

## Hans Brügelmann: Vermessene Schulen – standardisierte Schüler

Rezensiert von Andreas Vohns



Hans Brügelmann, Autor des hier zu rezensierenden Titels, war von 1993–2012 Professor für Erziehungswissenschaft mit Schwerpunkt Grundschulpädagogik und -didaktik an der Universität Siegen, nahezu zeitgleich mit dem den Leser(inne)n der MGDm vermutlich vertraueren Hans-

Werner Heymann, der dort die Professur für Erziehungswissenschaft mit dem Schwerpunkt Schulpädagogik und Didaktik inne hatte<sup>1</sup>. Die beiden eint neben der gemeinsamen Arbeit am Projekt „LISA & KO – Lernbiographien im schulischen und außerschulischen Kontext“ auch der Umstand, dass beide Erziehungswissenschaftler mit ausgeprägten fachdidaktischen Interessenbereichen sind – in Heymanns Fall bekanntlich in der Mathematikdidaktik, in Brügelmanns Fall in der Deutschdidaktik, genauer: im Bereich des Schriftspracherwerbs.

Seit Erscheinen von „Vermessene Schulen – standardisierte Schüler“ eint die beiden zudem neben der 2012 erfolgten Emeritierung auch noch der Umstand, dass beide es geschafft haben, im Zusammenhang mit ihren Arbeiten äußerst prominent Eingang in die Schlagzeilen der Zeitung mit den vier großen Buchstaben gefunden zu haben. Brügelmann hat es dabei in die Sonntagsausgabe vom 13.9.2015 geschafft, mit der Schlagzeile „Schulnoten sind überflüssig!“, die als Zusammenfassung seines Buches „Vermessene Schulen – standardisierte Schüler“ allerdings noch stärker verkürzt, als es „Zu viel Mathe ist Quatsch!“ für Heymanns „Allgemeinbildung und Mathematikunterricht“ tat.

Der Untertitel des Buches: „Zu Risiken und Nebenwirkungen von PISA, Hattie, VerA und Co.“ gibt da schon treffender die Stoßrichtung vor: Brügelmann legt gewissermaßen einen „Beipackzettel“ zur quantitativen empirischen Bildungsforschung vor, insbesondere zu jenen Spielarten, die

sich als Evaluations- oder Wirkungsforschung mit dem Ziel der Generierung „harter“ Evidenz für sogenannte „evidenzbasierte Bildungspolitik“ generieren. Brügelmann richtet sich mit seinem Buch zunächst an diejenigen, die man neudeutsch als „Stakeholder“ einer solchen Forschung bezeichnen würde, also „Interessierte in Politik, Medien, in den Schulen und Familien“ – auch und gerade mit dem Ziel, dass diese sich gegenüber unangemessenen Geltungsansprüchen „evidenzbasierter“ Urteile emanzipieren mögen. Brügelmann hofft überdies, zur Selbstreflexion derjenigen seiner Kolleg(inn)en beizutragen, die in die empirische Bildungsforschung selbst als Akteur(inn)e(n) involviert sind.

Nach einem grundsätzlichen Problemaufriss „Über das Spiel mit Zahlen hinaus – Grundprobleme einer ‚Evidenzbasierung‘“ widmet sich Brügelmann in vier Hauptkapiteln nacheinander der Meta-Studie von Hattie, PISA bzw. allgemeiner internationalen Vergleichsstudien, VerA und (was überraschen mag, aber den BAMS-Aufmacher erklärt) dem Problem der Leistungserfassung, -rückmeldung und Notengebung (Brügelmann ist als vehementer Gegner von Schulnoten bekannt, insofern trifft die Schlagzeile, das Thema im engeren kommt aber im Buch allenfalls am Rande vor). Das Buch kulminiert in zehn abschließend formulierten Thesen zur kritischen Diskussion empirischer Bildungsforschung bzw. von Evaluation im pädagogischen Bereich. Abgerundet wird das Buch durch eine kommentierte Bibliographie mit Leseempfehlungen<sup>2</sup> und ein knapp gehaltenes statistisches Glossar.

Die vier Hauptkapitel des Buches verlagern jeweils den Blickwinkel in Richtung eines der bereits erwähnten „Stakeholder“: Hatties Wert für die Bildungsforschung wird hinterfragt, PISAs möglicher Beitrag zur Steuerung von Bildungspolitik, VerAs konkreter Nutzen für die Schul- und Unterrichtsentwicklung vor Ort und schließlich die Frage angemessener Evidenzsammlung für die Rückmeldung des Lernfortschritts aus Sicht der Lernenden.

<sup>1</sup> Im Sinne der vollständigen Offenlegung: Der Rezensent selbst studierte von 1995–2002 an derselben Universität Lehramt und promovierte dort von 2000–2007.

