

# Feedback

Bibliografie:

Bernhard Hemetsberger:

Peripetien der Ziffernnoten.

*journal für lehrerInnenbildung*, 19 (1), 80-87.

[https://doi.org/10.35468/jlb-01-2019\\_07](https://doi.org/10.35468/jlb-01-2019_07)

Gesamtausgabe online unter:

<http://www.jlb-journallehrerinnenbildung.net>

<https://doi.org/10.35468/jlb-01-2019>

ISSN 1681-7028 (Printausgabe)

journal für lehrerInnenbildung  
jlb  
no. 1  
2019

**07**

*Bernhard Hemetsberger*

Peripetien  
der Ziffernnoten

Schulische Bewertungsformen und besonders Ziffernnoten sind wiederkehrend Teil von öffentlichen und politischen Debatten. Ihnen wird zugetraut der Angelpunkt von Schultheorie, pädagogischer Praxis und Schulerfahrungen gleichermaßen zu sein und deshalb den Kern von Schülerleistung zu erfassen (vgl. Tocci, 2010). Verständlich, dass sich daran Steuerungsphantasien und Konflikte entzünden, wenn mit guten Absichten, praktischer Erfahrung oder elterlicher Wunschvorstellung ein Umschwung zu vortrefflicheren Formen der Leistungsbeurteilung herbeigesehnt wird. Man fordert nicht weniger als bessere Rückmeldungen, anschlussfähigere Auskünfte und effizienteres Schullehalten. Dieser Forschungsbericht soll nicht mehr als Anmerkungen zu Grundlagen für Peripetien, historische und international-vergleichende Umschwünge von Ziffernnoten und abschließende Ausblicke anbieten.

## Grundlagen der Peripetien

In der aristotelischen Poetik wird Peripetie als Umschlag der Unternehmungen in ihr Gegenteil, und zwar, nach innerer Wahrscheinlichkeit oder Notwendigkeit gefasst (Gomperz, 1897, S. 23). Dass dies auf die Geschichte der Notengebung – wahrlich keine linear-konfliktfreie – zutrifft, zeigt sich auch an einer großen Zahl jährlich erscheinender internationaler Studien zur schulischen Leistungsbewertung (vgl. auf <https://eric.ed.gov/> erfasste Literatur). Sie bilden widersprüchliche Aussagen zur wissenschaftsbasierten Bevorzugung verschiedener Ausprägungsformen ab. Es gilt sie deshalb zu befragen: Stehen inner-schulisch-pädagogische Bezugsnormen im Fokus der Beurteilungsform oder sollen außerschulische Informationsbedürfnisse bei Eltern oder Arbeitgeber\*innen bedient werden? Welche Formen und Mischverhältnisse von Leistungsbeurteilungspraktiken werden eingesetzt und können daher miteinander verglichen werden? Welche Aspekte im Trubel der alltäglichen Unterrichtsgestaltung beziehen sich überhaupt offenkundig auf Leistungsbeurteilung? Verschiedentliche Antworten darauf und ihr jeweils divergierender Fokus in wissenschaftlichen Studien zeigt die Wankelmütigkeit von Leistungsbeurteilung, die in den Kategorien Bezugsnorm, Funktionen und Prüfungsformaten analysiert werden können.

