



Innovative Leistungsfeststellung und Leistungsbe- wertung an der Privaten Pädagogischen Hochschule der Diözese Linz – Vision Statement der *Interessensgemeinschaft Innovative Grading IG²*

Johannes Reitingger, Alfred Weinberger, Ursula Svoboda, Ernst Nausner,
Ute Vogl, Wolfgang Bachl, Marlene Bauer, Emmerich Boxhofer, Veronika Grabner,
Franz Keplinger, Cornelia Kittinger, Bernhard Kitzmüller, Anna Neulinger,
Clemens Seyfried und Elke Wöß^a

^aalle: Private Pädagogische Hochschule der Diözese Linz
johannes.reitingger@ph-linz.at

EINGELANGT 27 FEB 2018

ÜBERARBEITET 31 MAI 2018

ANGENOMMEN 07 JUN 2018

Der vorliegende Beitrag beschreibt eine Projektinitiative der Privaten Pädagogischen Hochschule der Diözese Linz im Bereich innovativer Leistungsfeststellung und Leistungsbewertung. Die neu gegründete Interessensgemeinschaft Innovative Grading IG² verfolgt dabei das Ziel, Alternativen zur herkömmlichen Ziffernzensur zu untersuchen und einzusetzen. Im Folgenden werden in Form eines Vision Statement der Interessensgemeinschaft erste Diskussionsergebnisse, ein Literaturreview sowie eine mittelfristige Begleitforschungsperspektive dargelegt.

SCHLÜSSELWÖRTER: Leistungsfeststellung, Innovative Grading, Kritik der Ziffernzensur,
hochschulisches Qualitätsmanagement

1. Ausgangssituation

Fragen der Leistungsfeststellung und Leistungsbewertung sind für jede Bildungsinstitution von Relevanz. In der Ausbildung von Lehrpersonen kann von einer zweifachen Relevanz gesprochen werden. Einerseits geht es wie in jeder Bildungsinstitution darum, dass die Leistungen der Lernenden auf der Basis lernpsychologischer und gesetzlicher Grundlagen festgestellt und bewertet werden. Andererseits setzen sich Lehramtsstudierende während ihrer Ausbildung selber mit der Thematik auseinander und es wird erwartet, dass sie in ihrem eigenen Unterricht geeignete Methoden verwenden. Auf diesem Hintergrund erweist es sich als sinnvoll, wenn das System der Leistungsfeststellung und -bewertung einer Ausbildungsinstitution auf der Basis neuer Erkenntnisse kritisch analysiert und adaptiert wird. Dieser Analyseprozess findet gegenwärtig an der Privaten Pädagogischen Hochschule der Diözese Linz statt. Dazu wurde eine Interessensgemeinschaft In-

novative Grading IG² implementiert, die das Ziel verfolgt, Alternativen zur herkömmlichen Ziffernzensur zu untersuchen und gegebenenfalls in der Hochschule einzusetzen. Dieses Vorgehen stellt eine Maßnahme des Qualitätsmanagements der Pädagogischen Hochschule dar. Der vorliegende Artikel berichtet über die ersten Resultate dieser Untersuchung auf der Basis eines Diskussionsprozesses und eines Literaturreviews.

2. Kritik der Ziffernzensur – Warum?

„So folgen die Menschen in ihren Wertigkeiten skalenmäßig aufeinander, und die Einteilung in wertvolle, minderwertvolle und geringwertige Individuen vollzieht sich nach den einfachsten Gesichtspunkten. Dabei ist für die ganze Betrachtungsweise charakteristisch, daß jede Minderleistung als eine Art Verschulden, für das man den Menschen verantwortlich macht, empfunden wird, als eine Undankbarkeit der Gnade, Mensch sein zu dürfen.“
K. Zeidler, 1928

„I remember the first time that a grading rubric was attached to a piece of my writing. Suddenly all the joy was taken away. I was writing for a grade – I was no longer exploring for me. I want to get that back. Will I ever get that back?“

Claire, eine Studentin; nach K. Olson, 2006

Die Forschung zeichnet ein ziemlich eindeutiges Bild, was Ziffernzensur betrifft:

