

# Wie unfair! Die neuen Prüfungsaufgaben sind doch viel schwieriger als die alten – ein empirischer Lösungsansatz

Joy Backhaus<sup>1</sup>, Herbert Pointingl<sup>2</sup>, Chantal Rabe<sup>1</sup>, Sarah König<sup>1</sup> & Eva Hennel<sup>3</sup>

<sup>1</sup>Institut für Medizinische Lehre und Ausbildungsforschung Würzburg, Deutschland

<sup>2</sup>Psychologische Fakultät, Universität Wien, Österreich

<sup>3</sup>Lehrklinik der Medizinischen Fakultät Würzburg, Deutschland

E-Mail: [Backhaus\\_J@ukw.de](mailto:Backhaus_J@ukw.de), [herbert.pointingl@gmail.com](mailto:herbert.pointingl@gmail.com), [Rabe\\_Chantal@ukw.de](mailto:Rabe_Chantal@ukw.de), [Koenig\\_Sarah@ukw.de](mailto:Koenig_Sarah@ukw.de) & [Hennel\\_E@ukw.de](mailto:Hennel_E@ukw.de)

## Fragestellung/Zielsetzung

Prüfungsaufgaben müssen aus unterschiedlichen Gründen überarbeitet werden, wie zum Beispiel schlechte psychometrische Statistiken oder Bekanntheit der Prüfungsaufgaben bei den Teilnehmern. Die Erstellung neuer Aufgaben ist ein langwieriger und arbeitsintensiver Prozess. Umso wichtiger ist es, dass die Vergleichbarkeit zwischen den neu erstellten und bereits etablierten Aufgaben auf einfache Weise sichergestellt werden kann. In diesem Beitrag werden Ergebnisse eines Anwendungsbeispiels dargestellt. Im Rahmen der Objective Structured Clinical Examinations (OSCE) wurde geprüft, ob sogenannte Brückenaufgaben („linked items“) ein geeignetes Mittel zur Bewertung der Prüfungsaufgabenleistungen zwischen den Semestern darstellen.

## Material/Methoden

Bei den OSCEs wurde ein „linked item“ – Ansatz verwendet um insgesamt 427 Aufgaben (WS 2015/2016 & SS 2017) mittels Raschmodell zu analysieren. Dabei wurden Aufgaben, die bei allen zu untersuchenden Prüfungen vorgegeben wurden, als sogenannte Brückenaufgaben oder Linkaufgaben verwendet („anchored-item-design“). Diese verlinkten Aufgaben ermöglichen mit Raschmodell-Analysen die Aufgabenschwierigkeiten für den kompletten Aufgabensatz (bestehend aus etablierten und neu erstellten Aufgaben) und die Prüfungsaufgabenleistungen (über Schätzung eines Personenparameters, der die Prüfungsaufgabenleistung widerspiegelt) für alle Teilnehmer zu bestimmen.

## Ergebnis

Nach einer Überprüfung auf Raschmodell-Konformität mittels Modelltests ist ein fairer Vergleich zwischen den Aufgaben als auch zwischen den Prüfungsaufgabenleistungen aller Teilnehmer über die verschiedenen Prüfungstermine möglich. Die Parameterschätzungen für die Aufgaben und die Prüfungsaufgabenleistungen der Teilnehmer zeigen ein valides Raschmodell. Durch die Gültigkeit des RM-Modells ist die Vergleichbarkeit zwischen den Aufgaben als auch zwischen den verschiedenen Teilnehmern (unabhängig aus welchem Semester) gegeben.

## Take-Home-Message

Die Verwendung von Brückenaufgaben ist eine zuverlässige Methode um faire und effiziente Vergleiche zwischen Aufgaben aus verschiedenen Semestern durchführen zu können.

Gleichzeitig können auch die Prüfungsleistungen der Teilnehmer aus den verschiedenen Semestern fair und zuverlässig miteinander verglichen werden.

### **Unsere Intention im Knowledge Cafe**

Im Rahmen des Knowledge-Cafes möchten wir den Teilnehmern anbieten primär praktische Aspekte des Ansatzes zu diskutieren:

1. Wie viele Brückentests benötige ich, wenn ich diesen Ansatz implementieren möchte.
2. Wie wurde der Ansatz von den Studierenden und Prüfern aufgenommen.
3. Wie sind wir in den Folgesemestern verfahren.

#### *Keywords:*

Probabilistische Testtheorie, anchored-item-design, Fairness, Medizin

#### *Quellen:*

Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement*, 14(2), 97-116.

Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.