Der räumliche Zusammenhang reicht von der Klasse als Vergleichsmaßstab – zumeist treten hier krude Vorstellungen von einer Normalverteilung der Leistung, also wenige am sehr guten bzw. nicht genügenden Rand und die Mehrheit im befriedigenden Mittel, zu Tage – bis zu internationalen Vergleichsstudien als globale Bezugsnorm der Schulleistung – PISA und Co sollten lokale Schulstandorte beispielsweise von Shanghai und Wien in Beziehung miteinander setzen können. Neben der räumlichen Bezogenheit spielen verschiedene Zeitspannen für Leistungsbewertungen eine zentrale Rolle. Die zeitlichen Bezugsnormen fassen ebenso weitläufig das Spektrum von zurückgelegtem Lernweg, über erreichten Stand, bis hin zur prognostischen Potentialbekundung – was jemand möglicherweise in Zukunft zu leisten im Stande sein wird. Schließlich wären noch inhaltliche Rückgriffe zu nennen, die sich auf die Wiedergabe von sachlichen Einzeldingen bis dahinterliegende Kompetenzvermutungen erstrecken können. Schon zeichnet sich ab, dass Funktionen von Beurteilungsformen der Diagnose, Motivation, Allokation oder Legitimation Ausdruck verleihen, wenngleich sich zuweilen die Einzelfunktionen gegenseitig behindern (vgl. Eder, Neuweg & Thonhauser, 2009; Darling-Hammond & Wentworth, 2010). Motiviert die Beurteilung einerseits zu höheren Zielen, kann die Zugänglichkeit von besseren beruflichen Positionen oder Studienplätzen an Universitäten dennoch unerreichbar sein (vgl. Labaree, 2010). Ähnliches wird bei einer näheren Differenzierung von Prüfungsformaten, die schließlich beurteilt werden können, ersichtlich. Von offenen zu standardisierten, von einmaligen zu wiederkehrenden, von personenbezogenen zu kollektiven, von personalisierten zu vergleichenden, von lokalen, nationalen zu globalen, von freiwillig bis verpflichtenden Prüfungsformaten unterscheiden sich die Dimensionen, die Reichweite und Aussagekraft der Beurteilungsformen (vgl. Reese, 2013).

Steht das eine oder andere im gewünschten Zentrum einer Vorstellung von Leistungsbeurteilung, so liegt es in der inneren Wahrscheinlichkeit oder gar Notwendigkeit, dass Umschläge ins Gegenteil der jeweils vorherrschenden Leistungsrückmeldungspraktik gefordert werden. Dermaßen viele Einfallspunkte für Kritik fordern geradezu auf, alternative Konzepte anzustreben und befeuern Krisenerzählungen, dass Schulen nicht leisten würden was von ihnen erwartet wird.

## Historische Umschwünge

In der Geschichte von schulischem Unterricht haben sich mit der Reformation und Gegenreformation Ziffernnoten ab 1599 (in der *Ratio Studiorum* der Jesuiten) neben lokalen Bewertungspraktiken, bis ins späte 19. Jahrhundert vollständig, durchgesetzt (vgl. Hemetsberger, 2015). Über beinahe 300 Jahre wurden also offene und lokale Formen der Leistungsbewertung in zunehmend standardisierte und überregionale Formate überführt (vgl. Ziegenspeck, 1999). Das geschah vornehmlich aus organisatorischen Gründen zur Reduktion steigender Komplexität bei zunehmenden Schülerzahlen in allgemein zugänglichen und öffentlichen Schulsystemen (vgl. Smallwood, 1935). Dabei ist bei einem der ältesten und stabilsten Rückmeldungspraktiken an Schulen, den Schulnoten, ein Umschlag im Fokus von einer innerschulischen zur außerschulischen Kommunikationsausrichtung festzustellen (vgl. Schneider & Hutt, 2014). Nicht pädagogische Gründe alleine sind im Zentrum der Notengebung, sondern auch außerschulische Anforderungen derjenigen Personen, die während der Unterrichtszeit nicht anwesend sein können, jedoch Interesse am Schulgeschehen haben – Eltern, Arbeitgeber\*innen, u. a. Gruppen. Früh regt sich Kritik, Ziffernnoten würden zu allgemein, abstrakt und wenig förderlich für das Lernen der Schüler\*innen sein (vgl. beispielsweise Martinak, 1900) oder nur ein „Stenogramm kleinteiliger Ausschnitte der Schule“ (Engelbrecht, 2008, S. 65) sein. Einsetzende Psychologisierungstendenzen versuchten die Vergabe von Ziffernnoten in objektivere und scheinbar gerechtere Praktiken zu überführen. Anhaltende Definitionspräzisierungen von Leistungskategorien oder gegenwärtiger Kompetenzrasterbeschreibungen sind eine Ausprägung davon. Durch derlei Untersuchungen und Bemühungen konnten beurteilungsrelevante Faktoren abseits von Schulleistung aufgezeigt werden. Stimm- lage, Sprechgeschwindigkeit, Wortwahl, Sprachstil, und andere mehr (vgl. dazu ein Überblick in Ingenkamp, 1971) oder auch Schönheit (vgl. Dunkake, Kiechle, Klein & Rosar, 2012) offenbarten, dass jedwede Leistungsbewertung ihre interpersonellen Befangenheiten nicht ablegen kann. Alternative Beurteilungsformen versuchten seit der aufkommenden Kritik an Schulnoten neue Möglichkeiten zur „pädagogisch wertvolleren Rückmeldung“ zu entwickeln. Besonders im reformpädagogischen Milieu wurden „notenfreie“ und alternative Unterrichtssettings erprobt. Diese erfreuen sich bis zur Gegenwart hin

