

Paulitsch, Michael A¹; Sennekamp, Monika¹; Klingebiel, Thomas²

¹ Dekanat am Fachbereich Medizin der Goethe-Universität; ² Zentrum der Kinder- und Jugendmedizin des Universitätsklinikums FFM

Messfehler

Verfahren zur Überprüfung von Messfehlerfreiheit von OSCE-Ergebnissen sind weit verbreitet

Interne Konsistenz

- Bestimmt den Messfehler bei querschnittlicher Erhebung (über Stationen oder Beurteilungsbögen)
- Dabei Annahme von Itemhomogenität (anhand von Faktorenanalyse vorab überprüft)
- Bei OSCEs: Keine Itemhomogenität vorhanden, daher keine Annahme von interner Konsistenz!

Generalisierbarkeit

- Bestimmt die Verlässlichkeit von Messungen über verschiedene Bedingungen hinweg
- Voraussetzung 1: Stabilität des gemessenen Merkmals
- Voraussetzung 2: Testwert-Unterschiede eines Individuums sind in verschiedenen Messsituationen von einer oder mehreren Fehlerquellen abhängig und nicht von systematischen Veränderungen aufgrund von Reife oder Lernen (Wurzeln in der Persönlichkeitspsychologie)
- Bei OSCEs: Testwertunterschiede aufgrund verschiedener plausibler Quellen (unterschiedliches Fähigkeitsniveau der Prüflinge, unterschiedliche Schwierigkeitsniveaus der Stationen, Interaktion von Prüflingen und Stationen), allerdings durch Rater bedingte Varianz unerwünscht. Daher keine Aussage über Zuverlässigkeit eines Ergebnisses, da Variation über verschiedene Bedingungen plausibel.

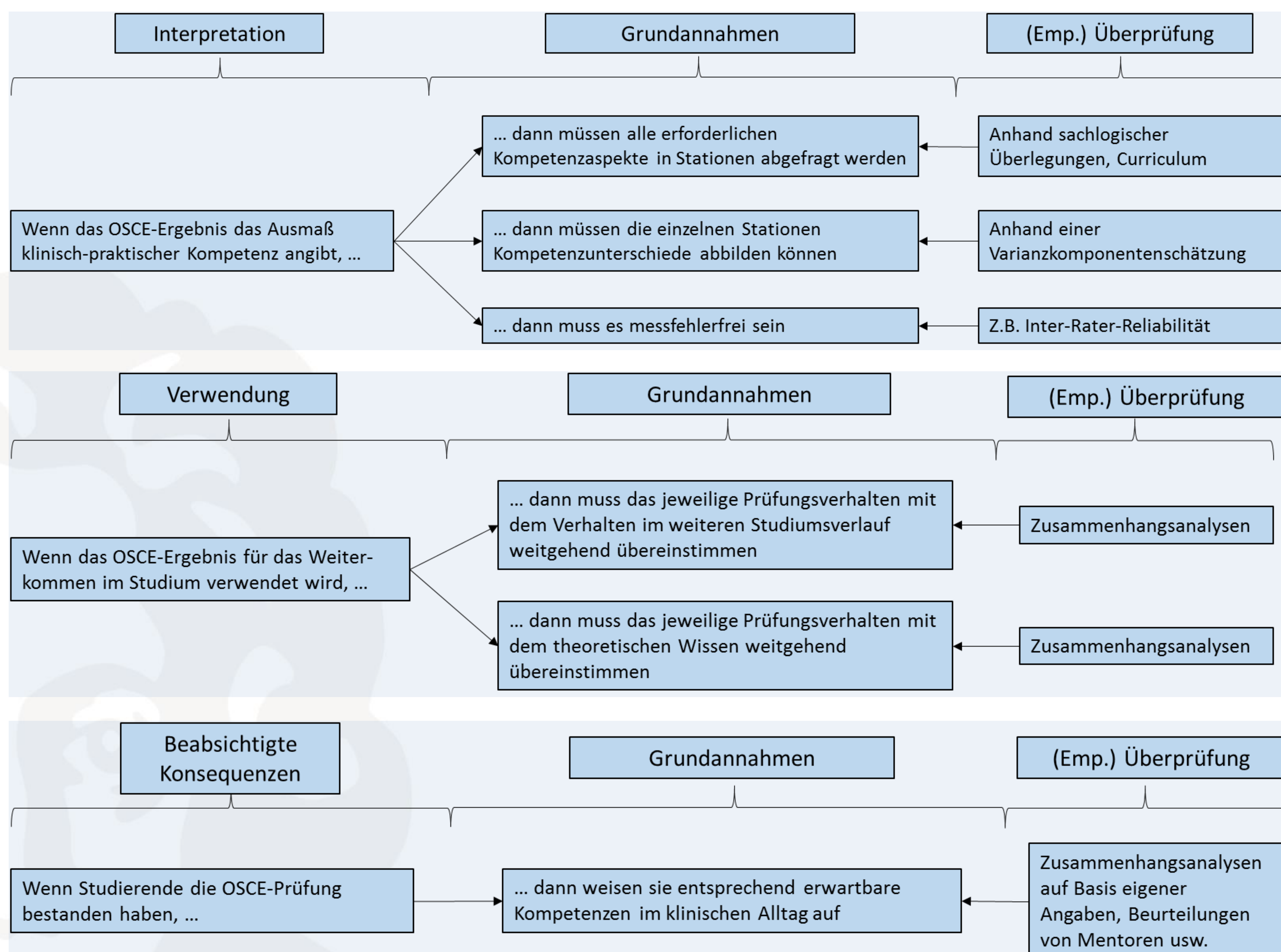
➔ **Aber Güte und Messfehlerfreiheit von OSCE-Ergebnissen grundsätzlich wichtig, da relevante Konsequenzen!**
Daher Vorschlag eines neuen Rahmenmodells zur Validitätsüberprüfung.

Validität

Wandel des Validitätsverständnisses

- Validität als Testeigenschaft ➔ Validität der Testergebnis-Interpretationen, Verwendungen und beabsichtigter Konsequenzen
- Verschiedene Validitätsarten ➔ Einheitliches Validitätskonzept
- Betrachtung von Einzelbefunden ➔ Kohärente Validitätsargumentation (Integration aller Befunde, auch Messfehlerbestimmung)

Validitäts-Argumentation



Schlussfolgerung

Nutzung eines argumentationsbasierten Ansatzes bietet folgende Vorteile

- Integration des aktuellen Validitätsverständnisses
- Keine Notwendigkeit einer starken Theorie wie bei Konstruktvalidität
- Spezifisch auf konkrete Anforderungen ausgerichtet

Literatur

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014). Standards for educational and psychological testing.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73.
- Schermelleh-Engel, K., & Werner, C. (2008). Methoden der Reliabilitätsbestimmung. *Testtheorie und Fragebogenkonstruktion*, 113-133.
- Shavelson, R. J., & Webb, N. M. (1991). Generalizability theory: A primer. Sage.