

Testen und Bewerten

Bibliografie:

Friederike Thole: Rezension zu
Reh, S., Bühler, P., Hofmann, M. & Moser, V. (2021).
Schülersauslese, schulische Bewertung und
Schülertests 1880-1980.
Verlag Julius Klinkhardt. 277 Seiten.
ISBN 978-3-7815-2458-3.
journal für lehrerInnenbildung, 22 (1), 90-95.
<https://doi.org/10.35468/jlb-01-2022-rez>

Gesamtausgabe online unter:

<http://www.jlb-journallehrerinnenbildung.net>
<https://doi.org/10.35468/jlb-01-2022>

ISSN 2629-4982

journal für lehrerInnenbildung
j l b
no.1
2022

Reh, S., Bühler, P., Hofmann, M. & Moser, V. (2021). *Schülersauslese, schulische Bewertung und Schülertests 1880-1980*. Verlag Julius Klinkhardt. 277 Seiten. ISBN 978-3-7815-2458-3

Verschiedene Formen der Begutachtung, der Testung und der Leistungsmessung sind wesentliche Aufgaben, die die Institution Schule erfüllen muss. Auch wenn die Praktiken, die dort zum Tragen kommen, fortlaufender Kritik ausgesetzt sind, sind sie doch konstitutiv für die uns bekannten Formen schulischer Bildung.

Diskussionen und kontroverse Debatten um Formen des *Prüfens* und *Bewertens* sind nicht nur heute – insbesondere vor dem Hintergrund der Bemühungen um Inklusion – hochaktuell, sondern lassen sich schon in den historischen Diskursen der letzten Jahrhunderte finden. Durch die wachsenden Versuche der Beschulung von Schüler*innen, die von der erwarteten Leistungsnorm abweichen, intensivierten sich diese Debatten. Sie umfassen, heute wie damals, nicht nur Praktiken der fachlichen Leistungsbewertung, sondern auch Praktiken der Diagnostik und damit verbundene Prozesse von In- und Exklusion.

Der Band *Schülersauslese, schulische Bewertung und Schülertests 1880-1980*, herausgegeben von Sabine Reh, Patrick Bühler, Michèle Hofmann und Vera Moser, widmet sich aus bildungshistorischer Perspektive den unterschiedlichsten Aspekten pädagogischer und schulpolitischer Debatten um Praktiken des Beobachtens, Prüfens und Beurteilens von Schüler*innen und den mit diesen verbundenen Verfahren und ihren Folgen.

Mit seinen insgesamt 15 Beiträgen gibt der Band Einblicke in unterschiedlichste Dimensionen der Diskussionen des vergangenen Jahrhunderts und beleuchtet, wie die Herausgeber*innen in der Einleitung formulieren, das „[...] entstehende Zusammenspiel zwischen der Etablierung und Differenzierung institutionalisierter Bildungsangebote, den Praktiken der Beobachtung, Prüfung, Testung und Begutachtung sowie den Diskursen über Begabung, Intelligenz und Leistung.“ (S. 9) So widmen sich mehrere Beiträge den Praktiken und Begründungsmustern des Ausschlusses von Kindern, die als „schwachbegabt“ (z. B. S. 65), „bildungsunfähig“, „anormal“, „idiotisch“ (z. B. S. 67) oder „schwachsinnig“ (z. B. S. 241) bezeichnet wurden, und den damit verbundenen Bemühungen um den Ausbau eines Hilfsschulsystems und den sich immer weiter ausdifferenzierenden Praktiken der Diagnostik.

