

Hartong, Sigrid

**"Wir brauchen Daten, noch mehr Daten, bessere Daten!". Kritische Überlegungen zur aktuellen Expansionsdynamik des Bildungsmonitorings**

formal überarbeitete Version der Originalveröffentlichung in:

formally revised edition of the original source in:

Pädagogische Korrespondenz (2018) 58, S. 15-30



Bitte verwenden Sie beim Zitieren folgende URN /  
Please use the following URN for citation:  
urn:nbn:de:0111-pedocs-211058 - <http://nbn-resolving.org/urn:nbn:de:0111-pedocs-211058>

in Kooperation mit / in cooperation with:



**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, vertrieben 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 use 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  
Internet: www.pedocs.de

Mitglied der  
  
Leibniz-Gemeinschaft

---

INSTITUT FÜR PÄDAGOGIK UND GESELLSCHAFT

---

# PÄDAGOGISCHE KORRESPONDENZ

---

HEFT 58

---

HERBST 2018

---

*Zeitschrift für  
Kritische Zeitdiagnostik  
in Pädagogik und  
Gesellschaft*

---

BUDRICH UNIPRESS OPLADEN & TORONTO

---

Die Zeitschrift wird herausgegeben vom  
Institut für Pädagogik und Gesellschaft e.V. Münster,  
im Verlag Budrich UniPress, Leverkusen

*Redaktionsadresse ist:*

Institut für Pädagogik und Gesellschaft e.V.  
Windmühlstraße 5, 60329 Frankfurt am Main

*Redaktion:*

Peter Euler (Darmstadt)  
Andreas Gruschka (Frankfurt/Main)  
Bernd Hackl (Graz)  
Andrea Liesner (Hamburg)  
Andreas Wernet (Hannover)  
Antonio Zuin (Saõ Carlos)

*Schriftleitung*

Harald Bierbaum (Darmstadt)  
Karl-Heinz Dammer (Heidelberg)  
Sieglinde Jornitz (Frankfurt/Main)  
Marion Pollmanns (Flensburg)

Manuskripte werden als Word-Datei an Sieglinde Jornitz ([jornitz@dipf.de](mailto:jornitz@dipf.de))  
oder Marion Pollmanns ([marion.pollmanns@uni-flensburg.de](mailto:marion.pollmanns@uni-flensburg.de)) erbeten und  
durchlaufen ein Begutachtungsverfahren.

*Abonnements und Einzelbestellungen:*

Institut für Pädagogik und Gesellschaft e.V.  
Windmühlstraße 5, 60329 Frankfurt am Main

Der Jahresbezugspreis der *Pädagogischen Korrespondenz*  
beträgt im Inland für zwei Ausgaben 23,- EURO zzgl. 4,- EURO Versand.  
Das Einzelheft kostet im Inland 12,50 EURO zzgl. 2,50 EURO Versand.  
Bezugspreise Ausland jeweils zzgl. gewünschtem Versandweg.  
Kündigungsfrist: schriftlich, drei Monate zum Jahresende.

*Copyright:*

© 2018 für alle Beiträge soweit nicht anders vermerkt sowie für  
den Titel beim Institut für Pädagogik und Gesellschaft, Münster.  
Originalausgabe. Alle Rechte vorbehalten.  
ISSN 0933-6389

*Buchhandelsvertrieb:*

Institut für Pädagogik und Gesellschaft e.V.

*Satz & Layout:* Susanne Albrecht, Leverkusen

*Anzeigen und Gesamtherstellung:*

Verlag Budrich UniPress Ltd., Stauffenbergstr. 7, D-51379 Leverkusen  
ph +49 (0)2171 344694 • fx +49 (0)2171 344693  
[www.budrich-unipress.de](http://www.budrich-unipress.de)

- 4      **ESSAY**  
*Andreas Gruschka*  
Von der Verpflichtung auf Rationalität zur Rationalisierung von Eigeninteressen. Beobachtungen zu einem globalen Phänomen
- 15     **AUS WISSENSCHAFT UND FORSCHUNG**  
*Sigrid Hartong*  
„Wir brauchen Daten, noch mehr Daten, bessere Daten!“ Kritische Überlegungen zur Expansionsdynamik des Bildungsmonitorings
- 31     **AUS WISSENSCHAFT UND POLITIK**  
*Annina Förtschler*  
Das „Who is who?“ der deutschen Bildungs-Digitalisierungsagenda – eine kritische Politiknetzwerk-Analyse
- 53     **REFORMRÜCKSCHLAG**  
*Michael Brandmayr*  
Chancengerechtigkeit und individualisiertes Lernen: Zur Verzerrung zweier Begriffe und ihrer ideologischen Funktion in Österreich
- 74     **AUS WISSENSCHAFT UND PRAXIS**  
*Oksana Baitinger*  
Auf der Suche nach pädagogischen Zielen der ganztägigen Betreuung von Grundschulkindern
- 86     **AUS WISSENSCHAFT UND LEHRE**  
*Nina Meister*  
„Wir kämpfen miteinander, dann können wir uns auch beim Vornamen nennen.“ — Eine Fallanalyse zu fachkulturellen Besonderheiten im Verhältnis von Hochschullehrenden und Lehramtsstudierenden
- 100    **THEORIE UND KRITIK**  
*Alf Hellinger*  
Das Materialismuskonzept Hans-Jochen Gamms und der Gedanke einer realhumanistischen Erziehung und Bildung
- 115    **AUS DEM GESTRÜPP DES INSTITUTIONALISMUS**  
Wie mit der Korrektur von Abiturklausuren zu verfahren ist

Sigrid Hartong

## „Wir brauchen Daten, noch mehr Daten, bessere Daten!“ Kritische Überlegungen zur aktuellen Expansionsdynamik des Bildungsmonitorings

### I

Dieser Beitrag beschäftigt sich mit dem in Deutschland relativ jungen Phänomen eines systematischen, auf Bildungsstandards, Tests und Berichterstattung beruhenden Bildungsmonitorings. Zu Beginn wird an Ursprung und Wandel von Versuchen der überschulischen Bildungsbeobachtung im internationalen Vergleich erinnert. Dann wird die Argumentation betrachtet, mit der Anfang des 21. Jahrhunderts ein systematisches Bildungsmonitoring in Deutschland eingeführt wurde, nämlich als ein Unterstützungsinstrument, das, anders als etwa in den USA, weder die professionelle Autonomie von Lehrkräften einschränken noch mit Sanktionen (*high stakes*) gekoppelt werden sollte. Dass dies möglich sei, wurde und wird vor allem von Seiten der Bildungspolitik immer wieder betont, in diesem Beitrag jedoch kritisch hinterfragt. Entsprechend werden *Mechanismen* dargelegt, die heutigen Bildungsmonitoringregimen eine Eigendynamik verleihen, die sich auch bei großer Anstrengung schwerlich außer Kraft setzen lässt, sobald sich einmal dazu entschieden wurde, über Bildungsmonitoring so genannte objektive Evidenz produzieren zu wollen. Entsprechend provoziert die Beobachtung dieser Mechanismen nicht nur deutliche Zweifel am Fortschrittsversprechen des Bildungsmonitorings, sondern wirft gleichermaßen die Frage auf, ob Bildungsmonitoring, wie es heute weltweit vorangetrieben wird, nicht zu einer systematischen Gefährdung von Bildung führt.

