

Annina Förschler & Sigrid Hartong

Datenpraktiken des Schulmonitorings in staatlichen Bildungsbehörden Beobachtungen jenseits des Regelkreises

Zusammenfassung

Der vorliegende Beitrag adressiert die Frage, inwieweit bestimmte Steuerungsvorstellungen, -modelle und -praktiken die zunehmende Datafizierung und Digitalisierung im Bildungssektor auf den verschiedenen Steuerungsebenen prägen, und thematisiert hierbei das Beispiel von Schulmonitoringsystemen in Schulbehörden auf Bundesland- bzw. Bundesstaatenebene. Es werden Ergebnisse eines laufenden DFG-Projektes vorgestellt, das anhand von Fallstudien in Deutschland und den USA die zunehmende Expansion und Zentralisierung digitaler Schulmonitoringsysteme (z.B. durch die Einrichtung zentraler Schulverwaltungssoftware, die Zusammenführung unterschiedlicher Leistungsdaten usw.) untersucht, um mehr über das „Wie“ der jeweiligen Datenproduktion, -verarbeitung und -nutzung zu erfahren. Wie die Ergebnisse zeigen, liegen der technischen Vorstellung und auch der Außendarstellung von Schulmonitoringsystemen zwar oftmals simplifizierte Regelkreismodelle zugrunde; in der Praxis handelt es sich jedoch um ein hochkomplexes Zusammenspiel unterschiedlicher, oftmals ambivalenter Logiken und Einflüsse, die wiederum jeweils spezifische Potenziale und Problematiken mit sich bringen. Während der Beitrag dies anhand von ausgewählten Beispielen diskutiert, wird gleichzeitig sichtbar, wo und wie sich Gemeinsamkeiten und Unterschiede zwischen Deutschland und den USA manifestieren.

Schlüsselwörter: Datafizierung, Digitalisierung, Schulbehörden, Monitoring, Datenpraktiken, Critical Data Studies

Data Practices in State Agencies' School Monitoring

Observations beyond the Loop Circle Model

Abstract

Addressing the question of how the rising datafication and digitalization of education comes along with specific ideas, models and practices of governance, this contribution discusses the ongoing transformation of school monitoring systems in state-level education agencies. The presented findings result from a DFG research project which analyzes the expansion and gradual centralization of school monitoring systems in Germany and the US, and which seeks to better understand the various data practices behind the actual "doing" of monitoring. As our findings illustrate, while loop circle models may frame the technical idea of school monitoring, the actual doing of monitoring appears as a complex entanglement of very different, ambivalent logics and problems. We explain selected ambivalences in more detail, also pointing to similarities and differences between the two countries under study.

Keywords: datafication, digitalization, state education agencies, school monitoring, data practices, critical data studies

1. Datenbasierte Regelkreise versus *Critical Data Studies*: Verortung der Studie

Mit dem bildungspolitischen *turn* zur evidenzbasierten Bildungssteuerung hat das Monitoring von Schulen über quantifizierbare und zunehmend digitale Daten nicht nur in Deutschland massiv an Bedeutung gewonnen (siehe z.B. Lawn, 2013; Böttcher, Bos, Döbert & Holtappels, 2008; Bellmann, 2015; Bormann, Hartong & Höhne, 2018). Entsprechend expandiert sind in den letzten Jahren neue Systeme bzw. Infrastrukturen (im Folgenden *Dateninfrastrukturen*) zur effektive(re)n Produktion, Sammlung, Verarbeitung, Aufbereitung und Nutzung dieser Bildungsdaten in Bildungspolitik, -verwaltung und Schulpraxis (Williamson, 2017; Hartong, Breiter, Jarke & Förschler, 2020; Breiter & Lange, 2019) sowie neue Technologien, die eine zunehmende Dateninteroperabilität¹ und Beschleunigung von Datenflüssen versprechen. Auf Bundeslandebene der Schulverwaltung haben diese wachsende Datafizierung und

1 Unter Dateninteroperabilität wird das zunehmende Standardisieren von Daten(formaten) verstanden, um den Datenaustausch sowie einfache System- und Datenverknüpfungen zwischen einzelnen Teilsystemen (z.B. Teilkomponenten von Schulinformationssystemen auf unterschiedlichen Ebenen; siehe Hartong et al., 2019, S. 6) möglichst reibungslos zu gestalten (vgl. auch European Interoperability Framework (EIF), https://ec.europa.eu/isa2/eif_en). Interoperabilitätslösungen sollen zudem einen schnellen Austausch komplexer Informationen in großen Mengen auch zwischen unterschiedlichen Systemen (z.B. Bildungsstatistik, Meldewesen etc.) ermöglichen und „beinhalten stets menschliche Operateure, Informationssysteme, eine Infrastruktur zum Austausch von Information und schließlich auszutauschende Information“ (Fraunhofer FKIE, 2018).

Digitalisierung u. a. dazu geführt, dass sich die kontinuierliche Produktion sowie das Management von Daten neben Schulaufsicht und -beratung zu zentralen Aufgaben von Schulbehörden entwickelt haben,² die die zunehmend eigenständigen Schulen auf Grundlage dieser Daten bedarfsgerecht unterstützen, aber auch kontrollieren sollen.

In Deutschland haben die meisten Bundesländer zur Bewältigung dieser Aufgabe in den vergangenen Jahren – mit Unterstützung von IT-Dienstleistern – entsprechende institutionelle Strukturen ausgebaut bzw. dezidierte Qualitäts-, Monitoring- bzw. auch Beratungsinstitute implementiert.³ Gleichzeitig lässt sich hierzulande ein eindeutiger Trend zur Datenzentralisierung identifizieren, etwa durch die Nutzung von *Data Warehouse*-Lösungen oder die Etablierung bundeslandeinheitlicher Schulverwaltungssysteme, die die Arbeit an und mit Schuldaten (sowohl Leistungs- als auch statistischen Daten), aber z. B. auch die Zusammenführung mit anderweitigen Verwaltungsdaten (z. B. Meldewesen, Sozialwesen) effizienter gestalten sollen (siehe auch Hartong et al., 2019; Breiter & Lange, 2019).⁴

Wie in den letzten Jahren umfangreich gezeigt wurde (siehe etwa die Beiträge in Bormann et al., 2018, oder auch Burchardt, 2017), hat die Durchsetzung datenbasierter Bildungssteuerung dabei auch zu einer neuen Prominenz von Regelkreismodellen geführt, die Daten oftmals als objektive Fix- bzw. Anhaltspunkte innerhalb von Schulentwicklung und -aufsicht verstehen (siehe Beispiel Abb. 1 auf der folgenden Seite). Derartige datenpositivistische Regelkreise laufen damit aber auch Gefahr – so die durchaus berechtigte Kritik (z. B. Bellmann, 2015; Rürup, 2018; Heinrich, 2018) –, dass die konstruktivistische und damit politische Seite von Daten(-produktion/-vermittlung/-rezeption), die in den von uns beforschten Bildungsbehörden durchaus kritisch wahrgenommen und thematisiert wurde (u. a. USA 15/17/18, DEU01/02/06/26/31/32/37), tendenziell ausgeblendet wird (vgl. Karcher, 2018, S. 85; siehe auch Maritzen & Tränkmann, 2015).

