

ANDERS PRÜFEN

**Prüfungen zählen zu den ungeliebten,
aber grundsätzlich notwendigen Aufgaben
im Lehrberuf. Seit Langem stehen sie in
der pädagogischen Kritik, insbesondere
in der Form summativer, individueller
Schülerbeurteilungen anhand von
Ziffernnoten beziehungsweise Zensuren.
Was lässt sich anders machen?**

MATTHIAS TRAUTMANN UND ALBRECHT WACKER

Schriftliche, mündliche und zuweilen auch praktische Prüfungen nehmen im Berufsalltag von Lehrpersonen eine besondere Stellung ein. Überwiegend werden sie von Lehrpersonen erstellt, zuweilen von der Bildungsverwaltung autorisiert, in einem für die Schüler:innen vorgegebenen Zeitrahmen bearbeitet und im Anschluss bewertet. Die Regeln für die Durchführung von Prüfungen sind in hohem Maß formalisiert, und häufig unterliegen sie in ihrer Beurteilung höheren Objektivitätsanforderungen: Beispielsweise werden Handreichungen zur Punktvergabe und Beurteilung beigelegt, doppelte Korrekturen von unabhängigen Personen verlangt oder die Prüfungen von schulfremden Prüfer:innen – wie Lehrpersonen benachbarter Schulen – abgenommen.

In einem weiten Sinne fungieren alle Verfahren als Prüfung, in denen Kompetenzen von Schüler:innen festgestellt und beurteilt werden. In diesem Begriffsverständnis ist eine Abgrenzung zur sonstigen Leistungsbeurteilung an Schulen nicht trennscharf möglich. Enger gefasst sind damit herausgehobene Anlässe gemeint, die Bildungs- und Lebenschancen eröffnen (oder erschweren) und mit Berechtigungen für den weiteren Lebensweg einhergehen. In diesem engeren Verständnis bilden Prüfungen eine Sonderform der Leistungsbeurteilung, stellen einen Verwaltungsakt dar und können – im Gegensatz zu üblichen Klassenarbeiten – mit Widerspruch und Klage angefochten werden.

Unabhängig vom jeweiligen Begriffsverständnis sind Prüfungssituationen in einem pädagogischen Sinn stets Leistungssituationen, die von den Lernsituationen getrennt und unterschieden werden sollten. Ihre grundlegende Idee bleibt eine auf das Individuum zugerechnete Beurteilung des jeweiligen Fähigkeitsstandes im Vergleich mit einem Standard/einer Norm und (meist) in Form einer Zensur.

BEGRÜNDUNG VON PRÜFUNGEN UND REFORMIDEEN

Prüfungen begründen sich nach der älteren Literatur in der Allokationsfunktion der Schule: In einer meritokratischen Gesellschaft werden gesellschaftliche Positionen – zumindest der Idee nach – durch Leistungen vergeben und die »Zuord-

nungen zwischen den Leistungen der Schülerschaft und ihren beruflichen Laufbahnen« erfolgen über das Prüfungswesen (Fend 2006, S. 50). Am Beispiel Frankreichs wird ersichtlich, wie das Bildungswesen in einem hohen Ausmaß von Prüfungen (Concours) durchzogen ist, die damit auch zu den bedeutendsten Gestaltungselementen der inneren Schularchitektur zählen.

Wo steht die Diskussion heute und welche Perspektiven hat die aktuelle Literatur auf das Themenfeld? Insgesamt scheint sie zwei unterschiedlichen Richtungen verpflichtet zu sein: mehr Vergleichbarkeit und Standardisierung auf der einen, mehr Individualisierung auf der anderen Seite. Eine höhere Vergleichbarkeit wird über die Perspektiven von Schulverwaltung und Bildungspolitik eingebracht und betrifft überwiegend (Abschluss-)Prüfungen im engeren Sinne, die eine Basis für weitere Karriereentscheidungen sind. In der

empirischen Forschung wie auch der öffentlichen Diskussion stehen deren Qualität und Vergleichbarkeit (zwischen Bundesländern, zwischen Einzelschulen) im Vordergrund (vgl. z. B. Hoffmann et al. 2022). Es gibt Anzeichen dafür, dass sich die Auslesefunktion von Prüfungen in den letzten Jahren deutlich reduziert hat. Wir wissen z. B., dass die

Noten, etwa im Abitur, tendenziell besser werden, was vermutlich – hier bräuchte es dringend Studien – nicht in gleichem Maße mit den Leistungen korrespondiert.