### II

Jede Form pädagogischer Handlung hängt unmittelbar mit Beobachtungen zusammen, die darauf gerichtet sind, Möglichkeiten oder Prozesse von Bildung oder Lernen zu erkennen und zu fördern. Mit der Entstehung eines „pädagogischen Diskurses über Erziehungsfragen“ (Koller 2017, S. 31) sowie, in einer Vielzahl von Ländern, der Etablierung einer staatlichen Schulverwaltung und -aufsicht – in Deutschland vor allem in der zweiten Hälfte des 18. bis Anfang des 19. Jahrhunderts; in Ländern wie den USA hingegen vor allem seit Mitte des 19. Jahrhunderts – rücken die Komplexitäten dieser pädagogischen Beobachtungssituation systematisch ins Zentrum der Aufmerksamkeit. Dies gilt auch für die Frage, wie durch Formen der staatlichen Beobachtung systematisch-steuernd in pädagogische Beobachtungen im Klassenraum eingegriffen beziehungsweise administrativ mit ihnen umgegangen werden kann. Dabei – und dies wird oftmals vergessen – spielen Daten, etwa in Form schulischer

Statistiken, im Prinzip von Anfang an eine zentrale Rolle, wenngleich diese lange Zeit mittels relativ hemdsärmeliger analoger Technologien produziert und verarbeitet werden.

Gleichzeitig lassen sich bezüglich der Rolle dieser Daten für die Schulsteuerung früh signifikante kulturelle Unterschiede erkennen. So finden sich in den USA bereits seit Mitte des 19. Jahrhunderts Versuche, Effizienz und Effektivität des Schulsystems über standardisierte, insbesondere über Leistungstests generierte Daten zu erhöhen – auch weil hier traditionell eher wenig in die Professionalität von Lehrkräften vertraut wird und sich die Schulsupervision vielmehr als eine managementorientierte Verwaltungsprofession etablieren kann (vgl. Mehta 2013, S. 61; Reese 2013). Ein erster Höhepunkt ist hier die sogenannte *Soziale Effektivitätsbewegung* um 1900 herum, bei der, in mitunter erstaunlicher Ähnlichkeit zu heutigen Reformen (vgl. Wells/Holme 2005, S. 25), fantastisch anmutende Ideen der Lern- und Bildungssteuerungsvermessung nach tayloristischen Prinzipien entworfen und mit Standards, Lernstufen, standardisierten Tests und technologischen Beobachtungsinstrumenten gekoppelt werden (vgl. Waldow 2012; Bellmann 2012). Gerade seit den 1980er Jahren gewinnt diese Entwicklung dann nochmals enorm an Fahrt, während sich gleichzeitig eine zunehmend nationale Standardisierung der Bildungsbeobachtung herauszukristallisieren beginnt (für eine detaillierte Erläuterung siehe Hartong 2018). Die kontinuierliche Expansion erhobener Bildungsdaten spielt hierbei eine entscheidende Rolle im Sinne des Versuchs, mit standardisierten Instrumenten des Bildungsmonitorings durch das in den USA hochkomplexe Mehrebenensystem der Bildung „hindurch“ zu steuern, um einerseits Bildungsgerechtigkeit und Qualität zu verbessern, und um andererseits die Wirkungen teurer politischer Investitionen zu kontrollieren. Nach der Jahrtausendwende, konkret mit dem viel diskutierten *No Child Left Behind*-Gesetz (NCLB), wird die Leistungsbeobachtung beziehungsweise die Leistungszuwachsbeobachtung – denn dass man einen jährlich nachweisbaren „adäquaten Leistungszuwachs“ (*adequate yearly progress*) bei jedem Kind beobachten müsse, wird hierbei vorausgesetzt – dann neben einer nochmaligen Ausweitung von Daten (lokales Handeln muss nun bewiesenmaßen „evidenzbasiert“ sein) das erste Mal mit einem umfassenden Sanktionssystem (*high stakes*) gekoppelt, das es unter anderem möglich macht, Schulen zu schließen, zu privatisieren oder Lehrkräfte zu entlassen.

In Deutschland sah und sieht die Situation bis heute (noch) anders aus. Zwar findet auch hier ab Ende des 18. Jahrhunderts im Prinzip dieselbe Etablierung einer staatlichen Schulsteuerung und damit ein ähnlicher Aufschwung der Standardisierung von Schulbeobachtung statt. Anders als in den USA ist die gesellschaftliche Konstellation der Etablierung des Bildungswesens im Kontext eines selbstbewussten Bildungsbürgertums stark geprägt von der Idee einer autonomen Lehrprofession, die staatlich ausgebildet sowie examiniert und deren „Leistungsoutput“ wenig kontrolliert wird (vgl. Hopmann 2003). Und so lässt sich für den deutschen Kontext festhalten, dass es zwar auch hier staatliche (jedoch gleichzeitig lückenhafte) Bildungsstatistiken für Planungszwecke (Inputsteuerung) gibt, die jedoch bis Anfang des 21. Jahrhundert weder

auf nationaler noch auf Bundeslandebene Leistungsdaten beziehungsweise umfassende Teststrukturen umfassen.

Dass man damit offenbar im Blindflug Bildungssteuerung betrieben hat ohne zu wissen, was „hinten rauskommt“, wird spätestens mit dem Einschlag der PISA-Studie (*Programme for International Student Assessment*) Anfang des 21. Jahrhunderts zum neuen Mantra der Bildungspolitik (vgl. Niemann 2010; Hartong 2012), mit dem Effekt einer politisch forcierten Nachfrage nach umfassenden „eigenen“ (kulturell auch sensibilisierenden) Schulleistungsdaten (als bei PISA) (vgl. Aljets 2015). So wird binnen erstaunlich kurzer Zeit ein umfassendes Reformpaket geschnürt und in der KMK-Gesamtstrategie zum nationalen Monitoring (<https://www.kmk.org/themen/qualitaetssicherung-in-schulen/bildungsmonitoring.html>) verankert, das die Entwicklung und Etablierung nationaler Standards für alle Kernfächer, verschiedener Testinstrumente (Teilnahme an internationalen Tests, nationale Bundesländervergleichstests und bundeslandweite, aber mittlerweile national entwickelte Vergleichsarbeiten), Berichterstattungen auf mehreren Ebenen sowie Programme sogenannter „Kapazitätsbildung“ (zur Sicherstellung der gewünschten Erfüllung der jeweiligen Programme) umfasst.

Gleichzeitig – und dies ist an dieser Stelle bedeutsam – ist diese doch radikale bildungspolitische Wende hin zur Outputorientierung durchaus von großen Bedenken begleitet, sodass ein großer Teil der damaligen Initiierungsarbeit insbesondere auch darin liegt zu betonen, dass man auf keinen Fall werden wolle wie die USA mit ihren omnipräsenten Tests und wettbewerbsorientierten Rankings, sondern das Bildungsmonitoring lediglich zur Unterstützung einsetze als ein *low-stakes*-System, das keinesfalls die Autonomie der Lehrkräfte einschränken oder aber datenbasierte Sanktionen zulassen dürfe. So mahnten etwa die Autoren einer der zentralen nationalen Expertisen zur Bildungsreform im Jahr 2003:

Viele Probleme, die im Zusammenhang mit den sogenannten high stakes assessments in den USA entstanden sind [...], hätten sich vermeiden lassen, wenn dort sorgfältiger zwischen Bildungsmonitoring, Evaluation auf Schul- und Klassenebene und Individualdiagnostik unterschieden worden wäre. [...] Testapplikationen mit individual-diagnostischer Zielsetzung sollten in der Verantwortung des Lehrers stehen (Klieme et al. 2003, S. 108f).

