

## Kriteriumsorientierte adaptive Klausuren an Hochschulen

Aron Fink<sup>1</sup>, Sebastian Born<sup>1</sup>, Andreas Frey<sup>1,2</sup>, Christian Spoden<sup>1</sup>,

<sup>1</sup>Lehrstuhl für empirische Methoden der erziehungswissenschaftlichen Forschung, Friedrich-Schiller-Universität Jena, Deutschland

<sup>2</sup>Centre for Educational Measurement (CEMO), University of Oslo, Norwegen

E-Mail: aron.fink@uni-jena.de, sebastian.born@uni-jena.de, andreas.frey@uni-jena.de, christian.spoden@uni-jena.de

Das Ausmaß, mit dem Studierende die Studienziele ihres Hochschulstudiums erreichen, wird im Kern am Abschneiden bei Prüfungen festgemacht. In der Europäischen Union erfolgt die Spezifikation der Studienziele seit Inkrafttreten der Bologna-Reform flächendeckend als Kompetenzen, die man nach Klieme und Leutner als domänenspezifische kognitive Leistungsdispositionen definieren kann [1]. Im Sinne der Kompetenzorientierung ist es daher nötig, dass auch die an Hochschulen durchgeführten Prüfungen dahingehend angepasst werden, dass sie überprüfen, über welche Kompetenzen die Studierenden am Ende eines Moduls verfügen. Da ohne erfolgreich absolvierte Prüfungen ein Studienabschluss nicht erlangt werden kann, kommen Prüfungen eine hohe persönliche Relevanz für die Studierenden zu. Dieser hohen Bedeutung werden insbesondere schriftliche Prüfungen in Form von Klausuren typischerweise nicht gerecht. Vielmehr klafft hier eine erhebliche Lücke zwischen dem an deutschen Hochschulen üblicherweise praktizierten Vorgehen und dem wissenschaftlichen Kenntnisstand der relevanten Teildisziplinen (v. a. Educational Measurement und Psychometrie).

Problematisch an derzeitigen Hochschulklausuren sind im Wesentlichen vier Aspekte: 1. Die in den Modulkatalogen verankerten, kompetenzorientierten Lernziele werden durch die Klausuraufgaben meistens nicht angemessen abgebildet. Zur Lösung der Problematik müsste zunächst der Messgegenstand der jeweiligen Klausur eindeutig spezifiziert und anschließend angemessen durch Klausuraufgaben operationalisiert werden. 2. Klausuren sind überwiegend als normorientierte beziehungsweise sich an willkürlich festgelegten Kriterien orientierende Verfahren konzipiert. Um Klausurergebnisse als Ausmaß des Erreichens kompetenzorientierter Lernziele interpretieren zu können, sollten kriteriumsorientierte Tests zum Einsatz kommen, wobei die jeweiligen Kriterien über sogenannte Standard Setting Verfahren festgelegt werden. 3. Klausuren sind zwischen verschiedenen Testzeitpunkten (z. B. Jahren) nicht statistisch miteinander verbunden, so dass Unabhängigkeit der vergebenen Bewertungen (z. B. Noten) von Kohortenleistungsfähigkeit und Klausurschwierigkeit nicht gewährleistet sind. Dieses Problem kann vermieden werden, indem mit Modellen der Item Response Theory (IRT; z. B. [2]) Klausuren über Testzeitpunkte verbunden werden, so dass individuelle Klausurergebnisse immer auf der selben Metrik verortet werden und so über die Klausurzeitpunkte hinweg vergleichbar bleiben. 4. Übliche Klausuren messen im mittleren

Leistungsbereich am präzisesten, wohingegen die Messpräzision in den Randbereichen der Merkmalsverteilung stark abnimmt [3]. Eine Angleichung der Messpräzision über den gesamten Merkmalsbereich kann mit computerisierten adaptiven Testen (CAT; z. B. [4]) erreicht werden.

Im Vortrag wird erörtert, wie diese Probleme durch die Kombination der angesprochenen Methoden aus den Bereichen Educational Measurement und Psychometrie gelöst werden können. Es wird sich dabei an modernen Ansätzen der theoriebasierten Konstruktion von Kompetenztests [5], die beispielsweise bei internationalen Schulleistungsstudien wie PISA und oder dem Ländervergleich verwendet werden, orientiert. Konkret wird ein hochschul- und fächerübergreifendes Konzept zur Konstruktion, Administration und Auswertung kriteriumsorientierter adaptiver Klausuren an Hochschulen vorgestellt.

*Keywords:*

E-Klausuren, Kompetenzdiagnostik, Computerisiertes Adaptives Testen

*Quellen:*

- [1] Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. Zeitschrift für Pädagogik, 52, S. 876–903.
- [2] van der Linden, W. J. (Hrsg.). (2016a). Handbook of item response theory. Volume one: Models. Boca Raton: Chapman & Hall/CRC.
- [3] Dolan, R. P., & Burling, K. S. (2012). Computer-based testing in higher Education. In: C. Secolsky & D. B. Denison (Eds.), Handbook on measurement, assessment, and evaluation in higher education. S. 312-335. New York, NY: Routledge.
- [4] Frey, A. (im Druck). Computerisiertes adaptives Testen. In H. Moosbrugger & A. Kelava (Hrsg.), Testtheorie und Fragebogenkonstruktion (3. aktualisierte und überarbeitete Auflage. Berlin, Heidelberg: Springer.
- [5] Frey, A., & Hartig, J. (im Druck). Kompetenzdiagnostik. In M. Gläser-Zikuda, Harring, M., & Rohlf, C. (Hrsg.), Handbuch Schulpädagogik. Münster: Waxmann.