großer Beliebtheit als Gegenargument zu Schulnoten, da „[r]eformpädagogische Methoden [...] nie einen Risikostempel [tragen; Anm. H.B.], sie werden in bester Absicht verwendet [...] stossen aber auf eine harte Wirklichkeit, die den guten Absichten oft im Wege steht“ (Oelkers, 2012, S. 5).

## International: Oszillation in Ruheposition?

Gegenwärtig sind daher international vielerlei Mischvarianten von benoteten und unbenoteten, standardisierten und offenen sowie zentralen und lokalen Prüfungsformen verbreitet (vgl. Darling-Hammond & Wentworth, 2010). Die Einsicht, dass Ziffernnoten keine leistungsbezogene „Wahrheit“, aber – wenn auch problematische – Annäherungen an die Assimilation von Schüler\*innen an die schulinternen Praxen liefern, die über Schülerleistung hinausgehen, erhöht das Lehrerwohlbefinden bei der Vergabe möglicherweise nicht (vgl. Falkenberg, 2018). Kommen doch diejenigen besser weg, die ohnehin aufgrund ihres bevorzugten Hintergrunds (sog. Bildungsnähe) sich leichter im Schulgeschehen zurechtfinden. Tendenzen zur Abschiebung der Verantwortung von Beurteilungen auf standardisierte und außerschulische Bewertungszentren (vgl. Hopmann, 2008) oder nachfolgende Institutionen durch „grade inflation“ (vgl. Boretz, 2010) oder einer Notenabwertung durch ausschließliche Vergabe der Bestnote sind zu beobachten. Gute Noten als Prognosen für Lebenserfolg sind klarerweise ohnehin wenig vertrauenserweckend, jedoch über die Länge der Schullaufbahn doch eher vorhersagetauglich, da sie über das Zurechtkommen mit dem Setting Schule (auch abseits der Leistungsanforderungen) relativ stabil Auskunft geben (vgl. Eder & Thonhauser, 2006).

Unterricht und fortlaufende Leistungsüberprüfung machen das Kerngeschäft aller Schulen aus. Es gehört mitunter zum Professionsstatus von Lehrer\*innen, über die Leistungen der Schüler\*innen in ihrem Fach (im rechtlichen Rahmen) autonom, aber nach offengelegten Kriterien und nach eigenem Urteil, aber konsistent, bestimmen zu können (vgl. Oelkers, 2012; Hemetsberger, 2015). Dem geschuldet ist, dass viele Länder, darunter auch die deutschsprachigen, Erziehungsberechtigten oder auch Schüler\*innen nur geringfügige rechtliche Mittel zum Einspruch von Beurteilungsergebnissen ermöglichen (vgl.

Waldow, 2013). Während jedoch international vielerlei Mischformen eingeführt wurden (vgl. Darling-Hammond & Wentworth, 2010), das Pendel sich metaphorisch in Ruheposition zwischen den beiden extremen Wendepunkten befindet, bewegen sich deutschsprachige Länder noch zusehends mit besten Absichten auf einem sich ausweitenden Standardisierungspfad (vgl. beispielsweise die „standardisierte Reifeprüfung“ in Österreich, den „Lehrplan 21“ in der Schweiz oder die „Bildungsstandards“ in Deutschland). Wie national respektive international die Entwicklung weiter verläuft ist gemäß den beschriebenen Pendelbewegungen vorauszuahnen.