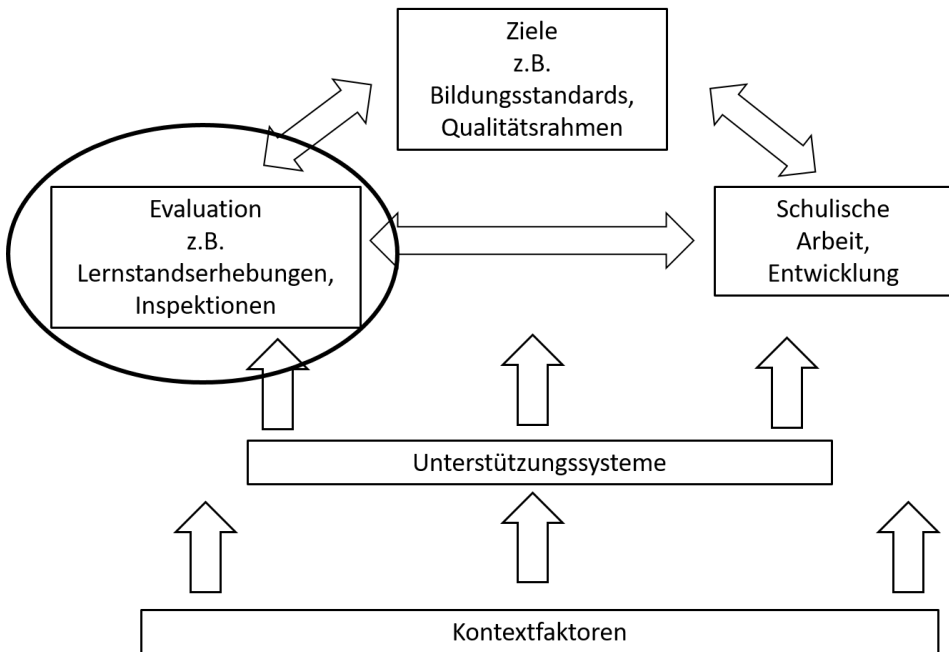
Der international sukzessive an Bedeutung gewinnende Forschungszweig sogenannter *Critical Data Studies* lässt sich als wichtige Antwort auf diese Problematik verstehen, indem sich hier vielmehr gezielt mit der Nicht-Neutralität, der Dynamik, aber auch der Macht von Strukturen beschäftigt wird, in denen Daten produziert,

2 Dies reflektieren auch unsere Studienergebnisse, etwa in den Interviews USA15/16/19 sowie DEU02/05/06. Für einen Überblick siehe Hartong & Förschler, 2019b.

3 Für einen aktuellen Überblick der Landesinstitute und Qualitätseinrichtungen der Länder siehe https://bildung-rp.de/fileadmin/user_upload/bildung-rp.de/pl/Newsletter/2018/8/12/Positionspapier_Transfer_31.10.18.pdf. Für einen kritischen Überblick siehe Rürup, 2014.

4 Trotz dieser Bemühungen berichteten uns einige Interviewpartner*innen von nach wie vor bestehenden Herausforderungen (sowohl für die Behörden/Institutionen als auch für die Schulen), mit der wachsenden Menge von Daten konstruktiv umgehen bzw. diese sinnvoll vermitteln oder nutzen zu können (u. a. DEU02/04/06/31, USA15/16/18).

Abb. 1: Beispiel eines Regelkreises datenbasierter Schulentwicklung



Quelle: Altrichter, Moosbrugger & Zuber (2016), S. 237, modifiziert

selektiert, modelliert und verarbeitet werden (Allert & Richter, 2016; Kitchin & Lauriault, 2014; Iliadis & Russo, 2016; Selwyn, 2015; Williamson, 2017; Ames, 2018). Wenngleich auch die deutschsprachige *Educational-Governance*-Forschung die zentrale Bedeutung komplexer, oftmals spannungsreicher Interdependenzbewältigung einer Vielzahl von Akteuren und Logiken im Zuge neuer Steuerung betont (z. B. Abs, Brüsemeister, Schemmann & Wissinger, 2015), so hat sie im Falle datenbasierter Schulsteuerung jedoch bislang vor allem auf Problematiken der Umsetzung und damit auf das Ziel eines besseren Reformgelingens abgestellt (z. B. *DDS – Die Deutsche Schule*, 2018, Heft 1; siehe hierzu auch Niedlich & Bormann, 2019; Hartong, 2019b). In diesem Sinne bieten die Ansätze der *Critical Data Studies* für derartige Forschung ein vielversprechendes Weiterentwicklungspotenzial.

Das laufende DFG-Forschungsprojekt *Bildungsdatenmanagement – Neue Wissens-, Interpendenz- und Einflussstrukturen im Kontext der Digitalisierung von Educational Governance* an der Helmut-Schmidt-Universität Hamburg ist in diesen Kontext einzuordnen, indem es darauf abzielt, ein besseres Verständnis für die konkrete Produktion und damit auch für die Komplexität des Monitorings – im Sinne von „data assemblages at work“ (Kitchin & Lauriault, 2014) – in Schulbehörden auf Bundeslandebene zu gewinnen. Gleichzeitig wird der Frage nachgegangen, inwieweit sich lokale bzw. kul-

turell-institutionelle Kontexte in den Praktiken der Datenproduktion manifestieren, konkret im Vergleich zwischen Deutschland und den USA.⁵

2. Jenseits des Regelkreises: Einblicke in Datenpraktiken des Schulmonitorings auf Bundesland-/Bundesstaatenebene

2.1 Methodische Vorbemerkungen

Im Rahmen der Studie wurde zunächst untersucht, wie sich Dateninfrastrukturen des Schulmonitorings und damit die Organisation der staatlichen Schulverwaltung in den letzten zwanzig Jahren verändert haben und wie sich der aktuelle Entwicklungsstand bezüglich Monitoringsystemen auf Bundesland-/Bundesstaatenebene in Deutschland und den USA darstellt. Hierbei wurden, neben Interviews mit politischen sowie länderübergreifend agierenden Akteuren (wie etwa der Kultusministerkonferenz in Deutschland oder der *Data Quality Campaign* in den USA) politische Initiativen und Gesetze analysiert, um Schlüsselveranstaltungen, zentrale Akteurskonstellationen sowie schließlich Gemeinsamkeiten und Unterschiede sowohl zwischen Deutschland und den USA als auch zwischen Bundesländern/-staaten in den jeweiligen Nationen herauszuarbeiten (siehe hierzu Hartong, 2019b; Hartong et al., 2019).