<sup>2</sup> Ein Teil der empfohlenen Texte steht als Onlinematerial zur Verfügung unter: [https://www.beltz.de/fachmedien/paedagogik/buecher/produkt\\_produktdetails/15130-vermessene\\_schulen\\_standardisierte\\_schueler.html](https://www.beltz.de/fachmedien/paedagogik/buecher/produkt_produktdetails/15130-vermessene_schulen_standardisierte_schueler.html)

Brügelmann gibt an, die einzelnen Kapitel seien für sich genommen lesbar, diesen Eindruck kann ich bestätigen, hatte aber beim Lesen dennoch nicht den Eindruck, dass der Anspruch der getrennten Lesbarkeit der Abschnitte die Lektüre des Gesamtwerks aufgrund unvermeidbarer Redundanzen spröde werden lässt. Dem Buch ist kaum anzumerken, zumindest an keiner Stelle anzulasten, dass sich Brügelmann bei dessen Zusammenstellung einer Reihe eigener, kürzerer Texte aus dem Zeitraum von 1978 bis 2015 als Fundus bedient hat.

Dem Problemaufriss zur Evidenzbasierung vorangestellt hat Brügelmann gewissermaßen gleich zwei Einleitungen, zunächst ein „Vorspiel“ in welchem er die Leser(innen) in das übergreifende Thema Evaluation im pädagogischen Bereich einführt. Er unterscheidet dabei zunächst drei Archetypen von Evaluatoren, denen er jeweils entsprechende Pendanten aus dem Schul- und Bildungsbereich gegenüberstellt: den „Warentester im Bereich der Produktionskontrolle“ (Pendant: „den technisch Messenden in den internationalen Leistungsstudien“), den „Kritiker in Kunst, Musik, Theater und Literatur“ (Pendant: „Gutachter bei der Lehrmittelzulassung“) und den „Richter im juristischen Prozess“ (Pendant: den „Prüfer in Examina“). Brügelmanns Hauptbotschaft ist nun, dass das, was „für das Gelingen von Lernen wesentlich ist“, sich gerade nicht auf dieselbe Art und Weise erfassen und beurteilen lässt, „wie man das Funktionieren technischer Geräte oder ökonomischen Erfolg misst“ (S. 12). Für Brügelmann bedeutet das im Umkehrzug nicht, dass Schule und Lernen keiner Evaluation bedürften, er postuliert aber, dass Didaktik und Pädagogik „andere Formen der Untersuchung und Bewertung von Leistung benötigen“. Brügelmann plädiert in seinem Vorspiel dann für eine *demokratische Evaluation*, die ihre Ziele und Kriterien weder allein der Bildungspolitik („bürokratische Evaluation“) noch der akademischen Sphäre („autokratische Evaluation“) verdankt, sondern zwischen den Interessen von Auftraggebern in der Bildungspolitik, Experten aus den einschlägigen Wissenschaften und den Betroffenen vermittelt und deren unterschiedliche Sichtweisen integriert. Ziel einer solchen Evaluation sei die Gewinnung „handlungsbedeutsamer Information“. Wer sich ein wenig mit Aktionsforschung, insbesondere mit Donald A. Schön und dessen *Reflective Practitioner* auseinandergesetzt hat, wird hier viele Parallelen erkennen, obwohl Brügelmann diesen theoretischen Bezug nicht explizit herstellt.

Bevor es ans Eingemachte geht, bietet Brügelmann seinen Leser(inne)n mit „Wozu Evaluation? Inszenierte Kontroverse in verteilten Rollen“ die

zweite, spielerische Einführung in das Themenfeld des Buches: Hier kommen im fiktiven Dialog Arbeitgeber(innen), Lehrer(innen), Eltern, Schulaufsicht und Bildungsforschung zusammen und streiten über Sinn und Unsinn von PISA, VerA, Hattie, Notengebung und Co. Obwohl ich eigentlich kein großer Freund solcher Inszenierungen bin, muss ich einräumen, dass diese Seiten in derart beeindruckender Weise und mit allenfalls geringen Überzeichnungen einen Einstieg in das Thema liefern, dass man sich in jeder Einführung in die Fachdidaktik oder Pädagogik eigentlich die Zeit und Mühe nehmen sollte, einen Teil dieser Inszenierung mit verteilten Rollen aufzuführen.