- Testleistung und tatsächliche Zensur stehen in der Praxis in einem sehr vagen Zusammenhang (fehlende Objektivität der Beurteilung; z.B. Roeder, Baumert, Sang & Schmitz, 1986; siehe auch Schmidinger & Vierlinger, 2012, S. 119ff).
- Die fehlende Objektivität der Ziffernzensur betrifft alle Segmente institutionalisierter Bildung und damit folglich auch den tertiären Bereich (vgl. Schweitzer, 1996).
- Die Ziffernzensur hat – über den Aspekt mangelnder Gütekriterienkonformität hinaus (Bohl, 2009a, S. 73) – auch psychologisch gesehen negative Wirkungen (z.B. Einfluss auf Motivation; vgl. Kohn, 2011; „Grades tend to reduce the quality of students´ thinking.“; ebd., S. 29).
- Es gibt Alternativen, die gut ausgearbeitet und erprobt sind (vgl. Bohl, 2009a, S. 132ff) und zumindest ansatzweise auch schon beforscht werden (z.B. Bos, Beutel, Berkemeyer & Schenk, 2010).
- Innovative Ansätze der Leistungsbeurteilung können nicht losgelöst von der Frage nach förderlichen Prozessen der Leistungsfeststellung gedacht und bewertet werden (ebd.).

Es gibt also eine Reihe von alternativen Ansätzen der Leistungsfeststellung bzw. -bewertung, wie z.B. schriftliche Bewertung, Kompetenzraster, Portfolio, Direkte

Leistungsvorlage (DLV; Schmidinger & Vierlinger, 2012) usw., die aber wiederum alle hinsichtlich ihrer Stärken und Schwächen zu analysieren sind. Zudem findet man im Diskurs und in der Praxis der Leistungsbewertung differenzierende Modi, die auf weniger als fünf Bewertungskategorien bauen, wie z.B. pass/fail grading (zwei Stufen; mit/ohne Erfolg teilgenommen). Hier ist zu klären, inwiefern bei einem pass/fail-Modus paradigmatische oder konzeptuelle Unterschiede zur klassischen fünfstufigen Ziffernbeurteilung tatsächlich vorhanden sind und ob es hinsichtlich der Wirkung einer solchen Beurteilungsform wissenschaftlich nachgewiesene Unterschiede gibt.

3. Innovative Grading im Sinne des pass/fail-Ansatzes

Erste vergleichende Forschungsergebnisse zeigen auf, dass es tatsächlich Unterschiede zwischen pass/fail grading und mehrstufigen Beurteilungsformen gibt: Pass/fail grading reduziert Wettbewerb und fördert Kollaboration, erhöht intrinsische Motivation (White & Fantone, 2009), schafft mehr positive Erfahrungen mit Leistungsfeststellung sowie einen authentischeren Zugang zur Auseinandersetzung mit Inhalten (Dahlgren, Fejes, Abrandt-Dahlgren, & Trowald, 2009). Weiters wird von einer Erhöhung des Wohlbefindens und von Stressreduktion berichtet (ebd.; Rohe, Barrier, Clark, Cook, Vickers, & Decker, 2006). Betreffend Lernoutput, der mittels standardisierten Inventaren gemessen werden kann, sind bisher Untersuchungen im Rahmen der medizinischen Ausbildung durchgeführt worden. Diese sind jedoch noch nicht eindeutig und sprechen von leicht sinkender (McDuff, McDuff, Farace, Kelly, Savoia, & Mandel, 2014) bzw. gleichbleibender Performance (ebd.; Bloodgood, Short, Jackson, & Martindale, 2009).

Neben der Berücksichtigung erster Forschungsergebnisse ist auf theoretischer Ebene zu klären, ob sich ein fünfstufiges Beurteilungssystem von einem zweistufigen tatsächlich auch paradigmatisch-konzeptuell unterscheidet. Dass eine Umstellung von fünf Bewertungsstufen (sehr gut, gut, befriedigend, genügend, nicht genügend) auf zwei Bewertungsstufen (pass/fail grading) etwas „Neues“ oder „Anderes“ darstellt, ist nämlich auf den ersten Blick nicht ad hoc plausibel. Eine nähere Analyse lässt aber erkennen, dass die tatsächliche Veränderung nicht ausschließlich in einer Reduktion der Kategorien besteht, sondern in einer Umstellung von einem Zwei-Ebenen-Kategoriensystem auf ein Ein-Ebenen-Kategoriensystem. Denn auch die herkömmliche fünfstufige Form besteht eigentlich aus zwei Hauptkategorien (pass/fail), von denen aber eine Kategorie (pass) weiter differenziert – sozusagen subkategorisiert – ist (sehr gut, gut, befriedigend, genügend). Eine Umstellung auf pass/fail grading hat zumindest hier den Vorteil, dass eine schwer objektivierbare und wissenschaftlich unzureichend begründbare Stufung der Qualität erbrachter Leistungen nicht mehr im Zentrum der Leistungsrückmeldung steht. Ein pass/fail-Modus – insofern dessen Bewertungsgrundla-

