

Mirko Krüger, Johannes Rosendahl

## Rezension zu

Demski, D. & Racherbäumer, K. (2017). What data do practitioners use and why? Evidence from Germany comparing schools in different contexts. *Nordic Journal of Studies in Educational Policy*, 3(1), 82–94. DOI: 10.1080/20020317.2017.1320934.

## Kommentierter Kurzbefund

Infolge des schlechten Abschneidens in verschiedenen internationalen Schulleistungsstudien um die Jahrtausendwende wurde und wird zunehmend datengestütztes Entscheiden und Handeln auf verschiedenen Ebenen des deutschen Bildungssystems gefordert (u. a. auf Schulebene). Gleichwohl mangelt es an Erkenntnissen darüber, inwieweit Schulpraktikerinnen und -praktiker verschiedene Arten von Daten als nützlich erachten, in welchem Maße sie diese Daten für ihre Arbeit heranziehen und welche Bedingungen aus ihrer Sicht eine daten- bzw. evidenzbasierte Schulentwicklung hemmen oder fördern. Insbesondere für Schulen in herausfordernden Lagen gilt datengestütztes Handeln als adäquate Strategie, doch gibt es in der Praxis Unterschiede zu Schulen in privilegierten Sozialräumen?

Demski und Racherbäumer präsentieren hierzu Befunde aus drei Studien. Untersucht wurden die wahrgenommene Nützlichkeit sowie die selbstberichtete Nutzung von 13 verschiedenen Datenquellen, aus denen Lehrkräfte und Schulleitungen Schlussfolgerungen für ihre Arbeit ziehen können (u. a. Schulinspektion, kollegiale Hospitation, Vergleichsarbeiten). Insgesamt basieren die Auswertungen auf Fragebogenangaben von 1.040 Lehrkräften und Schulleitungen des Sekundarbereichs sowie auf (halbstrukturierten) Interviews mit 44 Lehrkräften und 15 Schulleitungen.

Die Befunde zeigen, dass Lehrkräfte und Schulleitungen standardbasierte Reformmaßnahmen bzw. die damit gewonnenen „harten“ Daten als weniger nützlich einschätzen (v. a. Schulinspektion, Vergleichsarbeiten, Schulleistungsstudien). Stattdessen nutzen sie bevorzugt prozessorientierte und schulintern generierte Daten (z. B. Parallelarbeiten, kollegiale Hospitationen, Schülerfeedback). Die Datennutzung an Schulen in herausfordernden Lagen fällt – außer im Hinblick auf Vergleichsarbeiten – überwiegend geringer aus als an Schulen in privilegierten Sozialräumen, obwohl die Lehrkräfte an diesen Schulen die allgemeine Nützlichkeit der verschiedenen Datenquellen tendenziell eher höher einschätzen.

Aus den Interviews geht hervor, dass eine geringe Datennutzung auf fehlende Zeitressourcen zurückgeführt wird. Daneben bestehen Schwierigkeiten bei der Dateninterpretation und Übertragung auf den eigenen Handlungskontext. Zudem wird von einigen Lehrkräften die Datenqualität bemängelt, insbesondere von Schulinspektionen und standardbasierten Leistungstests. Eine bedeutsame Rolle kommt der Schulleitung zu, da sie durch ihren Führungsstil und eigene Aktivitäten den Umgang mit und die Nutzung von Daten an ihrer Schule wesentlich beeinflussen kann.

Die Autorinnen plädieren dafür, bei der Erhebung und Präsentation von Daten für die Schul- und

Unterrichtsentwicklung die Perspektive der Schulpraktikerinnen und -praktiker einzubeziehen und ihre Vorerfahrungen sowie ihre Situation zu berücksichtigen, da sie dann aus den Daten leichter anschlussfähige Informationen für ihre schulinterne Kommunikation, für (kollektive) Bedeutungsherstellung und Bezüge zur eigenen Praxis herstellen können.