Patrick Bühler greift diese Thematik auf mit dem Beitrag *Komplett pessimistisch eingestellt Hilfe und Heilung in der Schweizer Sonderpädagogik zu Beginn des 20. Jahrhunderts*, Michèle Hofmann mit seinem Beitrag *Grenzziehungen – Praktiken der Kategorisierung geistig anormaler Kinder um 1900 in der Schweiz*, Nadja Wenger mit ihrem Beitrag *Ihr gebt mich fort, weil ihr mich nicht gern habt. Gutachten der St. Galler Fürsorgestelle für Annormale in den 1940er Jahren* und auch Jona Garz, Vera Moser und Stefan Wünsch unter dem Titel *Die Kielhorn-Rede: Ursprungsmythos der deutschen Hilfsschulen*, sowie Michaela Vogt im letzten Beitrag des Bandes über *Das Hilfsschulnahmeverfahren als Grenzzone der Schülerschulung in BRD und DDR*.

Bühler und Hofmann nehmen zu Anfang ihrer Beiträge jeweils Bezug auf den Beginn des sukzessiven Ausbaus eines Hilfsschulsystems in der Schweiz. Bühler betont hierbei insbesondere die Institutionalisierung der Heilpädagogik und geht dann weiter auf die schon damals existierende Problematik des eher medizinisch verorteten Begriffs des „Heilens“ ein (vgl. S. 87). Hofmann konzentriert sich in seinem Beitrag auf die Ausdifferenzierung verschiedener Beschulungs- oder auch Verwahrungskonzepte für von der Norm abweichende Kinder und Jugendliche und erläutert dabei dezidiert die bis heute existierende Problematik der Herstellung einer Norm, von welcher dann eine Abweichung konstruiert wird.

Garz, Moser und Wünsch bereichern die Debatte um die Gründung von Hilfsschulen mit einem Beitrag, der die bisherige Historiographie des Sonderschulwesens kritisch beleuchtet. Die bekannten Erklärungsmuster der Gründung von Hilfsschulen, insbesondere die wachsende Industrialisierung und eine fortschrittsoptimistische Form der Geschichtsschreibung, werden durch die Autor*innen erweitert um „[...] ein vielschichtiges Netzwerk von zeitgenössischen Wissensordnungen und protestantischen Verheißungen [...]“ (S. 42). Sie kommen zu dem Schluss, dass pädagogische Erwägungen kaum eine Rolle spielten und psychiatrische und christliche Deutungsmuster die relevanten Lieferanten der neuen Wissensbestände waren, die zur Konstruktion von Hilfsschulen und somit auch dem Hilfsschulkind beitrugen (vgl. S. 42). Zentral sind in diesen Beiträgen immer wieder die Diskussionen um Verfahren der Feststellung von Abweichung sowie der anschließende Umgang der besonderen Beschulung – Fragen, die auch im heutigen Schulsystem trotz aller inklusiven Bestrebungen als nicht endgültig beantwortet betrachtet werden können.

Eine besondere Praktik der *Testung* beschreibt Jona Garz in seinem Beitrag *Schriftproben von schwachsinnigen resp. Idiotischen Kindern. Testwissen zwischen Psychatrie und Pädagogik um 1900*. Garz illustriert in diesem Beitrag, wie Schreiben als ein Diagnoseinstrument genutzt wurde, wobei insbesondere die Anordnung der Buchstaben und weniger der Inhalt zählte (vgl. S. 52). Die Handschrift wurde dabei als „materialisierter Ausdruck der Störung des Gehirns“ (S. 52) betrachtet. Mit Hilfe der Analyse von Schriftproben wurde davon ausgegangen so zwischen „verschiedenen Arten des Schwachsinnns zu differenzieren und den Schweregrad der Idiotie“ messen zu können. Diese Form des Tests, der als ein objektives Diagnosekriterium (vgl. S. 58) galt, zeigt exemplarisch, welche hohen Stellenwert standardisierte Verfahren hatten und der Beitrag verweist auch explizit darauf, welchen Erfolg diese auch in Teilen bis heute im Schulsystem haben.