ge nicht summative Lernergebnisse oder Scores bei punktuellen Prüfungen sind – ließe ferner eine rasterbasierte Beschreibung erbrachter Kompetenzen zu. Die Einigung darüber, welche Kompetenzen erreicht werden müssen, um zu einem erfolgreichen Abschluss (pass) zu kommen, ist u.a. aus dem Curriculum abzuleiten. Im besten Fall ist dieses Kompetenzraster mit den Studierenden nicht nur besprochen, sondern auch partizipativ geprüft und gegebenenfalls weiterentwickelt worden (partizipativer Ansatz von Leistungsfeststellung und Leistungsbewertung; siehe auch Kohn, 2001, S. 31). Die Darlegung erworbener Kompetenzen geschieht nicht mehr punktuell durch Prüfungen oder – wie im herkömmlichen Modus – summativ, sondern formativ im Zuge der gesamten Lehrveranstaltung.

Und nicht zuletzt böte eine kompetenzrasterbasierte, formative Leistungsfeststellung mehr Transparenz (vgl. Neuweg, 2014) für die Gruppe der Lehrenden wie auch die betroffenen Studierenden. „Der Schüler soll selbst kontrollieren können, wo er steht, was er noch üben muss und wo er Erfolg hat. [Im übertragenen Sinne würde das auch für Studierende gelten; Anm.d.V.]“ (Singer, 1981, S. 200).

4. Vision Statement

Im deutschsprachigen wie auch internationalen Raum findet Forschung zur Wirksamkeit innovativer Konzepte zur Leistungsfeststellung und Leistungsbewertung statt, tatsächlich aber nur in einem sehr bescheidenen Ausmaß (vgl. Bohl, 2009b, S. 492). Ferner wird davon ausgegangen, dass eine erfolgreiche Implementierung innovativer Ansätze an Schulen nur im Zusammenwirken mit den tertiären Institutionen erfolgen kann (vgl. Schmidinger & Vierlinger, 2012, S. 107). Dieser Appell an die Lehrerinnen- und Lehrerausbildung kann zum einen bedeuten, dass dort die diesbezügliche Begleitforschung verstärkt wird. Zum anderen wird es aber auch nötig sein, dass diese Institutionen selber innovative Formen der Leistungsfeststellung und Leistungsbewertung praktizieren, beforschen, weiterentwickeln und mit Studierenden gemeinsam zum Thema machen.

Angesichts dieser Überlegungen sowie der weiter oben angeführten Evidenzen ist an der Privaten Pädagogischen Hochschule der Diözese Linz von der Interessensgemeinschaft Innovative Grading IG² folgende Forschungsprojektinitiative angedacht (Implementierung in zwei Phasen):

- Projektphase 1: Umstellung der Beurteilung vom 5-stufigen System zum pass/fail-System in jenen Lehrveranstaltungen, in denen die Veranstaltungsleiter/-innen diesbezüglich bereits ihr Interesse bekundet haben, mit optionaler gleichzeitiger Umstellung der Leistungsbewertung in diesen Veranstaltungen auf einen kompetenzrasterbasierten, formativen und partizipativen Modus. Für die betroffenen Lehrenden werden entsprechende Fortbildungen und Coachings angeboten (z.B. zur Theorie des innovative grading; zur Praxis der Umsetzung; zur Begleitung und Bewertung von