## Hintergrund

Demski und Racherbäumer weisen einleitend darauf hin, dass das schlechte Abschneiden in verschiedenen internationalen Vergleichsstudien um die Jahrtausendwende zu einer neuen, den Output und Outcome betreffenden Steuerungslogik geführt habe. Hierzu wurden neue Steuerungsinstrumente eingeführt (z. B. Bildungsstandards, Schulinspektion, zentrale Abschlussprüfungen). Analog vollziehe sich in den letzten Jahrzehnten eine Schulautonomisierung mit neuen Entscheidungsfreiheiten und Gestaltungsspielräumen auf Einzelschulebene. In diesem Zusammenhang werde diskutiert, inwiefern Schulen im Blick auf ihre sozialräumliche Lage und damit verbundenen Risikofaktoren mit zusätzlichen finanziellen Ressourcen zur Herstellung bzw. Aufrechterhaltung von Chancengleichheit ausgestattet werden sollten. Hierzu würden durch die Bildungsadministration Daten erhoben und den Schulen u. a. zur Schul- und Unterrichtsentwicklung zur Verfügung gestellt. Da für dieses Vorgehen öffentliche Gelder verwendet werden, sei die Nutzung und Nützlichkeit der Daten aus Sicht der schulischen Akteure von Bedeutung. Datengestützte Entscheidungen und Handlungspraxen würden insofern als relevant für die Qualitätssicherung und -steigerung im Mehrebenensystem Schule angesehen. Da in Deutschland im Gegensatz zu anderen Ländern eine vergleichsweise geringe Rechenschaftslegung mit wenigen Konsequenzen für die Schulen besteht, sei die Datennutzung eher eine Frage der Bereitschaft schulischer Akteure als eine Frage des möglichen administrativen Drucks.

Zur Datennutzung liegen jedoch bislang kaum Erkenntnisse vor. Die beiden Autorinnen stellen heraus, dass in der Literatur keine Klarheit hinsichtlich des Evidenzbegriffs besteht. Dabei wird zwischen Evidenz im engeren Sinne (z. B. Informationen über Schülerleistungen aus standardbasierten Tests) und Evidenz im weiten Sinne (z. B. Schülerfeedback) unterschieden. Demski und Racherbäumer schließen sich dem an und differenzieren Evidenz im engeren Sinne mit einem stark ausgeprägten wissenschaftlichen Fokus und externalem Impetus (z. B. Ergebnisse aus Vergleichsarbeiten) von Evidenz im breiteren Sinne mit einem gering ausgeprägten wissenschaftlichen Fokus und internalem Impetus (z. B. Schülerfeedback).

Die Befundlage zur Datennutzung in Deutschland beschränkt sich eher auf einzelne Steuerungsinstrumente (z. B. Vergleichsarbeiten) und ist insgesamt überschaubar. Ähnlich verhält es sich in den Vereinigten Staaten. Zudem wird empfohlen, für die Betrachtung von Bedingungen, Prozessen und Wirkungen einer datengestützten Schulentwicklung eher eine vergleichende Analyse verschiedener Informationsquellen vorzunehmen. Dies wird u. a. damit begründet, dass eine Vielfalt der Datennutzung im Zusammenhang mit guten Schülerleistungen und günstigen Schulentwicklungsprozessen steht. Als potenzielle Einflussfaktoren werden die Eigenschaften der Evidenz an sich (v. a. Verfügbarkeit, Zeitlosigkeit und Validität der Information), Merkmale des organisationalen Kontextes (v. a. die Schulkultur und das Schulleitungshandeln) sowie Charakteristiken der individuellen Akteure (v. a. Überzeugungen, Wissen und Erfahrungen) erachtet. Nach dem bisherigen Forschungsstand in Deutschland scheint dabei die Datennutzung und die damit verschränkte Interpretation und Reflexion für Schul- und Unterrichtsentwicklungsprozesse zu variieren.

Vor diesem Hintergrund präsentieren die beiden Autorinnen Ergebnisse aus drei Studien. Sie arbeiten

heraus, welche Daten von Schulleitungen und Lehrkräften als nützlich eingeschätzt werden und welche sie tatsächlich nutzen. Zudem analysieren sie Bedingungsfaktoren hierfür. Da angenommen wird, dass der Schulkontext einen Einfluss auf die Datennutzung haben könnte, wird bei der Auswahl der Schulen systematisch zwischen solchen aus schwieriger und solchen aus privilegierter Lage variiert.

## Design

Zur Beantwortung der Fragestellungen werden Ergebnisse aus drei Studien berichtet.