Und sicherlich stimmt es, dass das Bildungsmonitoring in Deutschland, wie es dann implementiert wurde, nach wie vor deutlich weniger *high-stakes*-Charakter hat als in den USA und dass hierzulande bislang noch keine Lehrkräfte entlassen oder Schulen geschlossen werden, wenn sie von Bildungsmonitoring-instrumenten als *failing* identifiziert werden. Aber das Bildungsmonitoring bewegt sich – bereits in den letzten 10 Jahren hat es sich deutlich bewegt (siehe auch Bormann et al. 2018), und zwar zum einen in Richtung einer kontinuierlichen Ausweitung und sukzessiven Verfeinerung, zum anderen in Richtung einer wachsenden Zentralisierung von Dateninfrastrukturen.<sup>1</sup> Und in der Tat

---

1 Dateninfrastrukturen beschreiben das Netz aus Objekten (die Daten selbst, aber auch Hard- und Software oder Politikfragmente wie etwa Standards) und Subjekten (TechnikerInnen, VerwalterInnen, Schulakteure, intermediäre Akteure etc.), die um Bildungsdaten und ihre sozio-technische De- und Rekontextualisierung herum arrangiert sind und diese beeinflussen (vgl. Anagnostopoulos, Rutledge und Jacobsen 2013, S. 8).

wirft eine Betrachtung dieser Bewegungen, wie sie dieser Beitrag anhand von Beispielen sowohl aus den USA als auch aus Deutschland ins Zentrum stellen will, die Frage auf, ob dem Bildungsmonitoring nicht systematisch unterschätzte Eigendynamiken innewohnen, die ihre Wirkung in jedem Kontext entfalten – wenngleich vielleicht in unterschiedlichem Tempo –, sobald man sich einmal bildungspolitisch dazu entscheidet, vermeintliche Evidenz über maximal rationale Datenproduktionen generieren und Entscheidungsfindungen hieran ausrichten zu wollen. Ich möchte argumentieren, dass es plausible Gründe gibt, diese Frage zu bejahen, und diese Gründe im Folgenden genauer darlegen.

### III

Evidenzbasierte Bildungssteuerung lebt von der Maxime, bildungspolitische und -praktische Entscheidungen auf eine objektive und rationale Informationsgrundlage in Form von (idealerweise quantifizierten, weil dadurch leichter vergleichbaren) Daten zu stellen. Doch was so eindeutig scheint und was heute in so vielen Leitfäden, Fortbildungsinstrumenten und Ideen zur Etablierung von Monitoringinstrumenten verankert und verbreitet wird, nämlich dass man lediglich einer linearen Produktionskette aus Problemdefinition, Datenerhebung (oder -sammlung), -auswertung und Handlungsanpassung zu folgen habe, um evidenzbasiert zu handeln, genau das scheint bei näherer Betrachtung alles andere als einfach, sondern ein zentrales Problem: Wie lässt sich eine objektive Datenbasis fabrizieren (vgl. Piattoeva et al. 2018) – und zwar eine Basis, die die Vermessung und den Vergleich von Qualität ermöglicht?

Anhand eines Beispiels soll das Problem, das hierbei entsteht, verdeutlicht werden. Zwar handelt es sich dabei um ein Beispiel aus der Wissenschaftsteuerung, jedoch ließe es sich genauso auf den Schulbereich übertragen. Seit nicht allzu langer Zeit hat ja auch in der deutschen Wissenschaftspolitik die „Notwendigkeit“ einer stärkeren Outputorientierung Einzug gehalten, die unter anderem umfasst, Wissenschaftler/innen zunehmend auf befristeten Stellen zu beschäftigen und sie gleichzeitig verstärkt in den Wettbewerb um externe Forschungsgelder treten zu lassen. Diese Forschungsgelder soll nur der/diejenige bekommen, der/die zuvor nach objektiven Kriterien seine/ihre Förderwürdigkeit belegt hat und damit auch, dass es rational und sinnvoll ist, ihn/sie mit Geld zu betrauen. Doch wie lässt sich definieren, was „gute“, objektive Kriterien sind, die standardisiert an sämtliche Antragsteller/innen gleichermaßen angelegt werden können?

Wichtig, so meinen viele Geldgeber, sei vor allem, dass der/die Forscher/in international tätig und vernetzt ist. Verglichen lässt sich entsprechend zunächst, ob es internationale Publikationen gibt und, wenn ja, wie viele (wodurch man Zahlenwerte hätte, die sich leicht miteinander vergleichen ließen). Dann lassen sich bereits die ersten Antragsteller/innen aussortieren. Doch wie lässt sich nun entscheiden, ob solch eine internationale Publikation wirklich qualitativ hochwertig ist? Sicherlich könnte der/die Autor/in Auskunft darüber geben, wie viel Arbeit und Zeit in jeweilige Publikationen geflossen ist,

aber die Zeit drängt und wer weiß denn, ob diese Empfindungen nicht viel zu subjektiv sind? Auch hat sich der/die ein oder andere Antragsteller/in vielleicht primär in Sammelbände befriedeter Herausgeber/innen „hineingenetzwert“ und damit dem Urteil der Wissenschaftscommunity (das man pauschal als höher bewertet als die Urteile befriedeter Herausgeber/innen) nicht ausgesetzt. Besser wäre es also, auf internationale Zeitschriften zu setzen und zwar auf solche, die nach dem Prinzip des *double blind peer review* (mehr oder weniger) zufällig anonyme externe Gutachter/innen für eingereichte anonymisierte Manuskript benennen. Am besten wäre dann natürlich, wenn der/die Autor/in alleine den Beitrag verfasst hat. Und so lässt sich nach dieser Entscheidung wieder eine neue Aussortierung von Kandidat/innen vornehmen.

Aber leider gibt es inzwischen recht viele derartiger *peer review*-Journale, sodass erneut die Frage entsteht, wie sich nun wiederum feststellen ließe, welche dieser Journale wirklich die qualitativ hochwertigsten sind. Tatsächlich hilft hier inzwischen die automatisierte Technik, basierend auf dem sogenannten *impact factor*, bei der Algorithmen auf Grundlage der Auswertung von *big data* vermessen (sollen), wie oft ein Artikel zitiert und damit, so hofft man, von der Wissenschaftscommunity zur Kenntnis genommen wird.

Die Kette ließe sich endlos weiterführen. Was verdeutlicht werden soll und was zentral erscheint, wenn wir uns mit der Produktion von Bildungsmonitoringdaten beschäftigen wollen: Die Fabrikation objektiver Daten der Beobachtung von Welt – im genannten Beispiel der Qualität von Wissenschaftlern/innen – ist niemals ein einfacher, unproblematischer und noch viel weniger ein neutraler Prozess, sondern vielmehr begleitet von zahlreichen Momenten der bewussten oder unbewussten Entscheidung auf Grundlage hintergrundiger Annahmen (etwa: internationale Publikationen messen internationale Tätigkeiten, jede Publikation in *peer-reviewed* journals mit hohem *impact factor* ist gleichwertig usw.) und von der permanenten Wahl zwischen Alternativen, z.B. zwischen Indikatoren, von denen man meint, sie seien bestmöglich Ausdruck für etwas anderes. Dies scheint umso mehr zu gelten, je stärker versucht wird, der Komplexität von Welt wenigstens ansatzweise Genüge zu tun. Und während man meint, damit ein besseres, ein genaueres Instrument zu schaffen (wie gesagt, man könnte ja auch einfach Publikationen zählen), so handelt sich die Datenproduktion mit jeder Verfeinerung eine neue Kette von Problemen ein, die es wiederum notwendig machen, neue Kriterien zu definieren und neue Entscheidungen zu fällen.