- Lernprozessen; siehe auch Svoboda & Vogl 2017; zur kompetenzorientierten Leistungsfeststellung; siehe auch Gerick, Sommer & Zimmermann 2018). Inwiefern auch Modi der Selbstbewertung (vgl. Bohl, 2009a, S. 124ff; Reitinger, 2013, S. 155f) implementiert werden, ist noch zu diskutieren. Die diesbezüglichen Erfahrungen werden dokumentiert, im Sinne von Aktionsforschung begleitet und postaktional (also nach dem Lehrveranstaltungs-ende) – gemeinsam mit betroffenen Studierenden – in der Interessensgemeinschaft Innovative Grading IG² diskutiert. Zentrale Fragestellungen in dieser Phase sind u.a., inwiefern Studierende a) den Lernprozess und b) das persönliche Lernergebnis anders wahrnehmen als gewohnt. Die abhängigen Variablen sind Wahrnehmung von Wettbewerb, von Stress, von Angst, Wohlbefinden und Kollaboration. Ferner wird intrinsische Motivation, Erfahrung mit Leistungsfeststellung, Authentizität des Zugangs zu Lerninhalten und der Lernoutput erhoben. Zusätzliche abhängige Variablen werden im Verlauf der weiteren Diskussion in der Interessensgruppe ergänzt.
- Projektphase 2: Weiterführung des pass/fail-Systems unter Berücksichtigung der Ergebnisse und Diskurse aus der Projektphase¹ mit gleichzeitiger multimethodischer Begleitforschung in quasi-experimentellen Forschungsdesigns, um Unterschiede zum herkömmlichen Beurteilungssystem zu überprüfen.

Literatur

- Bloodgood, R. A., Short, J. G., Jackson, J. M., & Martindale, J. R. (2009). A Change to Pass/Fail Grading in the First Two Years at One Medical School Results in Improvement of Psychological Well-Being. *Academic Medicine*, 84(5), 655–662.
- Bohl, T. (2009). *Prüfen und Bewerten im Offenen Unterricht* (4. Aufl.). Weinheim und Basel: Beltz.
- Bohl, T. (2009b). Veröffentlichung und Würdigung von Unterrichtsergebnissen. In K.-H. Arnold, U. Sandfuchs & J. Wichmann (Hrsg.), *Handbuch Unterricht* (S. 490–493). Bad Heilbrunn: Klinkhardt.
- Bos, W., Beutel, S.-I., Berkemeyer, N., & Schenk, S. (2010). *LUZI. Leistungsbeurteilung ohne Ziffernzeugnisse. Abschlussbericht der wissenschaftlichen Begleitforschung*. Technische Universität Dortmund.
- Gerick, J. Sommer, A., & Zimmermann G. (Hrsg.) (2018). *Kompetent Prüfungen gestalten*. Münster: Waxmann.
- Kohn, A. (2011). The Case Against Grades. *Educational Leadership*, 69(3), 28–33.
- Dahlgren, L. O., Fejes, A., Abrandt-Dahlgren, M., & Trowald, N. (2009). Grading systems, features of assessment and students' approaches to learning. *Teaching in Higher Education*, 14(2), 185–194.

- McDuff, S. G. R., McDuff, D., Farace, J. A., Kelly, C. J., Savoia, M. C., & Mandel, J. (2014). Evaluating a grading change at UCSD school of medicine: pass/fail grading is associated with decreased performance on precinical exams but unchanged performance on USMLE step 1 scores. *BMC Medical Education*, 14, 1–9.
- Neuweg, G. H. (2014). *Schulische Leistungsbeurteilung. Rechtliche Grundlagen und pädagogische Hilfestellungen für die Schulpraxis*. Linz: Trauner.
- Reitinger, J. (2013). *Forschendes Lernen. Theorie, Evaluation und Praxis*. Immenhausen bei Kassel: Prolog.
- Roeder, P. M., Baumert, J., Sang, F., & Schmitz, B. (1986). Über die Zusammenhänge zwischen Zensur und Testleistung. In H. Pettilon, J. W. Wagner & B. Wolf (Hrsg.), *Schülergerechte Diagnose* (S. 31–60). Weinheim: Beltz.
- Rohe, D. E., Barrier, P. A., Clark, M. M., Cook, D. A., Vickers, K. S., & Decker, P. A. (2006). The benefits of pass-fail grading on stress, mood, and group cohesion in medical students. *Mayo Clinical Proceedings*, 81(11), 1443–1448.
- Schmidinger, E., & Vierlinger, R. (2012). *Zeitgemäße Leistungsbeurteilung*. Wien: Jugend & Volk.
- Schweitzer, W. (1996). *Gute und sehr gute Noten in Wirtschaftswissenschaft. Nachrichten und Berichte*. Universität Passau.
- Singer, K. (1981). *Maßstäbe für eine Humane Schule*. Frankfurt: Fischer.
- Svoboda, U., & Vogl, U. (2017). Lernen beraten. *Pädagogische Horizonte*, 1(1), 123–137.
- White, C. B., & Fantone, J. C. (2010). Pass-fail grading: laying the foundation for self-regulated learning. *Advances in Health Sciences Education*, 15, 469–477.