Eine verstärkte Individualisierung wird dagegen hauptsächlich über variable Prüfungsformate realisiert, bei denen auch formatives Assessment in den Blick kommt und verstärkt die individuelle Bezugsnorm zur Beurteilung herangezogen wird. In diesem Themenfeld liegt viel pädagogische Literatur vor, die auf einem weiteren Prüfungsbegriff basiert: Ausgehend von der Kritik an Noten seit den 1970er-Jahren wurden und werden Prüfungen beziehungsweise Leistungsbeurteilungen generell in den entsprechenden Empfehlungen durch verbale Urteile und gesprächsorientierte Rückmeldungen ergänzt, etwa mittels Portfolios, Projektberichten oder Lerntagebüchern (vgl. z. B. Stern 2008; Bohl 2009; Winter 2016); derartige Instrumente und Verfahren scheinen in der Schulpraxis auch zunehmend Verbreitung zu finden. Dabei wird die

Mehr Vergleichbarkeit und Standardisierung auf der einen, mehr Individualisierung auf der anderen Seite.

Viel Aufregung haben in letzter Zeit KI-basierte Sprachgeneratoren wie ChatGPT erfahren.

Beurteilung durch die Lehrkraft durch Selbsteinschätzungen der Schüler:innen erweitert, welcher laut Hattie ein äußerst hoher Effekt zukommt (Zierer 2022, S. 48), und die Selbstbeurteilung mit der Fremdeinschätzung verglichen. Auch wurden einhergehend mit der Kompetenzorientierung der 2000er-Jahre Prüfungs- und Testaufgaben propagiert, in denen statt abfragbarem Wissen (welches als wenig wichtig markiert wurde) nunmehr Kompetenzen mittels komplexer Problem- und Anwendungssituationen geprüft werden sollten (vgl. z. B. HRK 2015; Lötscher et al. 2021); auch hier spielen Selbsteinschätzungs- beziehungsweise Rückmeldeformate wie z. B. Kompetenzraster eine wichtige Rolle. Aktuelle Veröffentlichungen werben entsprechend für Verfahren, in denen Feedback zum Lernen, Dialog und Reflexion, Selbstbeurteilung und Orientierung an den Stärken der Schüler:innen großgeschrieben werden sollen, ein erweiterter Leistungsbegriff überfachliche Kompetenzen einschließt und auch digitale Formate in Prüfungen und die Leistungsbeurteilung einbezogen werden (vgl. Winter 2016; HRK 2015).

LEGITIMATIONSDEFIZITE VON PRÜFUNGEN

Stimmen die älteren Begründungsfiguren für Prüfungen noch, die häufig ihre Anleihen in der Gesellschaft der 1960er-Jahre machen und theoretisch im Strukturfunktionalismus verortet sind? Selektion und Allokation waren damals insbesondere erforderlich, weil eine größere nachwachsende Generation auf eine kleinere Anzahl von gesellschaftlichen Positionen zielte, was zu einer Aufwertung von Prüfungen und ihrer Legitimation führte. Hier hat sich mittlerweile in demografischer Hinsicht, mit Blick auf die heute vorhandene Optionsvielfalt, die Schulabgänger:innen zur Verfügung steht, und auch in den theoretischen Perspektivierungen viel verändert:

- Die damaligen geburtenstarken Jahrgänge, von denen der Jahrgang 1964 der größte war, stehen heute in ihren letzten Arbeitsjahren und vor dem Eintritt in den Ruhestand. Nicht zuletzt der täglich in der Presse angezeigte Fachkräftemangel verdeutlicht, dass eine kleinere Kohorte der nachwachsenden Generation die vorhandenen gesellschaftlichen Positionen und Plätze heute und in Zukunft kaum mehr auszufüllen vermag.
- Gleichzeitig haben sich die schulischen Anschlussoptionen für Jugendliche insgesamt ausgeweitet. Sie zeigen sich

an etwa 400 definierten Ausbildungsberufen und einer Fülle an möglichen Studienfächern, die in ihrer Breite bisher nicht anzutreffen war. In größerem Ausmaß als früher sind Seiten- und Quereinstiege (oder -ausstiege) möglich.