Entscheidend ist jedoch, dass dieser gesamte Bereich der Datenproduktion – dessen In-Wert-Setzungsprozesse (vgl. Mau 2017) immer auch von Fragen finanzieller Möglichkeiten, persönlicher Interessen oder politischer Vorstellungen beeinflusst werden – normalerweise nicht derjenige Teil des Bildungsmonitorings ist, der mit Betroffenen debattiert wird oder für deren Problematik Politik und Öffentlichkeit sensibilisiert werden. Und die Problematik erscheint umso relevanter, je schwieriger es wird, angesichts zunehmender Digitalisierung und damit der Verwendung immer komplexerer algorithmischer Berechnungstechniken die Datenproduktionen nachvollziehbar zu machen.

Die wachsende Komplexität der PISA-Studie, deren Produktion die OECD im Rahmen umfangreicher technischer Reports dokumentiert,<sup>2</sup> ist hierfür ein ebenso anschauliches Beispiel wie die in vielen amerikanischen Bundesstaaten trotz ihrer Umstrittenheit nach wie vor gängige Praxis der sogenannten leistungsbasierten Lehrkräfteevaluation (*value-added modeling*), bei der die Leistung einer Lehrkraft daran gemessen wird, welchen Leistungszuwachs sie bei den Schüler/innen innerhalb eines bestimmten Zeitraums erzielt hat. Doch was zunächst relativ simpel erscheint – etwa der Vergleich von Testwerten vor und nach einem Jahr des Unterrichtens durch die Lehrkraft –, wird zunehmend kompliziert, wenn beispielsweise der Kontext der Klasse oder der Schule berücksichtigt werden soll (vgl. Abbildung 1):

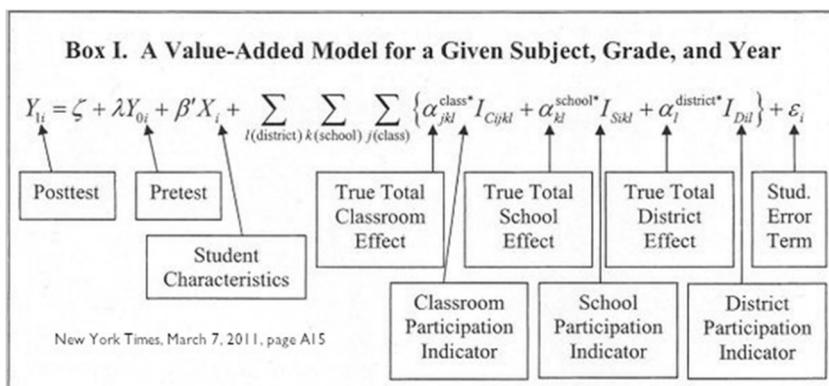


Abb. 1: Beispiel für die Modellierung leistungsbezogener Lehrkräfteevaluation. Quelle: Jarke/Breiter 2015.

Aber auch im deutschen Kontext ist an vielen Stellen das Problem sichtbar, Objektivität mit Bildungsmonitoringinstrumenten mühsam herstellen zu wollen und dabei einen kontinuierlichen Komplexitätszuwachs zum Preis einer von außen zunehmend sinkenden Nachvollziehbarkeit zu erzeugen. Beispielsweise müssen Hamburger Lehrkräfte in der Grundschule inzwischen fast 90 Kompetenzkategorien pro Zeugnis ausfüllen – für Deutsch sind es allein 22 Kompetenzen, zum Glück nur fünf für Bildende Kunst und Sport. Die Erläuterungen zum korrekten Ausfüllen umfassen ein weiteres, fast 20-

2 Auf der einen Seite legt die OECD damit ganz bewusst offen, wie der Produktionsprozess aussieht, und immunisiert sich damit auch gegen Kritik der Intransparenz. Gleichzeitig sprechen die Kommunikationsinstrumente, mit denen die OECD am Ende an Politik, Medien und Öffentlichkeit herantritt – nämlich Rankings, klare politische Empfehlungen oder „eindeutige“ Kurzreports –, eine in der Regel deutlich andere Sprache, bei der der Produktionskomplex zur Black Box wird.

seitiges Dokument. Wie genau diese Kompetenzen zustande kommen,<sup>3</sup> die wiederum in je fünf Niveaus unterteilt werden (von denen das beste jedoch nicht unbedingt ausgefüllt werden soll, weil es zusätzliche Fördermaßnahmen durch die Schule notwendig machen könnte), bleibt für die Lehrkräfte wenig bis nicht nachvollziehbar.

Um den Punkt und damit auch einen ersten grundlegenden Expansionsmotor des Bildungsmonitorings nochmals zu verdeutlichen: Je mehr man sich bemüht, die Komplexität von Realität – die im Bildungsbereich unfassbar groß ist – durch standardisierte, idealerweise quantifizierbare Daten abzubilden und hierfür objektive Monitoringinstrumente zu produzieren, desto länger scheint die Kette von notwendigen Produktionsentscheidungen zu werden, die zwar nach außen hin den Anschein von Objektivität und Rationalität vergrößern, damit aber auch die Anzahl von impliziten In-Wert-Setzungen, die in die Instrumente eingewoben werden, während ihre Produktion gleichermaßen zunehmend technisch detaillierter und damit von außen immer weniger nachvollziehbar wird.

Mit dieser wachsenden Komplexität ist unmittelbar die Frage verbunden, wer derartige Instrumente denn eigentlich (noch) produzieren kann, was sich nicht nur auf die vor allem technische Expertise bezieht, sondern ebenso auf die finanziellen Ressourcen. Denn Datenproduktionen, sollen sie dem hohen Anspruch des Bildungsmonitorings genügen, sind natürlich alles andere als günstig zu haben.

So beschäftigt sich etwa die amerikanische Forscherin Patricia Burch (2009) intensiv mit den Folgen des oben bereits angeführten *No Child Left Behind*-Gesetzes (NCLB) in den USA, das Anfang des 21. Jahrhunderts nicht nur eine detaillierte Leistungsvermessung sämtlicher amerikanischer Schüler/innen im Jahresrhythmus vorsah, sondern ebenso die Zuweisung nationaler Gelder an den Nachweis datenbasierter Evidenzen für die Finanzierung lokaler Implementierungsprogramme koppelte. Was Burch für die folgenden Jahre beobachtet, ist eine systematische Welle der Geldverschiebung in den Staaten und Distrikten aus bestehenden Grundfinanzierungsprogrammen für die Schulen heraus und in die Produktion von Daten, Standards und Evidenz hinein. Da die Staaten und Distrikte hierfür in vielen Fällen weder ausreichend technische noch personelle Ressourcen haben, fließt das Geld zu weiten Teilen an externe Dienstleister. Mit anderen Worten identifiziert Burch für die Jahre nach der Verabschiedung von NCLB eine wahre Explosion neuer Märkte und Dienstleister und damit eine signifikante Kommerzialisierung des Bildungsmonitorings, die sich gleichzeitig sehr implizit (etwa über die Auslagerung bestimmter Aufgaben der Datenproduktion an Dienstleister, über Kooperations- oder

---

3 Im Prinzip sind sie das Resultat eines verworrenen Produktionskomplexes aus nationalen Bildungsstandards, der Entwicklung daran anknüpfender Kompetenzmodelle, der schwierigen Definition von Kompetenzniveaus sowie *Cut Scores* zwischen diesen Niveaus (also wann zählt etwas zum einen, wann zum anderen Niveau), auf Standards basierenden Länderrricula und so weiter.

Partnerschaftsverträge etc.) und damit in weiten Teilen unter dem Radar der Öffentlichkeit vollzieht.