### Stichprobe

Die Stichprobe der ersten Studie umfasste 424 Lehrkräfte und 129 Schulleitungen des Sekundarbereichs aus Rheinland-Pfalz. Alle Schulen des Landes waren zur Teilnahme an der Untersuchung eingeladen worden. Zur zweiten Stichprobe zählten 487 nordrhein-westfälische Lehrkräfte von 33 Schulen des Sekundarbereichs, die sich in sozioökonomisch benachteiligten Gebieten befanden, hohe Quoten an Schülerinnen und Schülern mit Migrationshintergrund und/oder geringe sprachliche Leistungen aufwiesen und sich an einem Forschungsprojekt zu Multikulturalismus und Mehrsprachigkeit beteiligten. In der dritten Stichprobe befanden sich 8 Schulleitungen und 16 Lehrkräfte von Schulen aus deprivierter Lage in Nordrhein-Westfalen. Die Auswahl dieser Schulen erfolgte über eine durch die Qualitäts- und Unterstützungsagentur NRW (QUA-LiS NRW) vorgenommene Kontrastierung auf der Grundlage erwartungswidrig bzw. erwartungskonform guter Ergebnisse – basierend auf sozialräumlichen Indikatoren – bei den landesweiten Lernstandserhebungen in Klasse 8. Weiterführende Informationen etwa zu demographischen Hintergrundvariablen der involvierten schulischen Akteure werden nicht berichtet.

### Untersuchungsablauf und Konstrukte

In den Studien 1 und 2 wurden Einschätzungen zur wahrgenommenen Nützlichkeit und die selbstberichtete Nutzung von 13 Datenquellen mittels Fragebogen erhoben (5-stufige Likert-Skala: „gar nicht nützlich“ bis „sehr nützlich“):

1. Berichterstattung zu Schulleistungsvergleichen (z. B. PISA)
2. Landesweite Lernstandserhebungen/Vergleichsarbeiten
3. Externe Evaluation/Schulinspektion
4. Auswertungen der Schulstatistik
5. Aufgabensammlungen/-beispiele (z. B. vom IQB)
6. Schulfachbezogene Zeitschriften (z. B. Praxis Deutsch)
7. Überfachliche schulbezogene Zeitschriften (z. B. Pädagogik)
8. Bildungsteil von Tages-/Wochenzeitungen oder Magazinen
9. Innerhalb der Schule/des Unterrichts eingesetzte Tests
10. Schülerfeedback zum Unterricht
11. Selbst bzw. von der Schule durchgeführte Befragungen (z. B. mit SEIS Deutschland)
12. Gemeinsame Unterrichtsentwicklungsmaßnahmen (z. B. kollegiale Hospitationen)
13. Schulbezogene Parallelarbeiten innerhalb von Jahrgangsstufen

An sieben Schulen aus Studie 1 und acht Schulen aus Studie 3 wurden (halbstrukturierte) Interviews mit Schulleitungen und Lehrkräften geführt, um Gründe und Motive für eine datengestützte Schulentwicklung oder für die Nicht-Nutzung von vorliegenden Daten zu untersuchen. Die Interviews wurden transkribiert, mit MAXQDA kodiert und kategorisiert.

## Statistische Analysen

Für die Beantwortung der Forschungsfragen wurden deskriptive Analysen und Signifikanztests vorgenommen.

## Ergebnisse

Die Analyse der Fragebogenangaben ergibt, dass schulintern gewonnene Daten, d. h. „weiche“ Daten bzw. Evidenz im weiteren Sinne – beispielsweise gemeinsame Unterrichtsentwicklungsmaßnahmen (z. B. kollegiale Hospitationen), Parallelarbeiten oder Schülerfeedback – nützlicher eingeschätzt werden als Daten, denen schulexterne Konzepte zugrunde liegen, beispielsweise Ergebnisse aus bundesweit durchgeführten Vergleichsarbeiten bzw. Lernstandserhebungen oder aus externen Evaluationen (Schulinspektion), wobei Schulleitungen diesen Informationsquellen mehr Nutzen zubilligen als Lehrkräfte. Demnach scheinen Daten, die auf Grundlage interner Konzepte generiert wurden, Schulleitungen und Lehrkräften eher für die schulischen Bedürfnisse geeignet zu sein, was die Wirksamkeit von z. B. Daten aus Vergleichsarbeiten, zentralen Abschlussprüfungen oder von Resultaten aus Schulinspektionen zumindest fraglich erscheinen lässt. Vielmehr wird die Annahme bestätigt, dass nicht die bloße Existenz schulextern generierter Daten bereits Schul- und Unterrichtsprozesse initiiert.