Eingebettet in die Ergebnisse dieser ersten Phase wurden im zweiten Teil der Studie Schulmonitoringsysteme in einzelnen Bundesländern/-staaten detaillierter untersucht. Hierbei ging es zunächst um eine Sichtung der Organisationsstruktur, der zentralen Akteur*innen, der produzierten Datentypen/-flüsse und -schnittstellen, die zusammen das Bildungsmonitoring „hervorbringen“. Hierzu zählen statistische Erhebungen und die Sammlung von Verwaltungsdaten ebenso wie die Produktion und Nutzung von Leistungsdaten (etwa für die Steuerung behördlicher Interventionen).

Methodisch basieren die Falluntersuchungen zum einen auf der kategoriengeleiteten Analyse einer Dokumentensammlung, die Organigramme, Tätigkeitsberichte, Beschreibungen von und Anleitungen zur Nutzung einzelner Monitoringinstrumente bzw. digitaler Datenplattformen, Produktportfolios von Softwareherstellern oder Pressematerial für jedes Bundesland bzw. jeden Bundestaat umfasste. Auf der anderen Seite wurden 39 leitfadengestützte, narrative Interviews mit Mitarbeiter*innen der Schulbehörden, aber auch von Qualitätsinstituten sowie IT-Organisationen in zwei deutschen und zwei amerikanischen Bundesländern/-staaten geführt. Hierbei stand

5 Deutschland und die USA markieren in Sachen Datafizierung und Digitalisierung dabei sowohl ähnliche (massiver Ausbau von Dateninfrastrukturen in den letzten 20 Jahren, föderale Strukturen mit starker Länderautonomie) als auch gegensätzliche (lange Tradition datenbasierter Steuerungsideen in den USA versus radikalere Bildungsreformen in der jüngeren Vergangenheit in Deutschland) Fälle. Für die Begründung der Fallauswahl siehe auch Hartong, 2019a, sowie Hartong & Förschler, 2019, 2020.

neben einer weiteren Rekonstruktion von Dateninfrastrukturen und -flüssen des Schulmonitorings die Befragung der Mitarbeitenden zu ihrer konkreten *Datenarbeit* im Zentrum. So wurde zum Beispiel gefragt, wie Indikatoren von Datenerhebungen generiert werden, wie Modellierungsprozesse von Daten (etwa zur Erstellung steuerungsrelevanter Kennzahlen oder Indizes) erfolgen oder welche Daten für welche Formen der Entscheidungsfindung wie einbezogen werden.

Beide Projektphasen wurden durch einen kontinuierlichen Prozess des fallinternen und -übergreifenden Vergleichens begleitet, sowohl der zwei analysierten Nationen als auch der Bundesländer/-staaten, einzelner Datenpraktiken oder Akteurskonstellationen. Hierdurch wurde angestrebt, auf der einen Seite möglichst tiefe Einblicke in unterschiedliche Fälle und Kontextspezifitäten und damit auch in die Komplexität der betrachteten Phänomene zu erlangen, auf der anderen Seite aber auch ein gewisses Maß an Generalisierbarkeit zu ermöglichen.

Ein zentrales, fallübergreifendes Ergebnis war hierbei ein klares Bewusstsein der von uns interviewten Behördenmitarbeitenden für die enorme Komplexität und Dynamik von Schulmonitoring, das auf dem Papier einer klaren Logik folgen mag – Daten werden gesammelt, ausgewertet, berichtet und für Steuerungszwecke genutzt –, in der Praxis jedoch ein selbst von den Mitarbeitenden kaum zu durchdringendes Gemengelage von Datenflüssen, -prozessen und -strukturen darstellt (siehe auch Abb. 2 in Abschnitt 3). Mehrere Aspekte wurden hierbei von den Interviewten herausgestellt: Zum einen umfassen Datenpraktiken des Bildungsmonitorings, also etwa das Produzieren neuer Erhebungsinstrumente und Analysemodelle oder das Validieren von Daten, stets ein hohes Maß an Experimentieren, das auch als „*Kochen mit diversen Zutaten*“ (USA18), als „*Umbiegen*“ (DEU08), als „*irgendwie hinbekommen*“ (DEU08) oder als „*messing around*“ (USA15) beschrieben wurde. Hierbei spielt ebenfalls eine Rolle, dass regelmäßig eine Vielzahl durchaus auch unvorhersehbarer Einflüsse zum Tragen kommt, etwa politische Positionswechsel, Ressourcenprobleme, persönliche Vorlieben, aber durchaus auch eigenlogische Dynamiken der Dateninstrumente und Algorithmen selbst. Mit anderen Worten zeigte sich hier, was Rob Kitchin, Tracey P. Lauriault und Gavin McArdle (2015, S. 16–18) als „[...] [data being a] product of many minds working within diverse situations, framed and shaped within contexts and structures“ beschreiben.

Zum anderen wurde in allen untersuchten Behörden ein hochdynamisches Nebeneinander unterschiedlichster, in ihrer jeweiligen Komplexität kontinuierlich expandierender Dateninstrumente sichtbar, mit *je eigenen Regelkreisen*, Zeitfenstern (etwa tägliche Datenerhebungen über die Anwesenheit von Schüler*innen versus statistische Jahreserhebungen oder Tests), Abläufen (etwa Organisation der Datenerhebung) oder Datenpraktiken (etwa Formatierung, Modellierung oder Visualisierung). In der Tat haben die Behörden in beiden Nationen angesichts dieser Problematiken in den letzten Jahren zunehmend eine datenökonomische Strategie implementiert, die

nicht nur darauf abzielt, unnötige oder doppelt erhobene Daten zu eliminieren, sondern gleichzeitig das Datenmanagement durch Standardisierung, Automatisierung und Zentralisierung zu optimieren. Dennoch stellt die schiere Masse an Daten für die Behörden eine wachsende Herausforderung dar und damit auch die Frage, welche Daten(auswahl) überhaupt und in welcher Form an die Schulen weitergegeben werden sollen, um diese nicht mit Daten „zuzuschütten“ (DEU06).