In Deutschland sieht die Situation, jedenfalls bislang, etwas anders aus. Hier sind es weniger gewinnorientierte Privatfirmen, die den Markt der Datenproduktion für das Bildungsmonitoring besetzen (auch wenn es sie in zunehmendem Maße gibt; siehe hierzu auch den Beitrag von Annina Förschler in dieser Ausgabe), sondern empirische Bildungsforschungsinstitute, die sich im Nachgang der PISA-Studie in eine zentrale Position bei der Begleitung der Bildungsmonitoringimplementation gebracht und dabei zunehmend ausdifferenziert haben (vgl. Aljets 2015; Zapp/Powell 2016), um der wachsenden (v.a. technischen) Komplexität der Datenproduktionen gerecht zu werden. Gleichzeitig wird von Seiten der empirischen Bildungsforschung immer wieder die Expansion von Daten (aktuell v.a. Längsschnittdaten) sowie deren zentralisiertere Sammlung (etwa in Forschungsdatenzentren) eingefordert. Denn wo möglichst viele Daten sind, können wiederum spezifischere Forschungsfragen gestellt und die vermeintliche Evidenz für Bildungspolitik erweitert werden. Außerdem können große Datenbanken (z.B. die PISA-Datenbank, die zu einer der umfangreichsten Bildungsdatenbanken der Welt geworden ist) auf mögliche Kausalzusammenhänge und damit auf Hinweise für erfolgreiche Steuerung und Reform durchforstet werden (*data mining*) und eröffnen damit für die Forschung eine völlig neuartige Spielwiese der Erkenntnissuche.

Insgesamt findet also auch in Deutschland – wenngleich auf durchaus (noch) andere Weise als in den USA – eine sukzessive Verschiebung finanzieller Ressourcen, politischer Wahrnehmung und damit von Produktionsmacht für politikrelevantes Bildungswissen statt, bei der das Investieren vieler Millionen Euro in Datenproduktion, -aufbewahrung und -management zunehmend zum politischen Automatismus wird.

#### IV

Ein weiteres Grundproblem der Produktion einer objektiven Datenbasis zu Bildungsmonitoringzwecken ist, dass die Daten oftmals von Personen der Bildungspraxis bei laufendem Alltagsgeschäft eingespeist werden müssen – und zwar genau so, dass sie den Standards des Systems entsprechen und in diesem weiterverarbeitet werden können. Doch offenbar fängt die technizistische, nach außen hin durchrationalisierte Idee der Datenproduktion und -verarbeitung immer dann an zu schwächeln, wenn sie auf die Lebenswelt trifft, die sie auf der einen Seite braucht, um die Daten zu generieren, die auf der anderen Seite jedoch auch vermessen, evaluiert und gesteuert werden soll.

Der PISA-Prozess ist mittlerweile so organisiert, dass sogenannte unabhängige Testmanager in die Schulen gehen, um die (subjektive) Störgröße Lehrkraft aus dem Testprozess zu eliminieren und den PISA-Test maximal standardisiert durchzuführen. Begleitet werden die Testmanager zusätzlich von Qualitätsmanagern, geschult mit standardisierten Leitfäden, die darüber aufklären, was in welcher Situation genau zu tun ist.

Eine solche Praxis ist möglich, wenn der Datenerhebungsprozess an einem Sample von Schulen alle drei Jahre erfolgt – doch wie wird sichergestellt, dass die dauerhafte Erhebung von Leistungen (in Deutschland etwa durch die jährlich durchgeführten Vergleichsarbeiten) oder sonstiger Daten (auch Statistiken), wie sie das Bildungsmonitoring vorsieht, ausreichend objektiv und standardisiert erfolgt? Denn natürlich ist die Sammlung und Dokumentation von Daten insbesondere für Lehrkräfte und Schulleitungen normalerweise eine zusätzliche Tätigkeit, die als Priorität nicht unbedingt ganz oben steht. Am deutlichsten wird die Problematik wohl anhand der in den USA viel diskutierten Praxis des *teaching to the test*, bei der Lehrkräfte sich auf die für sie ggf. folgenschwere Produktion möglichst guter Testwerte konzentrierten und hierfür sogar Betrug oder gezielte Testhilfe für Ihre Schüler/innen in Kauf nehmen (siehe z.B. Nichols/Berliner 2007). Für den hohen Anspruch des Bildungsmonitorings, das ja die Realität – und damit auch „echte“ Leistungsverteilungen – messen will und nicht ein verzerrtes Abbild, ist dies natürlich hochgradig problematisch. Aber auch bei Datenproduktionen ohne unmittelbaren Sanktionswert, etwa statistische Angaben zu Schüler/innen, werden ungenaue, lückenhafte und fehlerbehaftete Dateneingaben zunehmend als Problem diskutiert.

Antworten auf dieses Problem sind in der Regel zum einen die Ausweitung von Schulungen, Beratungen, Fortbildungen und dergleichen, zum anderen die technische Verfeinerung der Datensysteme. Dies kann unter anderem am Beispiel der Entwicklung des sogenannten *EdFacts*-Systems der nationalen Bildungsbehörde in den USA illustriert werden ([www.ed.gov/open/plan/edfacts](http://www.ed.gov/open/plan/edfacts), Zugriff: 26.06.2018). Die Behörde nutzt dieses System seit ca. zehn Jahren, um eine große Menge standardisierter Daten von den Bundesstaaten zu sammeln und für Monitoringzwecke auszuwerten. Implementiert wurde es, weil die Bundesstaaten zuvor zwar Excel-Listen mit (lückenhaften) Daten lieferten, deren Qualität jedoch weder standardisiert gesichert noch nachprüfbar war. So wurde stattdessen ein Gesamtpaket geschaffen, das nicht nur ein national standardisiertes, digitales System der Datenübertragung und -sammlung beinhaltet, in das die Staaten die Daten nunmehr direkt eingeben, sondern auch für jedes zu liefernde Datum – insgesamt für über 100 Indikatoren je ca. 20 Seiten – Datenstandards und sogenannte *business rules* definiert, d.h. Regularien, wie ein Datum zu erheben und was im Zweifelsfall wie zu codieren ist. Kurze Zeit später wurde ein eigenständiges Aus- und Fortbildungsnetzwerk geschaffen, dessen Kernaufgabe es seitdem ist, dafür zu sorgen, dass die Datenstandards auch erfüllt bzw. das System an der Basis „wirklich verstanden“ wird. Dies bedeutet auch, ein möglichst dichtes Netzwerk zuständiger Personen zu installieren, die die Dateneingabe überwachen bzw. die Datenqualität sichern, beispielsweise die sogenannten *data stewards* ([www2.ed.gov/about/initis/ed-edfacts/edfacts-data-stewarding-overview.pdf](http://www2.ed.gov/about/initis/ed-edfacts/edfacts-data-stewarding-overview.pdf), Zugriff: 26.06.2018). In der Behörde selbst wurden wiederum unterschiedliche Qualitätsgremien implementiert, die beispielsweise entscheiden, wenn die Notwendigkeit einer neuen Standarddefinition entsteht. Kernidee des Gesamtprozesses ist es dabei, Daten, die den Standards nicht entsprechen, gar nicht erst in das zentralisierte System

„hineinzulassen“, sondern bereits im Kontext der Dateneingabe zu verhindern, dass die Daten vom System angenommen werden – meist durch das Erscheinen eines Error-Fensters. Auf diese Weise scheinen die Datenlieferanten gezwungen, die Daten erneut auf ihre Richtigkeit zu prüfen und zu korrigieren. Da es jedoch sehr viele Standards und Regeln gibt und dem Datenlieferanten nicht nur mitgeteilt werden muss, dass es Probleme mit den Daten gibt, sondern vielmehr welcher Art diese Probleme sind, sind schließlich Definition und Standardisierung von Error-Nachrichten notwendig, eine standardisierte, detaillierte Dokumentation von Error-Fällen, Error-Schulungen sowie die Etablierung neuer Experten für sogenanntes *error management* ([www2.ed.gov/about/initis/ed/edfacts/dms-state-user-guide.pdf](http://www2.ed.gov/about/initis/ed/edfacts/dms-state-user-guide.pdf), Zugriff: 26.06.2018).