Der kontrastierende Vergleich verdeutlicht, dass Lehrkräfte an Schulen in herausfordernden Lagen den potenziellen Nutzen der verschiedenen Datenquellen zwar tendenziell höher bewerten, allerdings – außer in Bezug auf Vergleichsarbeiten – eine eher geringere Datennutzung als Lehrkräfte an Schulen in sozial privilegierten Räumen berichten.

Wenngleich die Autorinnen keine Auswertungen dazu vorlegen, lassen die abgebildeten Balkendiagramme erkennen, dass die wahrgenommene Nützlichkeit im Schnitt höher eingeschätzt wird als das Ausmaß der Nutzung. Gründe für diese Diskrepanz und die im Mittel nicht sehr stark ausgeprägte Nutzung der verschiedenen Datenquellen liefern die Auswertungen der Interviews. Sie verdeutlichen die Schwierigkeiten von Schulleitungen und Lehrkräften, externe Evidenz für die eigene Handlungspraxis nutzen zu können. Als hinderliche Faktoren werden Zeitmangel und Interpretations- bzw. Auswertungsschwierigkeiten benannt. Einige Befragte zweifeln zudem an der Qualität externer Daten. Als Bedingungen, die die Datennutzung allgemein begünstigen, werden eine einfache Aufbereitung, die Nützlichkeit zur professionellen Weiterentwicklung und zur positiven Außendarstellung genannt.

## Diskussion und Einschätzung

### Hintergrund

Die Studie von Demski und Racherbäumer greift vor dem Hintergrund der Diskussionen um die Wirksamkeit neuer Steuerungsinstrumente und der Bedingungen für eine datengestützte Schulentwicklung ein für die Administration sehr relevantes Forschungsdesiderat auf. Es werden a) die wahrgenommene Nützlichkeit sowie die selbstberichtete Nutzung von 13 verschiedenen Informationsquellen und b) nutzungsförderliche sowie nutzungshinderliche Faktoren analysiert.

Für die Herausstellung der Relevanz ihrer Studie wird auf die Einführung von Instrumenten der Neuen Steuerung und auf die zunehmende Schulautonomisierung sowie damit verschränkte Annahmen Bezug genommen. Es wird betont, dass die beiden gegenläufigen Steuerungsmechanismen (top down vs. bottom

up) jeweils einen Betrag zur Qualitätssicherung im Mehrebenensystem Schule leisten sollen. Die beiden Autorinnen erläutern für die internationale Leserschaft einleitend die Ausgestaltung dieser beiden Steuerungslogiken im deutschen Bildungskontext. Sie stellen u. a. heraus, dass standardbasierte Testungen – beispielsweise im Gegensatz zu der High-Stakes-Testing-Situation in den Vereinigten Staaten – für Schulen in Deutschland eher geringe Konsequenzen nach sich ziehen. Anschließend legen sie den Forschungsstand dar. Auf dieser Grundlage wird die eigene Fragestellung abgeleitet. Die Argumentationsweise und die Hinführung zur eigenen Studie erscheinen insgesamt stringent. In der Einleitung tauchen einige Redundanzen auf (z. B. der Analysefokus). Eine theoretische Einbettung erfolgt jedoch nicht.

### **Design**

Das Studiendesign und die Durchführung werden ausführlich, aber im Blick auf die eigentlichen Samples bedingt gut bzw. unmittelbar nachvollziehbar benannt. Angaben zu den verwendeten Forschungsinstrumenten werden gegeben. Die Ergebnisdarstellungen beinhalten deskriptive Angaben und Mittelwertvergleiche, die mittels gruppierter Balkendiagramme visualisiert werden, allerdings werden nur vereinzelt Zahlen angegeben. Welche Signifikanztests zum Einsatz kamen, bleibt unerwähnt. Die methodische Vorgehensweise erscheint zur Eruierung dieses noch wenig inspizierten Forschungsfeldes angemessen. Inwiefern demographische Hintergrundvariablen zu den Lehrkräften oder Schulen erhoben und berücksichtigt wurden, bleibt offen.