## Ausblick: zu erwartende Pendelbewegung

Peripetien von lokalen und regionalen Bewertungspraktiken zur Standardisierung und Zentralisierung haben den betroffenen Lehrer\*innen das Unwohlsein über die Vergabe von ungerecht empfundenen, lokal gebundenen Bewertungen und damit einhergehend einen Teil ihrer Professionalität genommen, dabei aber nicht zu besseren und gerechteren Leistungsbewertungen geführt. Einerseits befördern diese Prüfungsformate „teaching to the test“, da geprüft wird, was den Testkonstruktionen zu entnehmen ist und nicht was im jeweiligen Unterricht lehrendenabhängig durchgearbeitet wurde. Dadurch kommt es zum nachweislich schleichenden Verschwinden nicht prüfungsrelevanter Inhalte an Schulen (vgl. Labaree, 2014). Andererseits werden dabei gerade außerschulische Unterschiede der Schüler\*innen, seien es soziokulturelle oder -ökonomische – verstärkt zur Einflussgröße in der Differenzmarkierung (vgl. Hopmann & Bauer, 2015; Nichols & Berliner, 2007). Die Rückmeldungspraxis an Schulen, wessen Informationsbedürfnis auch immer im Zentrum steht, kommt aber nicht ohne Bewertungsform aus (vgl. Hemetsberger, 2017). Stehen Lehrer\*innen und Schüler\*innen im Zentrum der Rückmeldungspraxis, können Schulnoten in ihrem allgemeinen, abstrakten Charakter bestehen bleiben, um außerschulische Anknüpfungsmöglichkeiten zu schaffen (übrigens eine erfolgreiche und etablierte Funktion), in der innerschulischen Praxis sollten Leerstellen, intersubjektive Prägungen, bekannte Schwierigkeiten und beschränkte Aussagekraft von Schulnoten diese aber an ihren gebührenden Platz stellen. Auch wenn in den deutschsprachigen Ländern das Ende zunehmender Standardisierungsbemühungen nicht

in Sicht ist, könnten lokale Formate der Leistungsbewertung etabliert werden. Dabei wäre die pädagogische Funktion von Noten, der Zwischenauskunft und Ausgangspunkt für weiteres Arbeiten und Fördern für Lehrpersonen angereichert mit lokalen, nicht-standardisierten und offenen Formen aus der Forschungsperspektive angeraten, wiewohl Praxis nur in der Praxis verändert werden kann (vgl. Heid, 2015, S. 392). Gute pädagogische Praxis findet in der Wissenschaft zwar keine allgemeingültigen Rezepte jedoch eine Informationsquelle.

## Literatur

- Boretz, E. (2010). Grade Inflation And The Myth Of Student Consumerism. *College Teaching*, 52 (2), 42-46.
- Darling-Hammond, L. & Wentworth, L. (2010). *Benchmarking Learning Systems: Student Performance Assessment in International Context*. Stanford: Stanford University, Stanford Center for Opportunity Policy in Education.
- Dunkake, I., Kiechle, T., Klein, M. & Rosar, U. (2012). Schöne Schüler, schöne Noten? Eine empirische Untersuchung zum Einfluss der physischen Attraktivität von Schülern auf die Notenvergabe durch das Lehrpersonal. *Zeitschrift für Soziologie*, 41 (2), 142-161.
- Eder, F., Neuweg, H. G. & Thonhauser, J. (2009). Leistungsfeststellung und Leistungsbeurteilung. In W. Specht (Hrsg.), *Nationaler Bildungsbericht Österreich Bd. 2. Fokussierte Analysen bildungspolitischer Schwerpunktthemen*. Verfügbar unter <https://www.bifie.at/buch/1024/b/6> [10.12.2018].
- Eder, F. & Thonhauser, J. (2006). Bildungsaspirationen, Noten und Berechtigungen am Übergang von der Volksschule in die Sekundarstufe I. *Erziehung und Unterricht*, 156 (3-4), 275-294.
- Engelbrecht, H. (2008). *Verhaltenssteuerung und Leistungsbeurteilung in der Schule: von den Anfängen bis zur Gegenwart auf dem Boden des heutigen Österreich*. Wien: Jugend & Volk.
- Falkenberg, K. (2018). *Gerechte Noten? Eine Grounded Theory-Studie zu Gerechtigkeitsüberzeugungen von Lehrkräften in Bezug auf schulische Leistungsbeurteilung im deutsch-schwedischen Vergleich*. Abstract zum Humboldt-Preis 2018. Verfügbar unter <https://www.hu-berlin.de/de/ueberblick/menschen/ehrungen/humboldtpreis/archiv/humboldt-preis-2018/gerechte-noten-eine-grounded-theory-studie-zu-gerechtigkeits-ueberzeugungen-von-lehrkraeften-in-bezug-auf-schulische-leistungsbeurteilung-im-deutsch-schwedischen-vergleich> [10.12.2018].
- Gomperz, T. (1897). *Aristoteles' Poetik. Übersetzt und eingeleitet*. Leipzig: Verlag von Veit & Comp.
- Heid, H. (2015). Bildungsforschung im Kontext gesellschaftlicher Praxis. Über (soziale) Bedingungen der Möglichkeit, Bildungspraxis durch Bildungsforschung zu beeinflussen. *Zeitschrift für Pädagogik*, 61 (3), 390-409.
- Hemetsberger, B. (2015). *Nicht Genügend ... Setzen! Zur Geschichte der Notengebung in Österreich*. Berlin, u. a.: LIT Verlag.

