren, Genehmigung, Kosten, Geheimhaltungspflichten etc.). Die größten Herausforderungen stellen sich Wissenschaftler:innen bei der Nutzung von Schulstatistiken, deren Verfügbarkeit aufgrund der landesspezifischen Handlungsrationale und -routinen in der Datengenerierung und Datenbereitstellung wesentlich schwerer überblickbar ist als in den übrigen, bundesgesetzlich geregelten Bildungsbereichen. Hier sind vor allem Politik und Administration gefordert, einheitlichere Festlegungen zu treffen, wie es zuletzt in einer Ländervereinbarung seitens der KMK 2020 erneut angekündigt wurde. Dafür müssten allerdings politische Vorbehalte überwunden werden, die eine Abstimmung in Zuordnungs- und Definitionsfragen mit unerwünschten Tendenzen einer bildungspolitischen Vereinheitlichung zwischen den Ländern assoziieren.

Bislang unterscheiden sich trotz bestehender Absprachen und Beschlüsse der KMK viele Erhebungskonzepte und -konstrukte erheblich; für länderübergreifende oder -vergleichende Analysen ist damit ein großer Teil der Schulstatistik nicht verfügbar. Beispielsweise erfassen Länder Merkmale eines Migrationshintergrunds uneinheitlich, sodass auf Bundesebene bis heute lediglich Angaben zur Staatsangehörigkeit der Schüler:innen vergleichend ausgewertet werden können. Differenziertere schulstatistische Analysen in Ländern, die einen breiteren Merkmalskranz (z. B. Geburtsort oder Familiensprache) erheben, können den Mehrwert solcher Daten für Fragen der Bildungsteilnahme verdeutlichen und damit Impulse für weitere Harmonisierungen geben. Gleiches gilt auch für Fragen der Erreichbarkeit von Bildungsangeboten. Nur in wenigen

	Veröffentlichungen	Bildungsdatenbanken	Sonderauswertungen	Mikrodatenzugriff
<b>Zugang</b>	Internetauftritte der öffentlichen Datenhalter	Öffentlich zugängliche Internetdatenbanken	Statistische Ämter; Sekretariat der KMK; Ministerien; Forschungsinstitute (z. B. BIBB, IAB)	Forschungsdatenzentren; Statistikstellen (z. B. Sekretariat der KMK oder Ministerien)
<b>Modalitäten</b>	Freier Datenzugang	Freier, oft interaktiver Datenzugang	Zustimmungs- und ggf. kostenpflichtige Aufbereitung durch Datenhalter	Kostenpflichtige, befristete und zweckgebundene Datennutzung auf Antrag
<b>Auswertungsoptionen</b>	Deskriptive Analysen mit Aggregatdaten in hoch standardisierten Tabellen	Deskriptive Analysen von Aggregatdaten nach Auswahl vorab definierter Merkmalsausprägungen	Deskriptive (ggf. auch multivariate) Analysen als externe Aufbereitung der Mikrodaten nach eigenen Vorgaben	Explorative deskriptive oder multivariate Analysen durch eigene Auswertung (anonymisierter) Mikrodaten

Quelle: Eigene Darstellung

Tab. 1: Datentypen und -zugänge der amtlichen Statistik im Überblick

Ländern liegen sowohl Angaben zum Wohn- als auch zum Schulstandort vor. Entsprechende kleinräumige Untersuchungen etwa zu Pendelbewegungen im Falle sonderpädagogischer Förderung könnten auch für die übrigen Länder den wissenschaftlichen wie auch politisch-administrativen Erkenntnisgewinn etwa für eine inklusive Schulentwicklungsplanung sichtbar machen.