In „Grundprobleme einer ‚Evidenzbasierung‘“ stützt Brügelmann seine These einer notwendig nicht im Sinne technischer Rationalität umsetzbaren Evaluation schulischer Lern- und Bildungsprozesse argumentativ ab. Dreh- und Angelpunkt seiner Argumentation ist dabei, dass „Großstudien mit standardisierten Tests“ aufgrund der ihnen eigenen Forschungslogik bei der Evaluation von Alternativen (Brügelmann nutzt als Beispiel verschiedene Methoden des Schriftspracherwerbs) immer nur den durchschnittlich größeren Erfolg einer Alternative gegenüber der anderen nachweisen können. Aber: Wenn im Durchschnitt z. B. die von Brügelmann favorisierte Methode „Lesen durch Schreiben“ für den Schriftspracherwerb sich als weniger erfolgreich herausstellt, als eine andere Methode, inwiefern kann dieses Ergebnis dann die Erfahrungen und größeren Erfolge, die einzelne Lehrer(innen) mit dieser Methode erzielt haben, überhaupt grundsätzlich entwerten? Brügelmann bestreitet dies in wenigstens zweifacher Hinsicht:

- Zum einen problematisiert er die Instrumente der Evaluation (also standardisierte Tests).
- Zum anderen wirft er die Frage auf, inwiefern die vielen Fälle, die quantitative empirische Bildungsforschung zwangsläufig benötigt, überhaupt in sich hinreichend homogene Realisierungen der jeweiligen Alternativen darstellen, um aussagekräftige Mittelwertbildungen zu erlauben.

Kritische Einwürfe im Sinne des ersten Punktes hat es im Umfeld von TIMSS, PISA und VerA auch in der Mathematikdidaktik nicht wenige gegeben. Brügelmann fokussiert hier auf das Problem, das quantitative empirische Bildungsforschung ein Streben nach *Eindeutigkeit von Daten* mit sich bringt, dass in Widerspruch zum Wunsch nach *bedeutungsvoller Information* geraten kann. Er lässt dazu fünf fiktive Schüler das Wort KINO schreiben:

- |         |          |
|---------|----------|
| 1. KINO | 4. KINNO |
| 2. KINO | 5. KIENO |
| 3. KINO |          |

Im Sinne eines standardisierten Tests wird man drei richtige und zwei falsche Lösungen konstatieren. Brügelmann gibt dann allerdings fünf fiktive Lernbeobachtungen an, die den Schluss zulassen, dass sich hinter der vermeintlich korrekten Schreibweise ebenso wie hinter den falschen völlig unterschiedliche Leistungsstände im Schriftspracherwerb verbergen können, wobei auch Falschreibungen für ein höheres Sprachverständnis stehen können, als die Richtigschreibung.

Der zweite Punkt, also die Frage, inwiefern Mittelwertbildungen über große Gruppen von nach unterschiedlichen pädagogischen Programmen unterrichteten Lernenden überhaupt aussagekräftig sind und konfligierende individuelle Evidenz entkräften können, motiviert Brügelmann zunächst durch eigene biographisch-episodische Evidenz: Seinem Eindruck nach lassen pädagogische Programme wie etwa „Lesen durch Schreiben“, ja überhaupt alle curricularen Entwürfe bis hin zu in kleinsten Schritten ausgearbeiteten Lehrprogrammen, immer noch vergleichsweise große Handlungsspielräume für Lehrende und zeitigen in unterschiedlichen Kontexten multiple Wechselwirkungen mit nicht im Programm erfassten Aspekten der Lehr-Lernsituation. In der Folge kann in sich an ein und demselben Lehrprogramm orientierenden Schulklassen immer noch vollkommen unterschiedlicher Unterricht beobachtet werden. Bekanntlich werden nun in hinreichend großen Untersuchungsgruppen immer kleinere Leistungsmittelwertunterschiede signifikant und dann nicht selten bereits als bedeutsam interpretiert, ohne Leistungsstreuungen in den Gruppen zu berücksichtigen. Brügelmann hält den Übergang zu Effektstärken, die Überlappungen und Streuung der Leistungsdaten berücksichtigen (und in der Hattie-Studie zuhauf berichtet werden), für eine technokratische Scheinlösung, da auch bei ihnen nicht berücksichtigt wird, dass eben nicht nur die Leistungen streuen, sondern auch die praktischen Realisierungen der Programme als solche. Das stellt allerdings die Evaluierbarkeit von relativ breiten Konzepten wie „offener Unterricht“ oder „Computereinsatz“ (zumindest in Feld- und Metastudien) ganz grundsätzlich in Frage. Unterrichtsmethoden, so Brügelmanns Schlussfolgerung (in meinen eigenen Worten), lassen sich eben nicht unabhängig von den Menschen, die sie anwenden, und unabhängig von den Menschen, auf die sie angewendet werden, und unabhängig von den Kontexten evaluieren, in denen dies geschieht. Das setzt empirischer Bildungsforschung Grenzen in der Erzielung eindeutiger Information und lässt Meta-Studien im pädagogischen Bereich als grundsätzlich hinterfragenswert erscheinen.