Im Kontext dieser allgemeinen Befunde möchten wir im Folgenden auf drei ausgewählte Ambivalenzen eingehen, die im Rahmen der Interviews fallübergreifend immer wieder und zu weiten Teilen auf ähnliche Weise problematisiert wurden (für einen umfangreicheren Überblick siehe Hartong & Förschler, 2019).

2.2 Datenvereinfachung versus Datengenauigkeit oder die Problematik der Unter-/Überkomplexität

Ein zentrales Ziel staatlicher Bildungseinrichtungen ist es, Schulen, Lehrkräften, Eltern oder der breiten Öffentlichkeit Anreize und Impulse in Richtung einer häufigeren Nutzung von (ihren) Daten und einer verbesserten datengestützten Kommunikation zu geben. Gleichzeitig problematisierten einige unserer Interviewpartner*innen⁶, dass vielen ihrer Adressat*innen die Zeit, das Fachwissen oder die Motivation fehle, um sich in die (steigende Menge und Komplexität der) Daten *richtig* einzuarbeiten und diese reflektiert zu nutzen. Entsprechend wichtig ist allen Behörden, die Daten möglichst *gebrauchsfertig* aufzuarbeiten und mit einer klaren Aussage bzw. „Geschichte“ (USA19) zu versehen, um die Nutzungszahl auch von Nichtfachkundigen zu erhöhen:

What we're trying to do right now is expand our outreach because we know that there's a huge opportunity for parents and kids, other audiences to use this data but they're not going to have as much experience with data, they are not going to be the ones to download it and put it into Tableau and run analytical reports to figure out which school has the best support program. [...] [W]e are trying to [...] really work [...] on data visualization and doing more actionable data with less interpretation of the data. So that we do it so that parents, we can reach that audience that is not data experts. (USA19)

Wir fanden verschiedene Beispiele für solch explizit vereinfachte bzw. verdichtete Dateninstrumente, darunter unter anderem handliche Schuldatenblätter oder Online-Datendashboards mit Schulprofilen, sowie eine wachsende Verwendung von Karten, Ampelsystemen oder bereits aufbereiteten Präsentationsformaten:

6 Bei der konkreten Nennung unserer Interviewpartner*innen verwenden wir im Folgenden zwecks Leserlichkeit sowie aus Anonymisierungsgründen nur die männliche Form.

[Da] macht es halt viel mehr Sinn, die Daten so zusammenzustellen, dass man damit nicht mehr so viel Aufwand selbst hat, sondern man kann das einfach so zack, rausziehen, zeigen [...], also es wurde so gemacht, dass die Karten, [...] die können halt in dem Format gedruckt werden, die dann auch für Präsentationen gebraucht werden können. (DEU05)

Die Kehrseite derartiger benutzerfreundlicher Vereinfachungen ist jedoch, Kontextspezifität und damit auch Alternativen der Dateninterpretation systematisch zu vernachlässigen bzw. auch „Fehleinschätzungen“ (DEU02) in Kauf zu nehmen. Diese Diskrepanz zwischen der Notwendigkeit der Vereinfachung komplexer Datenbestände, dem parallel existierenden Wunsch und Streben nach mehr Kontextsensibilität und damit Datengenauigkeit und schließlich der Gefahr von Fehlinterpretationen – die ein US-Behördenmitarbeiter als „*tons of ways to use this data inappropriately*“ (USA17) beschrieb – erachten eine Vielzahl der Befragten als höchst problematisch:

There are plenty of people who go overboard and they overreact to the data that they don't really understand sometimes and they make decisions, even though it's good data you can make a bad decision with good data. (USA18)

Ähnlich beschrieb es ein deutscher Behördenmitarbeiter bzgl. des Umgangs mit Ergebnissen von Lernstandserhebungen:

Im schlimmsten Fall gibt es Ärger. Also ich hab schon von [Schulaufsichtspersonen] [...] gehört, die also [...] gesagt haben, ihr schneidet aber schlecht ab und dann müsst ihr ja mal gucken, dass ihr besser werdet und so. Das ist genau nicht der Anspruch. (DEU32)

Neben den bereits genannten Herausforderungen wird die *richtige* Aufbereitung und Kommunikation von Daten zudem durch die zunehmende Erweiterung und auch Differenzierung des adressierten Datenpublikums erschwert:

I've been a little worried that we have so many data tools, as you've seen [...] and we develop them in so many different ways and we deploy them in so many different ways and I'm worried [...] that we are not clear on what these are all for, who should use which ones, what's the right audience and that kind of thing. (USA16)

Entsprechend erhoffen sich viele der von uns befragten Personen eine langfristige Lösung durch stärker individualisierbare, interaktive und flexiblere Dateninstrumente, die verschiedene, nutzerbezogene Anpassungsmöglichkeiten bieten und so helfen, eine höhere Datenkomplexität ohne Attraktivitätsverlust zu kommunizieren.

2.3 Datenrelationierung: Was und wen wie vergleichen?

Eine weitere zentrale Herausforderung staatlichen Schulmonitorings liegt in der Entscheidung, welche in Daten überführten Subjekte oder Objekte wie miteinander relationiert, also in Beziehung gesetzt und welche Räume des Vergleichs damit geschaffen werden. Gerade durch die zunehmende Datenexpansion, -zentralisierung, -interoperabilität, -standardisierung sowie durch Möglichkeiten automatisierter Datenauswertung werden dabei kontinuierlich neue Möglichkeiten des Vergleichs geschaffen, auch über die territoriale Verortung – etwa Nachbarschulen, Bundesland- oder nationale Grenzen – hinaus (siehe hierzu auch Lewis, 2017). Auf der anderen Seite geht jede Vergleichsentscheidung immer mit bestimmten In-Wertsetzungen (Mau, 2017) einher und hat damit nicht nur unmittelbare Relevanz für das damit erzeugte Datenbild von Schulen, Schüler*innen oder Lehrkräften, sondern auch für daraus abgeleitete Steuerungsentscheidungen.

Diese Problematik wurde im Rahmen der Interviews vor allem im Kontext sogenannter *fairer Vergleiche* bei der Auswertung von Leistungstests deutlich, die jedoch in den von uns beforschten Bundesländern/-staaten sehr unterschiedlich umgesetzt werden. Die Grundidee ist hierbei, bei der Auswertung von Schüler*innen- oder Schulleistungsdaten festgesetzte Gruppierungen wie etwa Klassenverbände oder regionale Verortung dadurch zu relativieren, dass alternative Kriterien wie etwa demographische Randbedingungen oder auch individuelle Leistungsverläufe – und damit als für Leistungoutput ebenfalls relevant erachteter Kontext – einbezogen werden. Positiv wurde dabei hervorgehoben, dass die Verwendung einer solchen angepassten (faireren) Relationierung die Datenakzeptanz an Schulen fördere und den Schulen gleichzeitig dabei helfe, ihre eigenen Leistungen realistischer einzuschätzen.