Mit der Datenverfügbarkeit eng verknüpft sind auch Defizite in der Datenbereitstellung. Werden Merkmale uneinheitlich (oder gar nicht) erfasst, kann jeweils nur der kleinste gemeinsame Nenner veröffentlicht werden. Neben dem Beispiel des Migrationshintergrunds betrifft dies unter anderem auch die Schulstrukturen. Die traditionelle Schulartgliederung in den bundesweiten Veröffentlichungen spiegelt schon lange nicht mehr die differenzierten Schullaufbahnoptionen in den einzelnen Ländern wider. So werden auf Bundesebene einerseits Schultypen zusammengefasst, die sich in ihrer Organisationsform deutlich unterscheiden (z. B. unterschiedlich ausgestaltete Gemeinschaftsschulen und Gesamtschulen). Andererseits werden aber umgekehrt sehr ähnliche Schultypen unterschiedlichen Schularten zugeordnet (z. B. Schulen, die sowohl zum Hauptschul- als auch zum mittleren Schulabschluss führen). Wünschenswert wäre hier nicht trotz, sondern wegen der schulstrukturellen Eigenheiten in den Ländern eine konsistentere Datenbereitstellung, die die institutionelle Schulartgliederung mit der individuellen Zuordnung zu den Bildungsgängen (Hauptschul-, Realschul-, Gymnasialbildungsgang und Integrierte Form) verknüpft und damit weiterhin den landesspezifischen System- und Handlungslogiken Rechnung tragen würde.

Die Voraussetzungen für derartige Anpassungen der Schulstatistik wären günstiger, wenn alle Länder den bereits 2003 gefassten KMK-Beschluss zur Einführung schulstatistischer Individualdaten umgesetzt hätten. Mit Einzeldatensätzen sind Verknüpfungen zwischen Angaben über die Einrichtung und über die Schüler:innen nicht nur einfacher,

sondern auch flexibel handhabbar. Solange in einigen Ländern noch zusammengefasste Summendaten an die Meldestellen übermittelt werden, sind jeder Datenauswertung entsprechende Grenzen gesetzt – auch der Mikrodatenzugang über die Forschungsdatenzentren geht dann nicht über die Schul- oder ggf. Klassenebene hinaus. Die wenigen Forschungsergebnisse mit Schulstatistiken jener Länder, die bereits weitgehend den Kerndatensatz der KMK eingeführt und teilweise auch mit einer Personenkennung verbunden haben, verdeutlichen eindrücklich den damit verbundenen Mehrwert für die Erforschung und die Steuerung des Systems (vgl. Böttcher & Kühne, 2017). Die flächendeckende Umsetzung des Kerndatensatzes wird umso wahrscheinlicher, je häufiger die bereits verfügbaren schulstatistischen Individualdaten von Wissenschaftler:innen nachgefragt und ihnen dann auch für Forschungszwecke zur Verfügung gestellt werden. Entscheidend für den Auf- und Ausbau dieser Dateninfrastrukturen bleibt nichtsdestotrotz der politische Wille, dem weniger Datenschutzbedenken als die Sorge entgegenstehen dürfte, infolge erweiterter Nutzungsmöglichkeiten mit immer neuen, auch unerfreulichen Forschungsbefunden konfrontiert zu werden, die das politische Handeln unter erhöhten Legitimationsdruck setzen (können). Beschleunigt werden könnte die politische Willensbildung durch die neueren Bemühungen zur Einführung eines Bildungsregisters (vgl. Gawronski, 2020). 2019 haben sowohl der Bundesrat als auch die Konferenz der Ministerpräsident:innen beschlossen, den Aufbau eines solchen Registers zumindest zu prüfen, welches unter Einführung von datenschutzkonformen Identifikationsnummern bildungsstatistische Informationen über die Bildungsbereiche hinweg zusammenführen soll. Ohne schulstatistische Individualdaten wird dies jedoch kaum gelingen können.