Anschlussfähig an die immer wieder aufkeimenden Debatten um die Möglichkeit einer neutralen oder objektiven Bewertung, insbesondere von Aufsätzen, zeigt sich Kerrin von Engelhardts Beitrag *Der papierenen Drache – Der Reifeprüfungsaufsatz zwischen 1890 und 1930*. Von Engelhardt beschreibt hier die in der Praktik der Reifeprüfungsaufsätze zu beobachtende „hybride Struktur“ (S. 172) der Bewertung, bei der sowohl fachspezifische Kriterien herangezogen wurden als auch übergeordnete, wie beispielsweise ein Ermessensspielraum, der sich aus weiterführendem Wissen der Fachlehrer*in über die zu bewertenden Schüler*innen ergab. Die Form des Changierens zwischen übergeordneten und individuellen Bewertungsmustern und die von Von Engelhardt aufgegriffene Thematik des Problems der individuellen Präferenzen und Sympathien, wird somit als schulische Herausforderung erkennbar, mit der sich pädagogisches Fachpersonal nicht erst in den letzten Jahrzehnten konfrontiert sah.

Einen bereichernden Blick über die Grenzen des deutschsprachigen Raums hinaus gibt Fanny Isensee in ihrem Beitrag *Intelligence tests were given in order to obtain a basis for classifying the pupils – Die reclassification Projects in New York City in den 1920er-Jahren*. Anhand des Beispiels Brooklyn diskutiert Isensee die Nutzung von Intelligenztests zur Einteilung von Schüler*innen in bestimmte Klassenstufen. Besonders aufschlussreich und neu ist Isensees Teilfragestellung, inwiefern die Ergebnisse aus solchen Tests auch Rückwirkung auf die Institution Schule und ihre Ausgestaltung hatten. Isensee kommt zu dem Ergebnis, dass Resultate von Intelligenztests auch genutzt wur-

den um die Schulen effizienter und wirtschaftlicher (vgl. S. 239) zu gestalten. Isensees Analysen zeigen zudem, was kaum verwundern mag, aber doch wieder ein Augenmerk auf die Beharrlichkeit solcher und ähnlich gelagerter Annahmen richtet, dass Schüler*innen mit Migrationshintergrund immer mit dem Vorwurf konfrontiert waren, schlechtere Schulleistungen zu zeigen und somit auch schlechter in IQ-Tests abzuschneiden (vgl. S. 238).

Neben diesen diversen Beiträgen, die sich eher der Problematik des Umgangs mit *negativ* von der Leistungsnorm abweichenden Schüler*innen befassen, greift Susanne Schregel in ihrem Beitrag *Übernormalen Pädagogik und Begabenschulen zwischen Kaiserreich und Weimarer Republik* die Thematik um Hochbegabte auf. Zentrales Thema des Beitrages sind Diskussionen um die Überführung von Hochbegabten in „Klassen für Übernormale“, „Sonderschulen für hervorragend Befähigte“ oder „Genie-Schulen“ (S. 135). Hierbei stellt sie durch die Analyse von zeitgenössischen pädagogischen und psychologischen Fachpublikationen heraus, dass die Verbreitung der Intelligenzdiagnostischen Verfahren vermehrt dazu führte, dass sowohl positive als auch negative Abweichung als fast gleichartige „Rand“-Phänomene betrachtet wurden (vgl. S. 147). Somit zeigt sie, anknüpfend an auch heute viel rezipierte kritische Positionen, dass diagnostische Verfahren die Gefahr bergen, eine künstliche Leistungsnormalität zu konstruieren, die jegliche Form der Abweichung als unnormal und als potenziell störend kennzeichnet.

Dass sich psychologische Tests und Diskussionen um Normalität und Abweichung nicht nur darauf beschränkten *innerhalb* des Schulsystems homogene Leistungsgruppen herzustellen oder Schüler*innen eine als angemessen angesehene Förderung zukommen zu lassen, zeigt der Beitrag *Das Kind vor verfehlter Wahl geschützt. Die Einführung eines berufspsychologischen Schülerbeobachtungsbogens in Berliner Schulen, 1917-1923* von Johanna Lerch. Sie erläutert, wie in der Weimarer Republik die Verbesserung der Berufswahl jugendlicher Schulabgänger*innen zu einem politischen Thema wurde und große Erwartung in psychotechnische diagnostische Verfahren gesetzt wurde, um potenzielle Arbeitskräfte ökonomisch nutzbringend auf verschiedenen Wirtschaftszweige zu verteilen (vgl. S. 209). Schon damals standen einige Pädagog*innen und Psycholog*innen solchen Tests kritisch gegenüber und stufen diese als „ungenügend, oberflächlich und zufällig“ ein (S. 211), weshalb sie alternative Formen der Beob-