In den deutschen Behörden wird hierfür u. a. ein sogenannter *Sozialindex* verwendet (siehe auch Groot-Wilken, Isaac & Schräpler, 2016), jedoch fließen in den unterschiedlichen Bundesländern jeweils andere Indikatoren, Datenquellen (z. B. Schulstandortdaten, Angaben aus Elternfragebögen, Daten aus Schulverwaltungssystemen oder durchschnittliche Prüfungsergebnisse) und auch Indexwert-Modellierungen ein.⁷ Dass es sich hierbei um einen hochkomplexen und auch politischen Prozess handelt, verdeutlicht die folgende Interviewaussage:

Man muss sich ja gut überlegen, was will ich denn damit steuern, was will ich da abbilden und dementsprechend muss man den Index basteln.
(DEU32)

7 Für eine differenzierte Auseinandersetzung mit ambivalenten Momenten und Effekten im Rahmen von sozialindizierter Steuerung im Bildungsbereich siehe unter anderem Sendzik, 2018.

Vorrangig werden Sozialindizes zur schulischen Ressourcenverteilung verwendet, jedoch ebenso z.B. für die Durchführung „fairer“ Vergleiche von Lernstandserhebungen. Diese Nutzung für unterschiedliche Zwecke wird von den Befragten durchaus problematisiert und beispielsweise als *entfremdet* oder als „Abfallprodukt“ (DEU07) beschrieben. Alternativ geht eine der Behörden inzwischen den Weg, für jeden Zweck unterschiedliche Sozialindizes zu entwickeln, was jedoch wiederum die Datenkomplexität erhöht (siehe Abschnitt 2.2).

Auch in den USA illustrierten die Befragten unterschiedliche Beispiele eines fairen Vergleichs, die wiederum in einzelnen Staaten different umgesetzt werden. So werden in einem Bundesstaat etwa bzgl. einer Vielzahl von Indikatoren jährlich Wachstumsziele definiert, die jede Schule erreichen soll. Diese Ziele werden jedoch nicht absolut festgelegt, sondern relativ (also *fairer*) auf Grundlage statistisch ähnlicher Schulen, die eine derartige Wachstumskurve bereits durchlaufen (und damit die Realisierbarkeit der Ziele belegt) haben. In einem anderen Bundesstaat werden wiederum einzelne Schüler*innen nach ihrem Vorjahrestestwert gruppiert und schul- bzw. klassenunabhängig in ihrer Testwertveränderung verglichen (in diesem Fall also wiederum unabhängig von anderen Kontextfaktoren), um hieraus einen gewichteten Durchschnittswert und damit eine alternative Maßzahl der Gesamtleistung pro Schule zu ermitteln. Gerade die Datenrelationierung jenseits der Klassenverortung wird dabei von der Behörde als hochgradig innovativ und maximal fair gepriesen.

2.4 Zwischen gewollten und ungewollten Effekten von Accountability

Während sich die bisherigen Ambivalenzen erstaunlich ähnlich in deutschen und amerikanischen Bildungsbehörden manifestieren, so zeigen sich im Bereich der Produktion und Nutzung von Daten für Accountability-Zwecke wesentlich stärkere nationale Unterschiede. Allem voran wurde im Rahmen unserer Interviews eine deutlich kritischere Haltung bezüglich standardisierter Leistungsmessung, der Veröffentlichung von Daten sowie der Kopplung von Outputwerten mit positiven und negativen Sanktionen (*high stakes*) in Deutschland sichtbar, was bzgl. der Daten(erhebung, -speicherung und -veröffentlichung) unter anderem in einer signifikant stärkeren Regulierung bzw. in Brüchen von Dateninfrastrukturen resultiert. Dies bedeutet jedoch nicht, dass sich nicht auch die Behördenmitarbeitenden in den USA bewusst sind, dass die Verwendung von Daten für Accountability-Zwecke nicht nur stark von zugrundeliegenden Normen und Werten beeinflusst wird, sondern sich die in diesem Beitrag skizzierten Ambivalenzen durch die Accountability-Nutzung quasi nochmals verstärken. Als entsprechend zentral wird beispielsweise die Indikatorenwahl wahrgenommen:

[W]e've done some pretty deep philosophical discussions when we are debating what indicators to include, how much improvement we should expect to see and it's an interesting balance of this technical side and then this normative side. Because in the end, you've got to say, did this school make it or not? (USA16)

Gleichzeitig zeigt sich auch für Deutschland, dass entgegen aller formalen Skepsis gegen High-Stakes-Modelle zunehmend Daten für eine sukzessiv engere Steuerung von Schulen verwendet werden – sei es durch kontinuierliches Datenfeedback (bzw. „freundliche Belagerung“: Wiarda, 2018), regelmäßige Zielvereinbarungsgespräche oder durch selektive Ressourcenzuweisung wie im Falle des Sozialindex (siehe Abschnitt 2.3). Dies ist auch politisch so gewollt:

„[Die] Stärkung schulischer Eigenverantwortung verlangt von einer Schule folgerichtig, dass Arbeitsziele formuliert werden und Rechenschaft über die erreichten Ergebnisse abgelegt wird. Daher sind neben Lernstandserhebungen und zentralen Abschlussprüfungen in vielen Ländern verschiedene Verfahren der schulischen Inspektion oder anderer externer Evaluation aufgebaut worden, die auch das Gefüge von Schulleitung und Schulaufsicht bzw. Schulbehörde verändern“ (KMK, 2010, S. 23).

Entsprechend betrifft auch die Problematik des *Wie* von Accountability letztendlich Behörden beider Länder auf ähnliche Weise. So ist erklärtes Ziel auf der einen Seite, dass sich die Schulen durchaus an den Daten(modellen) und Indikatoren ausrichten sollen, um z.B. Schulentwicklungsziele zu definieren. Gleichzeitig laufen die Behörden Gefahr, dass durch diese Ausrichtung der Schulen wiederum ungewollte Dynamiken oder *Trade-offs* entstehen. Hierzu gehören die Vernachlässigung nicht-indikatorisierter Bereiche der Schulentwicklung ebenso wie Datenmanipulationen oder das bereits vieldiskutierte *teaching to the test*:

From the accountability perspective, it could be tricky to put [in particular] [...] kind[s] of measures [...]. Like, for example, we're piloting a school climate survey. We could down the road consider putting that in, but you create these centers for teachers or principals or whoever to tell the kids, make sure you fill all those out, the top possible score in a way that's harder to do with an assessment, like it's harder to manipulate assessment graduation rates. So that's where I think the conversation gets more challenging, is around using accountability. (USA16)

So stellte eine deutsche Behörde unter anderem aufgrund von Manipulationsverdachtsfällen bei schulischen Statusauskünften sowie Validitätsproblemen bei Fragebogenerhebungen ihre Sozialindex-Erhebung auf eine ausschließlich aus Statistiken generierte Erhebung um.