Was das Beispiel verdeutlicht, ist eine weitere Expansionsdynamik im Sinne einer zunehmenden Verfeinerung von Bildungsmonitoringsystemen als notwendige Reaktion auf Performativitätslücken an der Basis, die dem Anspruch einer standardisierten und qualitativ gesicherten Datenqualität (als Voraussetzung für evidenzbasiertes Handeln) entgegensteht, bei der jedoch notgedrungen immer näher an die Datenlieferanten „herangerückt“ werden und der Komplex des Datenmanagements zunehmend ausgeweitet werden muss.

Gleichzeitig betonen die Befürworter/innen des heutigen Bildungsmonitorings immer wieder, dass die Datenerhebung, das Ausfüllen von Listen und dergleichen für die Betroffenen keine zu große Bürde werden dürfe, sondern dass man es letztendlich einfacher für die Personen vor Ort machen wolle, ihre Entscheidungen in Daten zu fundieren und sich gerne mit Daten zu beschäftigen. Tatsächlich wird mit der zunehmenden Automatisierung von Datenerhebungstechnologien, bspw. dem automatisierten Transfer von Statistiken von Schulen zu Behörden, der Umfang menschlicher Dateneingabepraktiken reduziert. Dennoch müssen die Daten, bislang jedenfalls, immer noch in den meisten Fällen zumindest einmal von Hand eingegeben werden. Wirklich eliminieren ließe sich dies nur, wenn die Dateneingabe in eine passive, technische Beobachtungstechnologie übersetzt würde.

Dass eine solche Möglichkeit kein Fantasiegespinst ist, zeigen aktuelle Entwicklungen im amerikanischen Silicon Valley. So wirbt etwa das Start-Up-Unternehmen *AltSchool* nicht nur damit, über einen systematischen Einsatz adaptiver Lernsoftware maximal personalisiert z.B. Lerninhalte zu steuern, sondern gleichzeitig über permanente Video- und Tonaufnahmen des Unterrichts *big* und *thick data* zu produzieren (z.B. über Motivation oder Konzentration der Schüler/innen, etwa abgelesen an Körperhaltungen), die in Echtzeit von Algorithmen ausgewertet und als Handlungsempfehlungen unmittelbar an die Lehrkraft zurückgegeben werden können. Diese Echtzeitsteuerung, so Alt-School, ermögliche ein Echtzeitmonitoring des Unterrichts, mit dem Effekt maximaler Leistung und Lehr-/Lerneffizienz. Leistungsoptimierung und Rechenschaftslegung blieben gleichzeitig nicht mehr auf die (kognitive) Testleistung von Schüler/innen beschränkt, die vielmehr in ihrem gesamten Unterrichts- und Lern-/Lehrverhalten erfasst werden könnten:

AltSchool software and algorithms created by Silicon Valley's top developers and data scientists would [...] search [...] for patterns in each student's engagement level, moods, use of classroom resources, social habits, language and vocabulary use, attention span, academic performance, and more (Herold 2016, S. 1f.).

Wenngleich die Entwicklungen in Deutschland auch hier auf den ersten Blick weit entfernt von derart technologisierten Szenarien zu sein scheinen, so heißt dies nicht, dass Bildungssteuerung und -praxis nicht auch hierzulande durch den Einsatz entsprechender Digitaltechnik in zunehmendem Maße verdichtet werden (siehe den Beitrag von Annina Förschler in dieser Ausgabe). Ein Beispiel sind sogenannte Lernmanagementsysteme, etwa von dem globalen Anbieter *itslearning* (<https://itslearning.com/de>), die in einer wachsenden Anzahl von Bundesländern eingesetzt werden, um nicht nur das digitale Lernen und Lehren zu fördern, sondern gleichermaßen das Datenmanagement von Schulen effizienter zu gestalten. Der Anbieter selbst wirbt hierbei auf der einen Seite:

Im Mittelpunkt Ihrer Bildungsaktivitäten haben Sie eine Lernumgebung, die von Beginn an intuitiv ist und Begeisterung schafft. So intelligent, in dem [sic!] es die Grenzen des physischen Klassenraums überschreitet und Lernmaterialien für jeden individuell parat hält. Und so inspirierend, dass es die Motivation am Unterrichten und Lernen verstärkt.“ (<https://itslearning.com/de/schule/lms-uebersicht>, Zugriff: 26.06.2018).

Auf der anderen Seite heißt es aber auch:

Wie machen sich die SchülerInnen? Sind alle in meinem Verantwortungsbereich auf dem richtigen Weg? Gibt es bestimmte Bereiche, wo im Unterricht nicht die Zielvorgaben erreicht werden? Nehmen die Lehrkräfte die neuen Lehrpläne an? Mit *itslearning* verfügen Schulleitungen und Schulverantwortliche über die erforderlichen Informationen, die sie für ihre Entscheidungen benötigen. Schauen Sie sich an, wie gut Kompetenzziele erreicht werden. Machen Sie sich ein Bild von der Umsetzung der Lehrpläne. Wählen Sie aus einem breiten Angebot von vorbereiteten Berichtsarten oder erstellen Sie Ihre eigenen (<https://itslearning.com/de/#tabbed-35>, Zugriff: 26.06.2018).

Schließlich ist ein wichtiges Feature vieler derartiger Plattformen der (Echtzeit-) Zugriff auf unterrichtsrelevante Informationen für Eltern, die, so die An-

bieter, wesentlich stärker am Schulleben ihrer Kinder (und damit an der Kontrolle dessen, was in Unterricht passiert) teilhaben sollen.

Als weiterer Expansionsmotor des Bildungsmonitorings operiert also schlichtweg die zunehmende Technologisierung, die immer neue Möglichkeiten schafft, eine größere Anzahl Daten auf schnellere Weise zu erheben, auszuwerten und per Interoperabilität ins Bildungsmonitoring einzuweben. Digitale Lernplattformen spielen dabei eine entscheidende, bislang jedoch kaum diskutierte Rolle, denn immer haben diese auch die Funktion, Daten über ihre Nutzer/innen zu erheben und für Steuerungszwecke nutzbar zu machen.

## V

Zusammengefasst sind also unterschiedliche Mechanismen auszumachen, bei denen sich eine zunehmende Datafizierung, dazu notwendige Standardisierung, die voranschreitende Technologisierung und schließlich eine sukzessive Macht-, Geld- und Aufmerksamkeitsverschiebung gegenseitig bedingen, wobei die im Grunde unbändige Suche nach einem objektiven Bildungsmonitoring zum kontinuierlichen, niemals nachlassenden Antriebsmotor wird. Das Resultat ist eine Ausweitung, Ausdifferenzierung, aber gleichzeitig wachsende Zentralisierung des Bildungsmonitorings, was Hempel und Kollegen (2010) mit der Entstehung eines Kontroll- und Sichtbarkeitsregimes beschreiben. Sie definieren dieses als

[...] soziale und technische Arrangements, die Ordnung stiften oder stabilisieren, Gefährdungen abwehren und Abweichungen korrigieren sollen und selbst eine Ordnung des Beobachtens und Beobachtetwerdens, des Zeigens und Verbergens etablieren. Sie wirken gleichermaßen auf das Handeln von Beobachtern und Beobachteten ein, lenken Blicke und dirigieren Aufmerksamkeiten; sie holen Verborgenes ans Licht oder sorgen dafür, dass es den Blicken entzogen bleibt; sie vergrößern Winzigkeiten oder zoomen Weitentferntes heran und machen es so für regulierende Zugriffe erreichbar; sie aggregieren Einzelbeobachtungen und schaffen dadurch erst bestimmte Objekte, z.B. statistische Normalverteilungen oder Risikogruppen (Hempel et al. 2010, S. 8f).