achtung vorschlugen. Anhand des Vergleiches verschiedener Beobachtungsbögen zeigt Johanna Lerch auf, „[...] dass die Abwendung von der experimentellen Psychologie hin zur Typenlehre und später zum psychologischen Test [...] sich in stetiger Auseinandersetzung mit ihren Anwender*innen weiterentwickelte.“ (S. 222) Diese Ergebnisse werfen die Frage auf, inwiefern auch heute Lehrkräfte als Anwender*innen von Tests die Ausgestaltung dieser mitbestimmen und somit als relevante Akteur*innen im Diskurs um Diagnoseverfahren angesehen werden müssen.

Die hier exemplarisch kurz vorgestellten Beiträge zeigen schon die Vielfältigkeit der Möglichkeiten, sich aus historischer Perspektive dem Phänomen der Schülersauslese und damit verbundenen Beurteilungs- und Testpraktiken zu nähern auf und die diskursiven Parallelen zu heutigen Debatten scheinen eklatant.

Obwohl einige Beiträge des Bandes weniger bekannte Praktiken von Leistungsbewertung und Tests aufgreifen, setzt er das Thema der Intelligenztests und die mit diesen verknüpften diskursiven Verstrickungen sehr zentral, sodass sich dieses im Vergleich zu anderen Praktiken der Diagnose und Leistungsmessung etwas überrepräsentiert zeigen. Weitere tiefere Einblicke in die Umsetzung der spezifisch methodisch unterschiedlichen Verfahren, wie sie beispielsweise der Beitrag von Garz zu Schriftproben oder von Von Engelhardt zum Reifeprüfungsaufsatz geben, wären ergänzend sehr lohnenswert gewesen.

Nichtsdestotrotz zeichnen sich die im Band versammelten Beiträge, sowohl hinsichtlich der zeitlichen Zuschnitte als der gewählten thematischen Einblicke durch eine große Heterogenität aus – ohne dabei einen roten Faden vermissen zu lassen. Der Band schafft es darüber hinaus, insbesondere für Leser*innen, die über aktuelle Debatten um das Schulsystem informiert sind, anhand von unterschiedlichen Aspekten die lange zurückreichende Kontinuität um ein Streben nach einer Homogenisierung von Lerngruppen (vgl. S. 101) zu illustrieren. Das Zusammenspiel von Inklusion als gesellschaftlicher Aufgabe und die Bearbeitung dieser mit schulisch oder weiter gefassten pädagogischen Mitteln, ist dabei immer mit zentral gesetzt.

In besonderem Maße zeigt sich aber auch immer wieder, dass konträre Positionen und Aushandlungsprozesse rund um Praktiken der Messung, Diagnostik, Prüfung und Bewertung, welche eng mit Ein- und Ausschlussprozessen verbunden waren und sind, existierten. Zudem

werden auch Ambiguitäten und Paradoxien rund um exkludierende, integrierende oder inkludierende schulische und pädagogische Prozesse aufgezeigt, in denen sich immer wieder fragmentarische Parallelen zu heutigen Debatten um das Inklusionsparadoxon oder auch das Etikettierungs- und Ressourcendilemma wiederfinden, was den Band auch für eine nicht genuin bildungshistorisch orientierte Leser*innenschaft äußerst interessant macht.

Friederike Thole, wiss. Mitarbeiterin
am Institut für Erziehungswissenschaft –
Fachgebiet Historische Bildungsforschung
der Ruhr-Universität Bochum

friederike.thole@rub.de