- Damalige Begründungsmuster zielten vor allem auf die (statische) Aufrechterhaltung der Gesellschaft, vernachlässigten jedoch die Schüler:innen als Individuen. In einer von Individualisierungsschüben durchzogenen Gesellschaft kommt hier den Schüler:innen und ihrem Wohlergehen in der Schule eine akzentuierte Bedeutung zu.

Die demografischen Entwicklungen schränken die Allokationsfunktion der Schule ein und überantworten ihr den Auftrag, individuelle Potenziale zu fördern und – dies ist ebenfalls wichtig – die Schüler:innen auf die Vielfalt der möglichen Anschlussoptionen vorzubereiten. Damit greifen die älteren Begründungsmuster nur noch eingeschränkt.

NEUE MÖGLICHKEITEN IM KONTEXT DER DIGITALISIERUNG

Die seit etwa 2019 Fahrt aufnehmende schulische Digitalisierung und der Distanzunterricht in der Covid-19-Pandemie schließlich führten und führen noch zu Aufgaben, die sich neue technische Möglichkeiten zunutze machen wollen und dafür ebenfalls den Gestus des Zeitgemäßen beanspruchen (z. B. Take-Home-, Open-Media-Klausuren, E-Books, Podcasts und Videos, vgl. prüfungskultur.de; Middendorf 2022). Viel Aufregung haben in letzter Zeit KI-basierte Sprachgeneratoren wie ChatGPT erfahren, mit denen schriftliche Aufgaben und Prüfungen ohne Aufsicht quasi in Sekundenschnelle gelöst werden können, auch bei individualisierten Problem- und Fragestellungen. Hier stellen sich künftig Fragen wie: Sollen derartige Werkzeuge in den Unterricht und in Prüfungen zugelassen und eingebaut werden oder eher nicht? Bedeuten die Instrumente das Aus für häusliche Arbeiten jeder Art? Die Diskussion ist noch zu neu, um die – rechtlichen und pädagogischen – Konsequenzen für das schulische Prüfungswesen (z. B. Facharbeiten, schriftliche Aufgaben allgemein, mündliche Prüfungen) abschätzen zu können. PÄDAGOGIK wird sich im nächsten Jahr in einem eigenen Themenschwerpunkt mit diesen Fragen befassen.

Was sind angemessene Aufgaben, Verfahren, Mitteilungs- und Rückmeldeformate für das 21. Jahrhundert, die auch den neuen Möglichkeiten digitalen Lehrens und Lernens Rechnung tragen?

IST UNSERE PRÜFUNGSKULTUR NOCH ZEITGEMÄß?

Es bleibt bilanzierend zu fragen: Ist unsere Kultur des Prüfens noch zeitgemäß? Bedarf es anderer Formate bei den Prüfungen? Was sind angemessene Aufgaben, Verfahren, Mitteilungs- und Rückmeldeformate für das 21. Jahrhundert, die auch den neuen Möglichkeiten digitalen Lehrens und Lernens Rechnung tragen? Anders prüfen könnte bedeuten, die schon bisher in der Literatur vertretene pädagogische Linie weiter zu akzentuieren und darin die an die Schüler:innen gerichteten Funktionen der Schule in den Blick zu nehmen, ihre Potenziale zu fördern, ihre Berufsfähigkeit zu stärken und ihr Wohlbefinden im Blick zu behalten (vgl. auch Schmid 2018). Im vorliegenden Themenheft werden deshalb Erfahrungen, die daran anknüpfen und das traditionelle Spektrum erweitern, in Auswahl und im Hinblick auf eine gelingende Praxis vorgestellt.