## Literatur

- Ackeren, I. van, & Hovestadt, G. (2003). *Indikatorisierung der Empfehlungen des Forum Bildung. Ein exemplarischer Versuch unter Berücksichtigung der bildungsbezogenen Indikatorenforschung und -entwicklung*. Bonn: BMBF.
- Artelt, C., Bug, M., Kleinert, C., Maaz, K., & Runge, T. (2019). Nutzungspotenziale amtlicher Statistik in der Bildungsforschung. Ein Überblick zu Erreichtem, möglichen Chancen und anstehenden Herausforderungen. In D. Fickermann & H. Weishaupt (Hrsg.), *Bildungsforschung mit Daten der amtlichen Statistik* (S. 21–37). Münster/New York: Waxmann.
- Autorengruppe Bildungsberichterstattung (2018). *Bildung in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Wirkungen und Erträgen von Bildung*. Bielefeld: wbv.
- Autorengruppe Bildungsberichterstattung (2020). *Bildung in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt*. Bielefeld: wbv.
- Carnap, R. von, & Edding, F. (1962). *Der relative Schulbesuch in den Ländern der Bundesrepublik 1952–1960*. Frankfurt a. M.: Hochschule für Internationale Pädagogik.
- Böttcher, W., & Kühne, S. (2017). *Schulstatistische Individualdaten zur Rekonstruktion von Bildungsverläufen: Perspektiven für die Weiterentwicklung des Sozial- und Bildungsmonitorings in Nordrhein-Westfalen* (FGW-Studie Vorbeugende Sozialpolitik 03). Düsseldorf: FGW – Forschungsinstitut für gesellschaftliche Weiterentwicklung.
- Edele, A., Jansen, M., Schachner, M. K., Schotte, K., Rjosk, C., & Radmann, S. (2020). School track and ethnic classroom composition relate to the mainstream identity of adolescents with immigrant background in Germany, but not their ethnic identity. *International Journal of Psychology*, 55(5), 754–768. <https://doi.org/10.1002/ijop.12677>.