- Hemetsberger, B. (2017). Schulnoten? – Das ist hier nicht die Frage! Historische und (system-)theoretische Bemerkungen. *Vierteljahrsschrift für wissenschaftliche Pädagogik*, 2, 255-267.
- Hopmann, S. (2008). No child, no School, no State Left Behind: Schooling in the Age of Accountability. *Journal of Curriculum Studies*, 40 (4), 417-456.
- Hopmann, S. & Bauer, S. (2015). Das „Equity Paradox“. In S. Rademacher & A. Wernet (Hrsg.), *Bildungsqualen – Kritische Einwürfe zum pädagogischen Zeitgeist* (S. 93-105). Wiesbaden: Springer VS Verlag.
- Ingenkamp, K. (Hrsg.). (1971). *Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte* (2. Aufl.). Weinheim & Basel: Beltz.
- Labaree, D. (2010). *Someone Has to Fail. The Zero-Sum Game of Public Schooling*. Cambridge, a.o.: Harvard University Press.
- Labaree, D. (2014). Let's Measure What No One Teaches: PISA, NCLB, and the Shrinking Aims of Education. *Teachers College Record*, 116 (9), 1-14.
- Martinak, E. (1900). *Psychologische Untersuchungen über das Prüfen und Classifizieren*. Wien: Hölder.
- Nichols, S. L. & Berliner, D. C. (2007). *Collateral Damage. How High-Stakes Testing Corrupts America's Schools*. Cambridge: Harvard Education Press.
- Oelkers, J. (2012). *Noten und Leistungen: Befunde und praktische Vorschläge*. Vortrag auf der LUXDIDAC am 3. Oktober 2012. Verfügbar unter [https://www.ife.uzh.ch/dam/jcr:00000000-4a53-efcc-0000-00004d513180/Luxemburg\\_Noten.pdf](https://www.ife.uzh.ch/dam/jcr:00000000-4a53-efcc-0000-00004d513180/Luxemburg_Noten.pdf) [10.12.2018].
- Reese, W. J. (2013). *Testing Wars in the Public Schools. A Forgotten History*. Cambridge, London: Harvard University Press.
- Schneider, J. & Hutt, E. (2014). Making the Grade: History of the A-F Marking Scheme. *Journal of Curriculum Studies*, 46 (2), 201-224.
- Smallwood, M. L. (1935). *An Historical Study of Examinations and Grading Systems in Early American Universities*. Cambridge: Harvard University Press.
- Tocci, C. (2010). An Immanent Machine: Reconsidering Grades, Historical and Present. *Educational Philosophy and Theory*, 7, 762-778.
- Waldow, F. (2013). Conceptions of Justice in the Examination Systems of England, Germany, and Sweden: A Look at Safeguards of Fair Procedure and Possibilities of Appeal. *Comparative Education Review*, 58 (2), 322-343.
- Ziegenspeck, J. W. (1999). *Handbuch Zensur und Zeugnis in der Schule. Historischer Rückblick, allgemeine Problematik, empirische Befunde und bildungspolitische Implikationen. Ein Studien- und Arbeitsbuch*. Bad Heilbrunn: Klinkhardt.

Bernhard Hemetsberger, MA,  
Univ.-Ass. für historische und vergleichende  
Schul- und Bildungsforschung  
an der Universität Wien.  
Arbeitsschwerpunkte: Schulnoten,  
Schulkrisenerzählungen



[bernhard.hemetsberger@univie.ac.at](mailto:bernhard.hemetsberger@univie.ac.at)