Ein Kritikpunkt betrifft die Kontrastierung von Schulen in herausfordernden Lagen mit Schulen in privilegierten Sozialräumen. Da erstere allein aus nordrhein-westfälischen Schulen rekrutiert wurden und letztere aus rheinland-pfälzischen, könnten die ermittelten Unterschiede teilweise länderspezifischen Strategien hinsichtlich der einzelnen Datenquellen geschuldet sein; beispielsweise sind die Durchführung von Vergleichsarbeiten bzw. Lernstandserhebungen, die Präsentation der Testergebnisse sowie die damit verbundenen Konsequenzen von Land zu Land unterschiedlich.

### **Ergebnisse**

Die Zielstellung der Untersuchung wird erreicht und die vorgenommenen Schlussfolgerungen erscheinen plausibel. Die Einordnung der Ergebnisse gelingt überzeugend, ist sehr differenziert und mit praktischen Implikationen für die Administration versehen. Es stellt sich die Frage, wie extern generierte Daten so aufbereitet und kommuniziert werden können, dass sie auch wirklich auf Schul- und Unterrichtsebene zur Qualitätsentwicklung genutzt werden. Hierzu bedarf es neuer und bedarfsorientierter Unterstützungssysteme. Als methodische Einschränkung wird die Datenqualität genannt, da die Ergebnisse auf Selbstberichten beruhen. Denn aus anderen Studien ist bekannt, dass diese in der Regel verzerrt sind und nicht die tatsächliche Handlungspraxis abbilden. Es wäre aus Sicht der Rezensenten wünschenswert gewesen, wenn die Ergebnisse zu den Lehrkräften in Abhängigkeit von verschiedenen Fachkulturen und damit verbundenen Berührungspunkten bzw. der praktischen Auseinandersetzung mit Daten im weiteren Sinne (z. B. mathematisch-naturwissenschaftliche vs. sprachliche Fächer) berichtet worden wären.

Es bleibt auch offen, warum trotz der im Forschungsüberblick erwähnten möglichen Einflussfaktoren für die Datennutzung (hier insbesondere Merkmale der Schule, etwa die Schulkultur, und der befragten Lehrkräfte und Schulleitungen, etwa ihre Überzeugungen, ihr Wissen und ihre Erfahrungen) in der Analyse selbst darauf nicht Bezug genommen wurde. Zudem bleibt aufgrund des fehlenden Berichts zur Verteilung demographischer Hintergrundvariablen in den einzelnen Samples sowie aufgrund fehlender Referenzen zur schulamtlichen Statistik unklar, inwiefern die Stichproben untereinander vergleichbar und die Ergebnisse verallgemeinerbar sind.

## Weiterführende Hinweise im Kontext

Demski, D. (2017). *Evidenzbasierte Schulentwicklung. Empirische Analyse eines Steuerungsparadigmas*. Wiesbaden: Springer VS.

## Reflexionsfragen für die Praxis

Nachfolgende Reflexionsfragen sind ein Angebot, die Befunde der rezensierten Studie auf das eigene Handeln als Lehrkraft oder Schulleitungsmitglied zu beziehen und zu überlegen, inwiefern sich Anregungen für die eigene Handlungspraxis ergeben. Die Befunde der rezensierten Studien sind nicht immer generalisierbar, was z. B. in einer begrenzten Stichprobe begründet ist. Aber auch in diesen Fällen können die Ergebnisse interessante Hinweise liefern, um über die eigene pädagogische und schulentwicklerische Praxis zu reflektieren.

### Reflexionsfragen für Lehrkräfte:

- Welche der im Artikel genannten Datenquellen halte ich für nützlich (z. B. Schülerfeedback, kollegiale Hospitation, Parallelarbeiten, Vergleichsarbeiten, vgl. Abschnitt Design)? Welche nutze ich systematisch und gewinnbringend? Woran scheitert bei mir ggf. eine (systematische) Datennutzung?
- Welche Unterstützungsangebote der Bildungsadministration oder von anderen Akteuren kenne ich, um die genannten Datenquellen zu nutzen (z. B. Instrumente zur Erhebung von Schülerfeedback, Rückmeldeportale zu den Vergleichsarbeiten, Aufgabensammlung des IQB, Hospitationsschulen)?
- Welche Daten nutzen meine Kolleginnen und Kollegen und wie gehen sie dabei vor? In welchen Zusammenhängen ist ein Austausch darüber zweckmäßig? Wie können wir schulinterne Daten systematisch generieren und es angesichts des Zeitmangels, der Schwierigkeiten bei der Interpretation und Rekontextualisierung schaffen, externe Daten stärker für die Schul- und Unterrichtsentwicklung zu nutzen?