Angesichts dieses Befundes ist es eine berechtigte Frage, ob man es in Deutschland dauerhaft wird schaffen können, die Daten, die mühsam erhoben und qualitätsgesichert werden, am Ende nicht doch zumindest ein wenig dafür zu nutzen, Handlungsdruck aufzubauen oder Formen der (positiven/negativen) Sanktion einzuführen. Dass dies zunehmend wahrscheinlich wird, zeigt etwa das Beispiel Berlin, das neuerdings auf Zielvereinbarungen zwischen Schulleitung und Schulaufsicht setzt:

Dort [in den Zielvereinbarungen, S.H.] soll dann zum Beispiel drin stehen, dass sich die Zahl der unentschuldigten Fehltage – vulgo: das Schwänzen – ganz deutlich reduzieren muss. Auf ähnliche Weise soll so verfahren werden, wenn die Zahl der Schulabbrecher auffallend hoch ist oder die Prüfungsergebnisse schlecht sind (Klesmann 2017).

Schließlich hängt damit auch die Frage zusammen, welche pädagogischen Implikationen mit der Etablierung derartiger Sichtbarkeitsregimes des Bildungsmonitorings einhergehen (können), etwa in Bezug auf das Verständnis von Bildung und Erziehung oder in Bezug auf Folgen für die pädagogische Profession.

In der Tat legen Beobachtungen, wie sie dieser Beitrag versucht hat darzulegen, nahe, dass moderne Sichtbarkeitsregime des datenbasierten Bildungsmonitorings nicht nur nicht in der Lage sind, Bildung zu beobachten (was sie zum Teil auch gar nicht behaupten, siehe PISA), sondern vielmehr der Idee von Bildung zunehmend entgegenstehen (vgl. Asmussen et al. 2017, siehe auch Tabelle 1):

Dimension	Idee von Bildung	Operationen im Sichtbarkeitsregime
Prozessuale Dimension	Keine linearen Entwicklungen, Offenheit des Prozesses	Linearität, Geschlossenheit im Datensystem, Stufenmodelle
Beobachtung (Prozess)	Indirekte pädagogische Beobachtung durch die Lehrkraft (unterstützt durch staatliche Aufsicht)	Durch Daten vermittelte Beobachtung durch Algorithmen, Tests, staatliche Aufsicht, Veränderung pädagogischer Beobachtung
Beobachtung (Idee)	Direkte Beobachtung von Bildung nicht möglich	Direkte(re) Beobachtung möglich(er), je ausgefester Sichtbarkeitsregime
Relevante Akteure	Schüler/in und Lehrkraft (Datenproduktion maximal als „Beiprodukt“)	Datenproduzenten, Qualitätsmanager, Datenvermittler, Dataufbereiter (Pädagogik als „Beiprodukt“)
Visualisierung und Rechenschaftslegung	Praxis der Selbstreflexion nach pädagogischen Kriterien, Zielen auf intersubjektive Verständigung	Praxis der Datenreflexion nach technisch-psychologisch-statistischen Kriterien, Produktion von Objektivität
Vertrauen	In Lehrkräfte	In Daten

*Tab. 1: Schematische Gegenüberstellung der klassischen Bildungsidee und der Operationen im Sichtbarkeitsregime.*

So steht der Bildungsbegriff in vielen Bildungstheorien<sup>4</sup> etwa für eine genuin konflikthafte, unbequeme und biografisch verankerte Auseinandersetzung des Ichs mit der Welt, die weder in letzter Instanz planbar noch in Skalen oder lineare Stufenmodelle zerlegbar ist. Vielmehr müssen Bildungsprozesse systematisch offen bleiben, damit sie zu solchen werden können, was der Geschlossenheit technischer Datensysteme an vielen Stellen widerspricht. Am deutlichsten wird dies etwa bei in der Bildungssteuerung immer beliebteren Längsschnittdatensystemen, bei denen Kinder bereits ab dem Kindergartenalter in

4 So ziehen beispielsweise Asmussen et al. (2017) unterschiedliche Bildungstheorien heran, um sie mit der aktuellen KMK-Strategie zur Bildung in der digitalen Welt zu kontrastieren.

ihren Entwicklungen *getracht* werden und unmittelbar interveniert wird bzw. werden soll, sobald sie sozusagen vom vorgezeichneten Datenkurs abkommen.

Und während Bildung in beiden Fällen von der Idee her in ihrer Komplexität letztendlich nicht beobachtet werden kann, so wird diese Erkenntnis im Sichtbarkeitsregime zur selbstverständlichen Fußnote, die zwar der Objektivitätsbewahrung halber eingefügt, der aber nicht wirklich Beachtung geschenkt werden muss. Dies ist möglich, weil es zunehmend Nicht-Pädagogen und nicht-pädagogische Praktiken (Datenmanagement, Qualitätsmanagement, Error Management) sind, die ins Zentrum rücken und das produzieren, was als Bildung wahrgenommen wird. Dies führt schließlich auch dazu, dass die für Pädagogik so entscheidenden Momente der Reflexion zur Aufgabe transformiert werden, Handeln vor dem Hintergrund der zur Verfügung gestellten und damit als vertrauenswürdig deklarierten Daten zu reflektieren. In der Tat scheint es zunehmend leichter, diese Prämissen auch tatsächlich umzusetzen, da mit jedem weiteren Heranrücken des Sichtbarkeitsregimes (und damit z.B. auch der Eltern, siehe III) an den Unterricht letztendlich der Raum schrumpft, in dem die Schüler/innen einer Lehrkraft unbeobachtet überlassen werden.

Entsprechend und im Sinne einer Gegenbewegung zur Expansionsdynamik der Bildungsmonitorings fordert beispielsweise der Erziehungswissenschaftler Matthias Rürup (2018), dass es wieder den bewussten Beobachtungsverzicht, d.h. die Wiedererschaffung bzw. den Erhalt unbeobachteter Räume geben müsse, und zwar müsse dieser Verzicht „[...] explizit formuliert und begründet werden – als Ausdruck einer wertschätzenden Haltung des öffentlichen Nicht-Wissen-Wollens, um die Eigenständigkeit und Eigenwertigkeit dieser pädagogischen Praxis nicht zu gefährden“ (Rürup 2018: 38). Seine zweite Forderung: Wieder die Divergenz von Berichterstattungen und Beobachtungen zulassen, anstatt zu versuchen (im Sinne der Produktion von maximaler Objektivität), alles zu einem Gesamtbild zu synthetisieren, deswegen (notwendigerweise) zu standardisieren und zu zentralisieren. Er nennt hier das Beispiel lokaler Bildungsberichterstattung und schlägt vor, dass nicht eine einzelne Instanz mit vermeintlich hoher Datenexpertise mit der Produktion objektiver Bildungsberichte beauftragt wird, sondern vielmehr eine Vielzahl lokaler Interessensgruppen aufgefordert werden, Berichte zu schreiben, die ihre jeweiligen Daten und Einschätzungen widerspiegeln – um dann gemeinsam in den Diskurs über diese Daten zu treten. Hierfür wäre es natürlich nötig, den Dissens von Daten bewusst zu provozieren, was dem Glauben an ihre mögliche Objektivierung und damit Neutralisierung entgegensteht. So argumentiert etwa auch Heidrun Allert (2018), dass sich Technologien der Datenproduktion überhaupt nur reflektieren lassen, wenn man verschiedene nutzt und die von ihnen produzierten Bilder kritisch vergleichen kann. Dies würde jedoch gleichzeitig bedeuten, Datenproduktionen des Bildungsmonitorings umfangreich(er) zu thematisieren – insbesondere in der Lehrerbildung – und für deren Problematiken, für die Macht der Daten und auch für die politischen und wirtschaftlichen Interessen dahinter zu sensibilisieren, anstatt angehenden Lehrkräften die Daten als objektive Evidenz ihres Handelns zu präsentieren.