Hiermit zusammenhängend ist es für die Behörden in beiden Ländern eine enorme Herausforderung, mögliche Überreaktionen auf Daten, die öffentlich gemacht werden (sollen), vorausszusehen und im besten Fall zu verhindern. Gerade in den USA steht hierbei viel auf dem Spiel, etwa bzgl. des elterlichen Schulwahlverhaltens und damit etwa auch Veränderungen des Immobilienmarktes. Auch berichteten mehrere Behördenmitarbeitende in beiden Ländern, dass die Schulen selbst mitunter höchst sensibel reagieren, wenn sich bspw. ihre Leistungsklassifizierung aufgrund einer Datenmodellanpassung verschlechtert.

Wenn Schulen merken, es wird damit etwas gemacht [...] sollte man genau überlegen, ob man das macht oder nicht. [...] [W]enn man zu öffentlich und offiziell mit diesen Daten umgeht [...] da sind die Schulen sehr empfindlich. (DEU06)

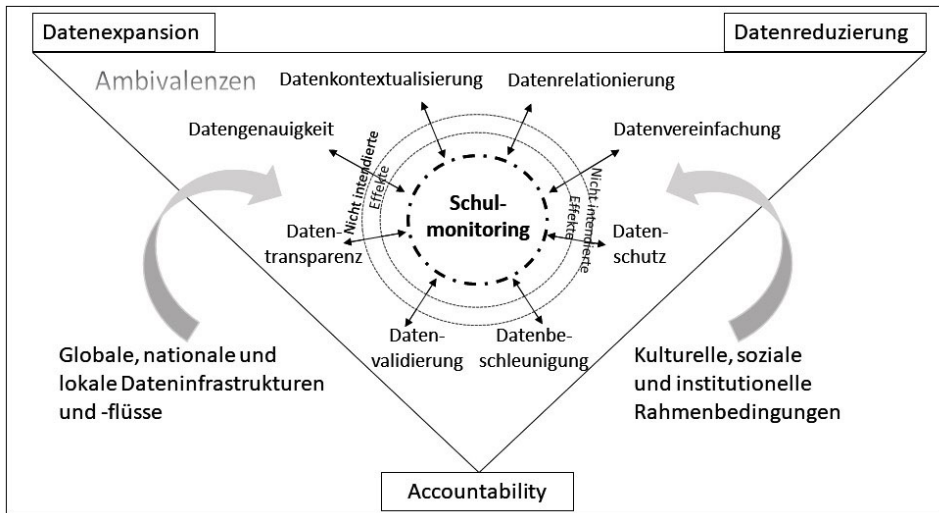
In einem US-Bundesstaat hat man diesbezüglich inzwischen eine Grenze eingeführt, die Modellanpassungen zulässt, sofern 95 Prozent der Schulen hierdurch nicht schlechter gestellt werden (USA37).

3. Schlussfolgerungen

Ziel dieses Beitrags war es, beispielhaft zu illustrieren, wie sich die wachsende Datafizierung und Digitalisierung von Bildungssteuerung auf der Ebene staatlicher Schulmonitoringsysteme auf Bundesland- bzw. -staatenebene manifestieren. Insbesondere ging es hierbei darum, zu zeigen, dass der technischen Vorstellung und auch der Außendarstellung von Schulmonitoringsystemen zwar oftmals simplifizierte Regelkreismodelle zugrunde liegen, die durchaus auch von den Bildungsbehörden selbst zur Außendarstellung genutzt werden. In der Praxis handelt es sich jedoch um ein hochkomplexes Zusammenspiel unterschiedlicher, auch ambivalenter Logiken und Einflüsse sowie Praktiken, die wiederum jeweils spezifische Potenziale und Problematiken mit sich bringen – und die den Behördenmitarbeitenden zum großen Teil sehr bewusst sind (siehe Abb. 2 auf der folgenden Seite).

Im Kontext des hier vorgestellten DFG-Forschungsprojektes wurde versucht, für diese Ambivalenzen eine erste Typisierung bzw. Heuristik zu entwickeln, die gleichzeitig bewusst einen Gegenentwurf zum Regelkreismodell und damit zu bisher prominenten Konzeptionalisierungen datenbasierter Bildungssteuerung darstellt. Obgleich es sich mit anderen Worten bei der Fabrikation von Schulmonitoring also um etwas handelt, das Neil Selwyn (2015) treffend als „messy realities of technology and education“ beschreibt, so lässt sich diese *messiness* dennoch greif-, untersuch- und konzeptionalisierbar machen.

Abb. 2: Ambivalente Datenpraktiken im Kontext des behördlichen Schulmonitorings



Quelle: Hartong & Förschler, 2019, S. 5, übersetzt

Während die ambivalenten Momente der Dateninfrastrukturalisierung von unseren Interviewten in beiden Länderkontexten dabei insgesamt erstaunlich ähnlich beschrieben wurden, so zeigten sich aber auch deutliche Unterschiede zwischen Deutschland und den USA. Insbesondere betreffen diese die Nutzung von Daten für Accountability-Zwecke oder die Veröffentlichung von Daten, was beides in Deutschland nach wie vor weitaus skeptischer beurteilt wird als in den USA.

Insgesamt weisen die Ambivalenzen ganz deutlich auf die – neben der technischen Dimension – hochgradig politische Seite von Schulmonitoring und Dateninfrastrukturen hin, die in Regelkreismodellen oftmals unterbelichtet bleiben. Mit anderen Worten: Dateninfrastrukturen sind immer wertbehaftete Instrumente der Selektion, In-Formation und In-Wert-Setzung von Welt (Thompson & Sellar, 2018), oder, mit den Worten von Kitchin und Lauriault (2014, S. 4–5), „[...] expressions of knowledge/power, shaping what questions can be asked, how they are asked, how they are answered, how the answers are deployed, and who can ask them“. Sie fabrizieren also immer bestimmte datafizierte Repräsentationen von Schule, Lehrkräften, Schüler*innen etc., die Verwaltungsautoritäten wiederum nutzen, um im Namen der Schule, der Lehrkräfte oder der Schüler*innen zu sprechen. Wie unsere Ergebnisse zeigen, schaffen Dateninfrastrukturen dabei nicht nur neue Kategorien (z. B. Schulen mit einem bestimmten Sozialindex oder, wie im Fall der USA, *schools (not) meeting targets*), sondern verändern letztendlich auch die Bildungswelt selbst, indem sie Dinge auf besondere Weise (un)sicht-, wahrnehm-, benenn- und damit steuerbar machen (siehe auch West, 2017). Während neue Daten in diesem Zusammenhang im

Prinzip kontinuierlich einen Bedarf an noch mehr und noch besseren Daten erzeugen (Thompson & Sellar, 2018; Hartong, 2018), erhöhen sie gleichzeitig die Bedeutung bzw. den Raum des Datenmanagements, einschließlich zunehmend notwendiger Daten über die Datenproduktion selbst (Piattoeva, 2016).