- Forster, E. (2014). Kritik der Evidenz. Das Beispiel evidence-informed policy research der OECD. *Zeitschrift für Pädagogik*, 60(6), 890–907.
- Gawronski, K. (2020). Konzeption eines Bildungsregisters in Deutschland, WISTA – Wirtschaft und Statistik. *Statistisches Bundesamt (Destatis)*, 72(2), 37–45.
- Hartong, S., & Förtschler, A. (2020). Dateninfrastrukturen als zunehmend machtvolle Komponente von Educational Governance. Eine Studie zur Implementierung und Transformation staatlicher Bildungsmonitoringsysteme in Deutschland und den USA. In I. van Ackeren, H. Bremer, F. Kessl, H. C. Koller, N. Pfaff, C. Rotter, D. Klein & U. Salaschek (Hrsg.), *Bewegungen. Beiträge zum 26. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft* (S. 419–432). Opladen/Berlin/Toronto: Verlag Barbara Budrich.
- KMK (2020). *Politische Vorhaben zur „Ländervereinbarung über die gemeinsame Grundstruktur des Schulwesens und die gesamtstaatliche Verantwortung der Länder in zentralen bildungspolitischen Fragen*. Beschluss vom 15.10.2020. [https://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2020/2020\\_10\\_15-Politische-Vorhaben-LV.pdf](https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2020/2020_10_15-Politische-Vorhaben-LV.pdf) [10.06.2021]
- Köhler, H. (1980). Amtliche Bildungsstatistik im Wandel. In J. Baumert, A. Leschinsky & J. Naumann (Hrsg.), *Bildung in der Bundesrepublik Deutschland. Daten und Analysen. Band 2: Gegenwärtige Probleme* (S. 1215–1285). Stuttgart: Klett-Cotta.
- Kühne, S. (2015). *Zur Rekonstruktion schulischer Bildungsverläufe. Der Beitrag der Individualstatistik für die Entwicklung von Verlaufsindikatoren*. Münster: Waxmann.
- Kühne, S., Maaz, K., Mank, S., Ordemann, J., & Schulz, S. M. (2019). *Bildung in Sachsen im Spiegel der Nationalen Bildungsberichterstattung*. Frankfurt a. M./Berlin: DIPF.
- Lüttinger, P., & Schimpl-Neimanns, B. (1992). *Amtliche Bildungsstatistik und empirische Sozialforschung* (ZUMA-Arbeitsbericht Nr. 92/19). Mannheim: Zentrum für Umfragen, Methoden und Analysen (ZUMA).
- Makles, A. M., Schneider, K., & Terlinden, B. (2019). Schulische Segregation und Schulwahl. Eine Analyse mit den Daten der Bremer Schülerindividualstatistik. In D. Fickermann & H. Weishaupt (Hrsg.), *Bildungsforschung mit Daten der amtlichen Statistik* (S. 176–196). Münster/New York: Waxmann.
- Mundelius, M. (2019). Der Kerndatensatz auf der Basis von Individualdatenerhebungen in der Schulstatistik. Von Summendaten zu Einzeldaten. In D. Fickermann & H. Weishaupt (Hrsg.), *Bildungsforschung mit Daten der amtlichen Statistik* (S. 38–45). Münster/New York: Waxmann.
- Picht, G. (1965). *Die deutsche Bildungskatastrophe*. München: Deutscher Taschenbuch Verlag.
- Rjosk, C., Lütke, O., Richter, D., & Eccles, J. S. (2017). Ethnic composition and heterogeneity in the classroom: Their measurement and relationship with student outcomes. *Journal of Educational Psychology*, 109(8), 1188–1204. <https://doi.org/10.1037/edu0000185>.
- Weishaupt, H. (2019). Wie Bildungsstatistiken für die Forschung nutzbar gemacht werden können: Praktische Hinweise zum Umgang mit dem Datenangebot der Schulstatistik. In *bildungs-serverBlog*. Frankfurt a. M.: DIPF. <https://blog.bildungsserver.de/?p=6597> [22.06.2021].
- Williamson, B. (2016). Digital education governance: an introduction. *European Educational Research Journal*, 15(1), 3–13.
- Zühlke, S., & Christians, H. (2005). *Datenangebot und Datenzugang im Forschungsdatenzentrum der Statistischen Landesämter*. Beiträge zu den Nutzerkonferenzen des FDZ der Statistischen Landesämter 2005.

**Abstract:** In contrast to survey-based primary data, official statistics, while comprehensive, often offer fewer variables and data linkage opportunities. However, their advantage lies in the fact that they provide data material that is available both nationwide and in long time series. Official statistics may not only serve as a reference point for scientific investigations (e. g., sampling) but can also be used for independent research questions and analyses for which no alternative data are available (e. g., for smaller populations or disparities at local levels). These statistics are also increasingly available in a disaggregated format for individual units, such as students or teachers, and can thus be used not only for structural or time-series analyses but also for microdata analyses. This article presents in broad terms the range of available and accessible official education statistics. Selected research results are used as examples to illustrate the added value of these government statistics from the perspective of education monitoring in Germany.

**Keywords:** Education Statistics, School Statistics, Official Statistics, Research Data, Educational Monitoring

#### **Anschrift der Autor:innen**

Dr. Stefan Kühne, DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation,  
Warschauer Straße 34–38, 10243 Berlin, Deutschland  
E-Mail: [kuehne@dipf.de](mailto:kuehne@dipf.de)

Svenja Mank, M.A., DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation,  
Warschauer Straße 34–38, 10243 Berlin, Deutschland  
E-Mail: [mank@dipf.de](mailto:mank@dipf.de)

Stefan Schulz, M.A., DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation,  
Warschauer Straße 34–38, 10243 Berlin, Deutschland  
E-Mail: [Stefan.Schulz@dipf.de](mailto:Stefan.Schulz@dipf.de)

Prof. Dr. Kai Maaz, DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation,  
Rostocker Straße 6, 60323 Frankfurt a. M., Deutschland  
E-Mail: [maaz@dipf.de](mailto:maaz@dipf.de)