Hendrik Haverkamp zeigt in seinem Beitrag, wie klassische Formate der Leistungsbeurteilung durch alternative, digitale und lernförderliche Formate ergänzt werden können. *Sönke Zankel* befasst sich mit Open-Book-Klausuren, bei denen Schüler:innen alles verwenden dürfen, was sie mitbringen, inklusive zuvor ermittelter Suchergebnisse aus dem Netz. *Andreas Baumgärtner*, *Tamara Erschig* und *Monika Becker* berichten von mehrstufigen »Gelingensnachweisen«, die nicht zu einem einzigen Zeitpunkt, sondern innerhalb eines Zeitraums geschrieben werden. *Jens Lindström* skizziert in seinem Beitrag zeitgemäße Formate bei der Beurteilung mündlicher Leistungen, zu denen er auch die Beurteilung von Videos anführt. Auch *Joel Trappmann* betont die Förderaufgabe der Schule und berichtet Erfahrungen zur Implementation gelingender Coachinggespräche. *Kathleen Falkenberg* schließlich bietet einen Blick nach Schweden als Beispiel einer anderen Prüfungskultur und lenkt die Aufmerksamkeit auf strukturelle Fragen. Die Behörden werden sich in diesem Sinne auch mit der Frage befassen müssen, unter anderem angesichts der Belastung von Lehrkräften und Schüler:innen, inwiefern die Prüfungssysteme der deutschen Bundesländer hinsichtlich Qualität und Quantität reformiert werden müssen.

Alle Beiträge sind einem weiten Prüfungsbegriff verpflichtet, der in die Leistungsbeurteilung insgesamt hineinreicht, und zielen insbesondere auf die Stärkung der Schüler:innen, auf die Beurteilung ihrer Lernprozesse und auf Rückmeldun-

gen, die den Lernprozess befördern. Wie immer wollen sie Lehrer:innen Anregung für die Prüfungsgestaltung bieten. ♦

LITERATUR

- Bohl, T. (2009): Prüfen und Bewerten im Offenen Unterricht. 4., neu ausgestattete Auflage. Weinheim/Basel.
- Fend, H. (2006): Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen. Wiesbaden.
- Hoffmann, L./Schröter, P./Groß, A./Schmid-Kühn, S. M./Stanat, P. (Hrsg.): Das unvergleichliche Abitur. Entwicklungen – Herausforderungen – Empirische Analysen. Bielefeld.
- HRK (2015): Kompetenzorientiert prüfen. Zum Lernergebnis passende Prüfungsaufgaben (Nexus Impulse für die Praxis, Ausgabe 4). Verfügbar unter www.hrk-nexus.de/fileadmin/redaktion/hrk-nexus/07-Downloads/07-02-Publikationen/HRK_Ausgabe_4_Internet.pdf (20.2.2023).
- Lötscher, H./Naas, M./Roos, M. (2021): Kompetenzorientiert beurteilen. Bern.
- Middendorf, W. (2022): Schulische Leistungsbewertung im Kontext von Digitalisierung – erste Hinweise zu einer von der KMK angestrebten erweiterten Prüfungskultur. Verfügbar unter www.pedocs.de/frontdoor.php?source_opus=23908 (20.2.2023).
- Schmid, C. (2018): Beurteilung des Gelernten und Beurteilen für das Lernen. In: Berner, H./Fraevel, U./Zumsteg, B. (Hrsg.): Didaktisch handeln und denken. Bern, S. 256–272.
- Stern, T. (2008): Förderliche Leistungsbewertung. Verfügbar unter www.ozepts.at/wp-content/uploads/2011/07/Leistungsbewertung_Onlineversion_Neu.pdf (20.2.2023).
- Winter, F. (2016): Leitfaden: Neue Formen der Leistungsbewertung. Verfügbar unter www.iqesonline.net/unterrichten/kompetenzorientierte-leistungsbeurteilung/neue-formen-der-leistungsbeurteilung/
- Zierer, K. (2022): Hattie für gestresste Lehrer. Kernbotschaften und Handlungsempfehlungen aus John Hatties »Visible Learning« und »Visible Learning for Teachers«. Hohengehren.

MATTHIAS TRAUTMANN ist Professor für Erziehungswissenschaft mit dem Schwerpunkt Schulpädagogik und Allgemeine Didaktik an der Universität Siegen und Mitglied der Redaktion von PÄDAGOGIK.

✉ matthias.trautmann@uni-siegen.de

ALBRECHT WACKER ist Professor für Schulpädagogik der Sekundarstufe I an der Pädagogischen Hochschule Ludwigsburg und Mitglied der Redaktion von PÄDAGOGIK.

✉ albrecht.wacker@ph-ludwigsburg.de