Vor diesem Hintergrund plädiert dieser Beitrag dafür, Forschung im Bereich der *Critical Data Studies* auszubauen sowie einen Wissenstransfer zwischen deren Erkenntnissen und der täglichen Praxis von Datenproduzent*innen, Datenmanager*innen, IT-Expert*innen, aber auch Schulleitungen und Lehrkräften zu fördern, um ein kritisches Bewusstsein für die Produktionskomplexität und Nicht-Neutralität von Daten/Monitoring-Instrumenten/Algorithmen/Programmen zu schaffen. Auf Ebene der Behörden kann ein Verständnis für sowie ein Überblick über die Vielschichtigkeit und Komplexität aktueller Dateninfrastrukturen (v.a. in der jeweils „eigenen“ Landesbehörde) jenseits linearer Regelkreismodelle wiederum zu einem reflektierten Umgang und neuen Perspektiven auf die eigene Datenpraxis führen. Ziel wäre mit anderen Worten, was Jonathan Gray, Carolin Gerlitz und Liliana Bounegru (2018, S. 1) als Dateninfrastrukturkompetenz (*data infrastructure literacy*) beschreiben und von einer reinen Datenkompetenz (*data literacy*) im Sinne eines reflektierten Arbeitens mit Datensätzen abgrenzen. Vielmehr geht es ihnen um die Fähigkeit, die breiteren soziotechnischen Infrastrukturen, durch die Daten (nicht nur im Bildungssystem) in wachsendem Maße erzeugt, gespeichert und analysiert werden, besser zu verstehen, sich mit ihnen – und mit ihrer Nicht-Regelkreishaftigkeit – kritisch auseinanderzusetzen, auch um letztendlich aktiv und demokratisch an ihrer Gestaltung mitwirken zu können.

Literatur und Internetquellen

- Abs, H. J., Brüsemeister, T., Schemmann, M., & Wissinger, J. (Hrsg.). (2015). *Governance im Bildungssystem. Analysen zur Mehrebenenperspektive, Steuerung und Koordination*. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-06523-2>
- Allert, H., & Richter, C. (2016). Learning Analytics: subversive, regulierende und transaktionale Praktiken. *Theorieforum Magdeburg* 2016, 1–15.
- Altrichter, H., Moosbrugger, R., & Zuber, J. (2016). Schul- und Unterrichtsentwicklung durch Datenrückmeldung. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 235–277). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-531-18942-0_9
- Ames, M. G. (2018). Deconstructing the Algorithmic Sublime. *Big Data & Society*, 5 (1), 1–4. <https://doi.org/10.1177/2053951718779194>
- Bellmann, J. (2015). Symptome der gleichzeitigen Politisierung und Entpolitisierung der Erziehungswissenschaft im Kontext datengetriebener Steuerung. *Erziehungswissenschaft*, 26 (50), 45–54. <https://doi.org/10.3224/ezw.v26i1.19511>
- Böttcher, W., Bos, W., Döbert, H., & Holtappels, H. G. (Hrsg.). (2008). *Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive*. Dokumentation zur sHerbsttagung der Kommission Bildungsorganisation, -planung, -recht (KBBB). Münster: Waxmann.

- Bormann, I., Hartong, S., & Höhne, T. (Hrsg.). (2018). *Bildung unter Beobachtung. Kritische Perspektiven auf Bildungsberichterstattung*. Weinheim: Beltz Juventa.
- Breiter, A., & Lange, A. (2019). Die digitale Schulverwaltung. In H. H. Lühr, R. Jabkowski & S. Smentek (Hrsg.), *Handbuch Digitale Verwaltung* (S. 330–342). Wiesbaden: Kommunal- und Schul-Verlag.
- Burchardt, M. (2017). Bildung oder Selbstregulation? Zur Anthropologie des „selbst-gesteuerten Lerner“. In E. Hübner & L. Weiss (Hrsg.), *Personalität in Schule und Lehrerbildung: Perspektiven in Zeiten der Ökonomisierung und Digitalisierung* (S. 99–118). Opladen: Barbara Budrich. <https://doi.org/10.2307/j.ctvdf0h6x.7>
- Förschler, A. (2018). Das „Who is who?“ der deutschen Bildungs-Digitalisierungsagenda – eine kritische Politiknetzwerk-Analyse. *Pädagogische Korrespondenz*, 58 (2), 31–52.
- Fraunhofer FKIE (2018). *Interoperabilität in sicherheitskritischen Domänen*. Poster des Fraunhofer-Instituts für Kommunikation, Informationsverarbeitung und Ergonomie (FKIE). Zugriff am 04.12.2019. Verfügbar unter: https://www.sifo.de/files/Poster-10_Interoperabilität_Schmitz_IF18.pdf.
- Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data Infrastructure Literacy. *Big Data & Society*, 5 (2), 1–13. <https://doi.org/10.1177/2053951718786316>
- Groot-Wilken, B., Isaac, K., & Schräpler, J.-P. (Hrsg.). (2016). *Sozialindices für Schulen. Hintergründe, Methoden und Anwendung*. Münster: Waxmann.
- Hartong, S. (2018). „Wir brauchen Daten, noch mehr Daten, bessere Daten!“ Kritische Überlegungen zur Expansionsdynamik des Bildungsmonitorings. *Pädagogische Korrespondenz*, 58 (2), 15–30.
- Hartong, S. (2019a). Politikmobilität und datenbasierte Educational Governance: (Weiter-) Entwicklung einer topologischen Perspektivierung. *Bildung und Erziehung*, 72 (1), 6–23. <https://doi.org/10.13109/buer.2019.72.1.6>
- Hartong, S. (2019b). The Transformation of State Monitoring Systems in Germany and the US: Relating the Datafication and Digitalization of Education to the Global Education Industry. In M. Parreira do Amaral, G. Steiner-Khamsi & C. Thompson (Hrsg.), *Researching the Global Education Industry: Commodification, the Market and Business Involvement* (S. 157–180). Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-04236-3_8
- Hartong, S., Breiter, A., Jarke, J., & Förschler, A. (2020, im Druck). Digitalisierung der staatlichen Schulverwaltung. In T. Klenk, F. Nullmeier & G. Wewer (Hrsg.), *Handbuch Digitalisierung in Staat und Verwaltung*. Wiesbaden: Springer VS.
- Hartong, S., & Förschler, A. (2019). Opening the Black Box of Data-based School Monitoring: Data Infrastructures, Flows and Practices in State Education Agencies. *Big Data & Society*, 6 (1), 1–12. <https://doi.org/10.1177/2053951719853311>
- Hartong, S., & Förschler, A. (2020, im Druck). Dateninfrastrukturen als zunehmend machtvollste Komponente von Educational Governance. Eine Studie zur Implementierung und Transformation staatlicher Bildungsmonitoringsysteme in Deutschland und den USA. In *Tagungsband „Bewegungen“. Beiträge zum 26. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft*. Opladen et al.: Barbara Budrich.
- Heinrich, M. (2018). Ökonomisierung der Schule durch evidenzbasierte Schulentwicklung? Analysen zur Schulinspektion im Rahmen des Effizienzversprechens „Neuer Steuerung“. In S. Hartong, B. Hermstein & T. Höhne (Hrsg.), *Ökonomisierung von Schule? Bildungsreformen in nationaler und internationaler Perspektive* (S. 173–191). Weinheim: Beltz Juventa.
- Iliadis, A., & Russo, F. (2016). Critical Data Studies: An Introduction. *Big Data & Society*, 3 (2), 1–7. <https://doi.org/10.1177/2053951716674238>
- Karcher, M. (2018). Die neue Übersichtlichkeit. Monitoring, datenbasierte Steuerung und Entpolitisierung. In I. Bormann, S. Hartong & T. Höhne (Hrsg.), *Bildung unter*