Zentrale Fragen im Kontext einer solchen Thematisierung wären beispielsweise, was passiert, wenn pädagogisches Vertrauen systematisch durch Beobachtung und Kontrolle ersetzt wird, und zwar durchgeführt von zunehmend technisch-anonymisierten Kontrollinstanzen (etwa Lernplattformen), die die Lehrkraft oftmals weder lokalisieren noch in ihrer Produktions- bzw. Entscheidungslogik nachvollziehen kann. Oder was es für Kinder bedeutet, wenn sie ab dem Kindergartenalter einem Sichtbarkeitsregime ausgesetzt werden und sie ihre Datenbiografie mitnehmen (müssen), egal wohin sie gehen (wenn sie beispielsweise die Schule wechseln und ihr Datenportfolio inklusive Vermerke, Fehl- und Krankheitszeiten der erste Eindruck ist, den die neue Schule bekommt). Nicht nur in der Lehrerbildung bedarf es hierbei eines umfangreicheren Diskurses als er zurzeit geführt wird – ebenso wie darüber, was einem sich kontinuierlich expandierenden Bildungsmonitoring entgegengesetzt werden muss, um derartige Gefährdungen einzudämmen.

## Literatur

- Aljets, Enno (2015): Der Aufstieg der Empirischen Bildungsforschung. Ein Beitrag zur institutionalistischen Wissenschaftssoziologie. Wiesbaden.
- Allert, Heidrun (2018): Kritische Bestandsaufnahme: Bildungsverständnis und Digitalisierung. In: Synergie #6, September/2018.
- Anagnostopoulos, Dorothea/Rutledge, Stacey A./Jacobsen, Rebecca (Hrsg.) (2013): The Infrastructure of Accountability. Data Use and the Transformation of American Education. Cambridge, Massachusetts.
- Asmussen, Michael/ Schröder, Christoph/Hardell, Sonja (2017): Bildung in politischen Programmen. Eine pädagogische Revision der KMK-Strategie zur Bildung in der digitalen Welt. In: Leineweber, Christian/de Witt, Claudia (Hrsg.): Digitale Transformation im Diskurs. Kritische Perspektiven auf Entwicklungen und Tendenzen im Zeitalter des Digitalen. URL: <http://www.medien-imdiskurs.de>.
- Bellmann, Johannes (2012): The very speedy solution. Neue Erziehung und Steuerung im Zeichen von Social Efficiency. In: Zeitschrift für Pädagogik, Jahrgang 58, Heft 2, S. 143-158.
- Bormann, Inka/Hartong, Sigrid/Höhne, Thomas (Hrsg.) (2018): Bildung unter Beobachtung. Kritische Perspektiven auf Bildungsberichterstattung. Weinheim/Basel.
- Burch, Patricia (2009): Hidden Markets: The New Education Privatization. New York/London.
- Hartong, Sigrid (2012): Basiskompetenzen statt Bildung? Wie PISA die deutschen Schulen verändert hat. Frankfurt/Main.
- Hartong, Sigrid (2018): Standardbasierte Bildungsreformen in den USA. Vergessene Ursprünge und aktuelle Transformationen. Weinheim/Basel.
- Hempel, Leon/Krasmann, Susanne/Bröckling, Ulrich (2010): Sichtbarkeitsregime. Eine Einleitung. In: Hempel, Leon/Krasmann, Susanne/Bröckling, Ulrich (Hrsg.) Sichtbarkeitsregime. Überwachung, Sicherheit und Privatheit im 21. Jahrhundert. Leviathan Sonderheft 25. Wiesbaden, S. 7-25.
- Herold, Benjamin (2016): The Future of Big Data and Analytics in K-12 Education. In: Education Week, 13.01.16.
- Hopmann, Stefan T. (2003): On the evaluation of curriculum reforms. In: Journal of Curriculum Studies. Jahrgang 35, Heft 4, S. 459-478.
- Jarke, Juliane/Breiter, Andreas (2015): Datafying education: How digital assessment practices reconfigure the organisation of learning. Konferenzpapier: [www.researchgate.net/publication/286863165](http://www.researchgate.net/publication/286863165).

- Klesmann, Martin (2017): Handeln nach Zahlen. Schulleiter sollen mit Hilfe interner Daten ihre Schule besser machen. Jedes Jahr soll die Schulaufsicht die Ergebnisse kontrollieren. In: Berliner Zeitung, Ausgabe 246 (21./22.10.2017), S. 11.
- Klieme, Eckhard/Avenarius, Hermann/Blum, Werner/Döbrich, Peter/Gruber, Hans/Prenzel, Manfred/Reiss, Kristina, Riquarts, Kurt/Rost, Jürgen/Tenorth, Heinz-Elmar/Vollmer, Helmut J. (2003). Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Berlin.
- Koller, Hans-Christoph (2017): Grundbegriffe, Theorien und Methoden der Erziehungswissenschaft. Stuttgart.
- Mau, Steffen (2017): Das metrische Wir: über die Quantifizierung des Sozialen. Frankfurt/Main.
- Mehta, Jal (2013): The Allure of Order: High Hopes, Dashed Expectations, and the Troubled Quest to Remake American Schooling. Oxford.
- Nichols, Sharon L./Berliner, David C. (2007): Collateral damage: How high-stakes testing corrupts America's schools. Harvard Education Press.
- Niemann, Dennis (2010): Turn of the Tide – New Horizons in German Education Policy-making through IO Influence. In: Martens, Kerstin/Nagel, Alexander-Kenneth/Windzio, Michael/Weymann, Ansgar (Hrsg.): Transformation of Education Policy. New York, S. 77-104.
- Piattoeva, Nelli/Centeno, Vera/Suominen, Olli (2018): Governance by data circulation? The production, availability, and use of national large-scale assessments data. In: Kauko, Jaakkko/Rinne, Risto/Takala, Tuoman (Hrsg.): Politics of Quality in Education: A Comparative Study on Brazil, China, and Russia. Routledge (forthcoming).
- Reese, William J. (2013): Testing Wars in the Public Schools A Forgotten History. Cambridge.
- Rürup, Matthias (2018): Berichtet Bildungsberichterstattung über Bildung? Eine Auseinandersetzung mit KritikerInnen. In: Bormann, Inka/Hartong, Sigrid/Höhne, Thomas (Hrsg.): Bildung unter Beobachtung. Kritische Perspektiven auf Bildungsberichterstattung. Weinheim/Basel, S. 16-42.
- Waldow, Florian (2012): Taylorismus im Klassenzimmer: John Franklin Bobbitts Vorschläge zur standards-based reform. In: Zeitschrift für Pädagogik, Jahrgang 58, Heft 2, S. 159-176.
- Wells, Amy S./Holme, Jennifer J. (2005): Marketization in Education: Looking Back to Move Forward with a stronger critique. In: Bascia, Nina/Cumming, Alister/Datnow, Amanda/Leithwood, Kenneth/Livingstone, David (Hrsg.): International Handbook of Education Policy. Dordrecht, S. 18–53.
- Zapp, Mike/Powell, Justin (2016). How to construct an organizational field: Empirical educational research in Germany, 1995–2015. In: European Educational Research Journal, Jahrgang 15, Heft 5, S. 537-557.

*Diese Forschung wurde unterstützt durch Mittel der Deutschen Forschungsgemeinschaft.*