

## **Methodenbericht – Auswirkungen einer computerisierten teiladaptiven Klausur auf das emotionale Erleben von Klausuren seitens Studierender**

Ziel der Studie war die Untersuchung der Auswirkungen einer computerisierten, teiladaptiven Klausur im Vergleich zu herkömmlichen Klausuren auf das emotionale Erleben von Klausuren seitens Studierender. Eine zentrale theoretische Grundlage für die Evaluationsstudie stellt die Kontrolle-Wert-Theorie (Pekrun, Frenzel & Goetz, 2007) dar.

### **Stichprobe**

Die Stichprobe bestand aus  $N = 94$  Studentinnen und Studenten im ersten Semester (83 % Frauen) mit dem Studienschwerpunkt "Erziehungswissenschaft" (Bachelor-Studiengang; 56 % mit Nebenfach Psychologie) der FSU Jena. Insgesamt wurden drei Erhebungen durchgeführt (T1 (Baseline) = eine Woche vor der Klausur, T2 = direkt vor der Klausur, T3 = direkt nach der Klausur). 84 Personen haben die Fragebögen zu T2 und T3 bearbeitet, 56 Personen alle drei.

### **Instrumente**

#### *Intervention – Computerisierte teiladaptive Klausur*

Die computergestützte teiladaptive Klausur zur Vorlesung „Statistik“ für Erziehungswissenschaftler bestand aus 5 Itemclustern mit je 5 dichotom kodierten Items (= Testlänge von 25 Items), die in einem Balanced Latin Square Design über Clusterpositionen und –reihenfolgen ausbalanciert wurden, was in insgesamt zehn Testversionen resultierte. In einem der Cluster wurden die Items adaptiv aus einem bestehenden Itempool mit bekannten IRT-Schwierigkeitsparametern gewählt. Die Prüfung wurde in den PC-Pools der Universität durchgeführt und über den SaveExamBrowser (ETH Zürich, 2019) und ein eigens entwickeltes Softwarepaket (KAT-HS App, Fink, Spoden, Naumann & Frey, 2019) ausgeliefert. Der computerbasierte Testalgorithmus wurde wie folgt konfiguriert: Die adaptive Itemauswahl wurde nach dem Kriterium der maximalen Information realisiert (Lord, 1980). Zur Schätzung der Itemparameter wurde die MML-Schätzung (Bock & Aitkin, 1982) und zur Schätzung der Fähigkeitsparameter der WLE (Warm, 1989) verwendet. Die Studierenden hatten 90 Minuten für die Bearbeitung der Klausur. Um zur nächsten Aufgabe zu gelangen, mussten die Studentinnen und Studenten das vorhergehende Item beantworten (forced choice). Somit konnten sich fehlende Werte nur am Ende der Klausur ergeben. Itemreview war nicht erlaubt. Um zu vermeiden, dass Studierende zu eilig arbeiten bzw. nicht alle Aufgaben bearbeiten, wurde während der Klausur kontinuierlich die durchschnittliche Bearbeitungszeit mit der geschätzten für die übrigen Items noch benötigten Zeit abgeglichen, und zurückgemeldet, ob die Studierenden zu schnell bzw. zu langsam arbeiteten. Dieses Vorgehen führte dazu, dass alle Studierenden die Klausur vollständig bearbeiteten. Fehlende Werte ergaben sich somit allein aus der teiladaptiven Itemvorgabe. Die Reliabilität der Klausur betrug  $rel_{WLE} = .722$ .

#### *Fragebogen*

Als abhängige Variablen wurden das emotionale Erleben während der Klausur und das subjektive Kontrollempfinden mittels Online-Fragebogen erhoben. Als Kontrollvariablen

diente der persönliche Wert der Klausurleistung. Zur Erfassung der Emotionen während der Testbearbeitung wurde der Test Emotion Questionnaire in deutscher Sprache (TEQ; Pekrun, Goetz, Perry, Kramer, Hochstadt, & Molfenter, 2004) genutzt. Die Stimuli wurden so formuliert, dass sich die Items auf eine typische Prüfung in T1 und T2 und die gerade abgeschlossene Prüfung in T3 bezogen. Die wahrgenommene akademische Kontrolle wurde mit dem Perceived Academic Control Questionnaire (PAC; Molfenter, 1999; Perry, Hladkyj, Pekrun, & Pelletier, 2001) erhoben, der aus 8 Items bestand. Die Items wurden so formuliert, dass sie sich auf typische Prüfungen in T1 und T2 und auf die gerade abgeschlossene Prüfung in T3 bezogen. Der subjektive Wert wurde durch ein eigens entwickeltes Instrument zu T2 erhoben.

## Literatur

- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: an application of an EM algorithm. *Psychometrika*, *46*, 443–459. doi: 10.1007/BF02293801
- ETH Zürich, Lehrentwicklung und -technologie (2019). Safe Exam Browser (Version 2.2.3) [Computer Software]. Zürich: ETH Zürich, Lehrentwicklung und -technologie (LET).
- Fink, A., Spoden, C., Naumann, P. & Frey, A. (2019). KAT-HS APP (Beta-Version) [Computer Software]. Frankfurt am Main: Goethe-Universität Frankfurt. Verfügbar unter: <https://kat-hs.uni-frankfurt.de/materialien/software/>
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- Molfenter S. (1991). Prüfungsemotionen bei Studierenden [Dissertation]. Regensburg: Universität Regensburg.
- Pekrun, R., Frenzel, A. C., & Goetz, T. (2007). The Control-Value Theory of Achievement Emotions: An Integrative Approach to Emotions in Education. In P. A. Schutz, & R. Pekrun (Eds.), *Emotion in education* (pp. 13-36). Amsterdam: Academic Press. <https://doi.org/10.1016/B978-012372545-5/50003-4>
- Pekrun, R., Goetz, T., Perry, R.P., Kramer, K., & Hochstadt, M., & Molfenter, S. (2004). Beyond test anxiety: development and validation of the Test Emotions Questionnaire (TEQ). *Anxiety, Stress and Coping*, *17*, 287-316. <https://www.doi.org/10.1080/10615800412331303847>
- Perry, R. P., Hladkyi, S., Pekrun, R., & Pelletier, S. (2001). Academic control and action control in college students: A longitudinal study of self-regulation. *Journal of Educational Psychology*, *93*, 776-789. <https://www.doi.org/10.1037/0022-0663.93.4.776>
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54*, 427–450. doi: 10.1007/bf02294627