### Reflexionsfragen für Schulleitungen:

- Der Schulleitung kommt im Hinblick auf die Nutzung von Daten eine bedeutsame Rolle zu. Welche der folgenden Punkte, die für Schulen mit mehr Datennutzung in den Interviews nachgewiesen wurden, könnte ich stärken?
  - Die Schulleitung unterstützt die Entwicklung einer kooperativen und familiären Atmosphäre.
  - Die Schulleitung verteilt Leitungsaufgaben und Führungsverantwortung und delegiert Verantwortlichkeiten an Arbeitsgruppen.
  - Die Schulleitung verfügt über gute Fähigkeiten im Umgang mit Daten.
  - Die Schulleitung wirkt als Mentor für datengestützte Schulentwicklung und macht deutlich, dass sie der Nutzung von Daten einen großen Wert zumisst.
  - Die Schulleitung vertritt überzeugend den Standpunkt, dass ein ggf. zusätzlicher Arbeitsaufwand lohnenswert ist.
  - Die Schulleitung vertritt überzeugend den Standpunkt, dass externe Daten nicht der

Kontrolle durch die Schulaufsicht dienen, sondern eine Möglichkeit zu organisationalem Lernen darstellen und die Chance bieten, die Qualität der Schule nach außen sichtbar zu machen.

- Welche der im Artikel genannten Datenquellen werden von den Lehrkräften meiner Schule als nützlich eingeschätzt und systematisch genutzt (z. B. Schülerfeedback, kollegiale Hospitation, Parallelarbeiten, Vergleichsarbeiten, vgl. Abschnitt Design)? Woran scheitert ggf. die Datennutzung? Welche Einstellungen zur Nutzung interner und extern generierter Daten dominieren im Kollegium?
- Welche Datenquellen kann ich selber nutzen, um mögliche Handlungserfordernisse an meiner Schule zu identifizieren und zu belegen? Welche Indikatoren können dazu dienen, Veränderungen nachzuvollziehen? Welche Daten liefern mir möglicherweise Argumente, um die Anliegen meiner Schule gegenüber Dritten zu begründen?

## Rezensent/-in

Dr. Mirko Krüger, PD, Lehrer an der Georg-Müller-Gesamtschule in Wetter (Ruhr) und Lehrbeauftragter an der Fakultät für Bildungswissenschaften, Universität Duisburg-Essen. Arbeitsschwerpunkte: Schul- und Schulsportentwicklung, Sprachbildung im Sportunterricht, Professionalisierung von Lehrkräften

Dr. Johannes Rosendahl, Dipl. Psych., Referent an der Qualitäts- und Unterstützungsagentur – Landesinstitut für Schule (QUA-LiS NRW), Soest. Arbeitsschwerpunkte: Bildungsmonitoring und Evaluation zentraler Prüfungen, Kooperation mit Wissenschaft zur Gewinnung und Erschließung empirischer Evidenzen

## Zitiervorschlag

Krüger, M., Rosendahl, J. (2019). Rezension zu Demski, D. & Racherbäumer, K. (2017). What data do practitioners use and why? Evidence from Germany comparing schools in different contexts. *Nordic Journal of Studies in Educational Policy*, 3(1), 82–94. DOI: 10.1080/20020317.2017.1320934. *Forschungsmonitor Schule*, 52. Abgerufen von <https://www.forschungsmonitor-schule.de/print.php?id=66>

## Urheberrecht

Dieser Text steht unter der [CC BY-NC-ND 4.0 Lizenz](#). Der Name des Urhebers / der Urheberin soll bei einer Weiterverwendung wie folgt genannt werden: Mirko Krüger, Johannes Rosendahl (2019) für den [Forschungsmonitor Schule](#).