Die Unterüberschrift „Warum wir PISA & Co brauchen, die Ergebnisse aber nur mit Vorsicht nutzen sollten“ des PISA-Kapitels deutet bereits an, dass Brügelmann mit den internationalen Leistungsvergleichsstudien etwas weniger hart ins Gericht geht, als mit der Hattie-Studie. Brügelmann wirft hier die Fragen auf, inwiefern Fachleistungen einzelner Fächer schon ein Maßstab für die Qualität des Schulwesens als solches darstellen können, ob die Testaufgaben geeignet sind, Fachleistungen in angemessener Form zu erfassen, wie aussagekräftig und verständlich die gewählten Auswertungen der Daten sind und inwiefern Verbesserungen in den Testwerten unumwunden als Verbesserungen der zu Grunde liegenden Fachleistungen bzw. den Kreis schließend für die Qualität des Bildungssystems genommen werden sollten. In diesem Kapitel werden diejenigen, die sich mit der PISA-Kritik aus dem mathematikdidaktischen Bereich bereits intensiver beschäftigt haben, viele Argumentationslinien wiedererkennen. Brügelmann ist dabei ein durchgängig unaufgegrerter, um sorgfältige Abwägung von Pro- und Contra-Argumenten bemühter Stil zu attestieren, dem es sehr deutlich um eine produktive Weiterentwicklung dieser Unternehmungen geht, nicht um deren grundsätzliche Verdammung. Was Brügelmann aber m. E. deutlich macht, ist, dass eine produktive Weiterentwicklung voraussetzt, dass die Proponenten dieser Art von Forschung zu einer deutlich bescheideneren Einschätzung bezüglich der Geltungsansprüche ihrer Ergebnisse gelangen. Ich würde Brügelmanns Überlegungen wie folgt zusammenfassen: Forscher(innen), die mit notwendigen Grenzen der mit ihren Untersuchungsmethoden erzielten bzw. erzielbaren Ergebnisse offen und ehrlich umgehen und andere Arten von Forschung als komplementäre Zugänge zu schätzen wissen, wären deutlich produktiver für Pädagogik und Didaktik, als solche, die eine Art Alleinvertretungsanspruch „echter“ Forschung für ihre Untersuchungsmethoden proklamieren und in dem Glauben verharren, jedes Problem und jede Grenze des eigenen Zugangs sei nur dessen bisheriger mangelnder Entwicklung geschuldet, temporär und durch rein technokratische Optimierungen schon noch in den Griff zu bekommen.

Die letzten beiden Kapitel, die den Beitrag von VerA zur Unterrichtsentwicklung und die Frage des Beitrags unterschiedlicher Instrumente zur Lernstandserhebung und -rückmeldung für Schüler(innen) behandeln, richten sich noch deutlicher als die hier bereits besprochenen Abschnitte insbesondere an Schulpraktiker(innen) und sind nicht minder interessant und kenntnisreich geschrieben. Ich verzichte hier im Sinne der gebotenen Kürze auf eine eingehendere Darstellung.

Dieses Buch kann uneingeschränkt all jenen empfohlen werden, die bereit sind, sich (selbst-)kritisch mit Nutzen und Grenzen quantitativer empirischer Bildungsforschung auseinanderzusetzen, auch empathisch im Sinne der verschiedenen „Stakeholder“. Es sei ganz besonders all denjenigen empfohlen, die (zu Recht oder Unrecht sei dahingestellt) bislang den Eindruck hatten, beim Pro-Kontra gäbe es in diesem Bereich nur Maximalpositionen, verhärtete Fronten, Eitelkeiten verschiedener Schulen und unproduktive Pauschal- und Fundamentalkritik. Diese wird man bei Brügelmann nicht finden und es wäre der Mathematikdidaktik sehr zu wünschen, wenn sie ähnlich multiperspektivisch, aufgeschlossen und mit kla-

rem Blick auf die Schul- und Unterrichtsentwicklung die Möglichkeiten und Grenzen fachdidaktischer Wirkungsforschung im Sinne empirischer Bildungsforschung und die Bedeutung alternativer Forschungszugänge zum Lehren und Lernen von Mathematik auszuloten bereit wäre.

Hans Brügelmann: *Vermessene Schulen – standardisierte Schüler. Zu Risiken und Nebenwirkungen von PISA, Hattie, VerA und Co.* Weinheim und Basel: Beltz, 2015. Gedruckt: ISBN 978-3-407-25729-1, EUR 19,95, eBook: ISBN 978-3-407-29204-9, EUR 18,99.

Andreas Vohns, Alpen-Adria-Universität Klagenfurt, Sterneckstraße 15, 9020 Klagenfurt, Österreich, Email: andreas.vohns@aau.at