- Beobachtung. Kritische Perspektiven auf Bildungsberichterstattung* (S. 66–88). Weinheim: Beltz Juventa.
- Kitchin, R., & Lauriault, T. P. (2014). *Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work*. Zugriff am 05.08.2019. Verfügbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112.
- Kitchin, R., Lauriault, T. P., & McArdle, G. (2015). Knowing and Governing Cities through Urban Indicators, City Benchmarking and Real-Time Dashboards. *Regional Studies, Regional Science*, 2 (1), 6–28. <https://doi.org/10.1080/21681376.2014.983149>
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2010). *Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung*. Köln: Carl Link.
- Lawn, M. (Hrsg.). (2013). *The Rise of Data in Education Systems*. Oxford: Symposium Books.
- Lewis, S. (2017). PISA for Schools: Respatializing the OECD's Global Governance of Education. In A. W. Wiseman & C. Stevens Taylor (Hrsg.), *The Impact of the OECD on Education Worldwide* (International Perspectives on Education and Society, Bd. 31) (S. 181–206). Bingley: Emerald Publishing Limited. <https://doi.org/10.1108/S1479-367920160000031010>
- Maritzen, N., & Tränkmann, J. (2015). Normative Grundlagen des Bildungsmonitorings. *DDS – Die Deutsche Schule*, 107 (3), 232–247.
- Mau, S. (2017). *Das metrische Wir: Über die Quantifizierung des Sozialen*. Berlin: Suhrkamp.
- Niedlich, S., & Bormann, I. (2019). Steuerungswissen. In R. Langer & T. Brüsemeister (Hrsg.), *Handbuch Educational Governance Theorien* (S. 491–522). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-22237-6_23
- Piattoeva, N. (2016). The Imperative to Protect Data and the Rise of Surveillance Cameras in Administering National Testing in Russia. *European Educational Research Journal*, 15 (1), 82–98. <https://doi.org/10.1177/1474904115617767>
- Rürup, M. (2014). Landesinstitute für Qualität als neue Akteure im deutschen Schulwesen? Governancetheoretische Annäherungen an eine Alltagstheorie. *DDS – Die Deutsche Schule*, 106 (3), 201–218.
- Rürup, M. (2018). Berichtet Bildungsberichterstattung über Bildung? Eine Auseinandersetzung mit KritikerInnen. In I. Bormann, S. Hartong & T. Höhne (Hrsg.), *Bildung unter Beobachtung. Kritische Perspektiven auf Bildungsberichterstattung* (S. 16–42). Weinheim: Beltz Juventa.
- Selwyn, N. (2015). Data Entry: Towards the Critical Study of Digital Data and Education. *Learning, Media and Technology*, 40 (1), 64–82. <https://doi.org/10.1080/17439884.2014.921628>
- Sendzik, N. (2018). Faire Ungleichbehandlung durch Sozialindizes. Kann Evidenz helfen, Ungleiches ungleich zu behandeln? In K. Drossel & B. Eickelmann (Hrsg.), *Does „What works“ work? Bildungspolitik, Bildungsadministration und Bildungsforschung im Dialog* (S. 291–309). Münster: Waxmann.
- Thompson, G., & Sellar, S. (2018). Datafication, Testing Events and the Outside of Thought. *Learning, Media and Technology*, 43 (2), 139–151. <https://doi.org/10.1080/17439884.2018.1444637>
- West, J. (2017). Data, Democracy and School Accountability: Controversy over School Evaluation in the Case of DeVasco High School. *Big Data & Society*, 4 (1), 1–16. <https://doi.org/10.1177/2053951717702408>
- Wiarda, J.-M. (2018). Der kühle Blick. *Süddeutsche Zeitung vom 22.07.2018*. Zugriff am 05.08.2019. Verfügbar unter: <https://www.sueddeutsche.de/bildung/schulen-der-kuehle-blick-1.4064279>.

Williamson, B. (2017). *Big Data in Education. The Digital Future of Learning, Policy and Practice*. Los Angeles, CA: SAGE.

Annina Förschler, M. A., geb. 1987, wissenschaftliche Mitarbeiterin und Doktorandin im DFG-Projekt „Bildungsdatenmanagement“, Helmut-Schmidt-Universität Hamburg, Fachbereich Erziehungswissenschaft.

E-Mail: foerschler.annina@hsu-hh.de

Sigrid Hartong, PD Dr., geb. 1985, wissenschaftliche Mitarbeiterin, Projektleiterin des DFG-Projektes „Bildungsdatenmanagement“ sowie Teilprojektleiterin des BMBF-Projektes „DATAFIED“, Helmut-Schmidt-Universität Hamburg, Fachbereich Erziehungswissenschaft.

E-Mail: hartongs@hsu-hh.de

Korrespondenzadresse: Helmut-Schmidt-Universität Hamburg, Holstenhofweg 85, 22043 Hamburg