



Objektive Messung von Prüferverhalten im OSCE – Psychometrie des Mannheimer Feedbackbogens für OSCE- Prüfer

Katrin Schüttpelz-Brauns*, Elvira Pippel,
Helene Todtenhaupt, Kathrin Nühse



GEFÖRDERT VOM



Einführung

- Prüfertrainings zur Qualitätssicherung von OSCE¹⁻⁵
- Beispiel: Ma-TOP⁶ als Blended-Learning-Format
 - E-Learning: Allgemeiner Teil zu OSCE und Prüferverhalten
 - Vor-Ort-Video: Spezieller Teil zu eigener OSCE-Station
 - Strukturierte Beobachtung inkl. Feedback: Transfer auf Prüfungssituation
- Problem bei Überprüfung der Wirksamkeit:
 - fehlende objektive Operationalisierung des Prüferverhaltens

Einführung

- **Entwicklung eines Beobachtungsbogens**
 - Likert-Items + Kommentarfelder für Feedback
 - Inhaltsvalidität:
 - Itempool aus Literatur, Expertenbefragung und Prüferbeobachtung
 - Anwendung der Items bei Prüferbeobachtung
 - Überarbeitung der Items
 - Bilden von Kategorien
 - Anwendung der Items bei Prüferbeobachtung
 - Überarbeitung der Items...
- Ziel: Eindeutigkeit und Relevanz der Items

Einführung

- **Mannheimer Feedbackbogen (Ma-FOP)**
 - Formale Aspekte (N=4)
 - Aufmerksamkeit (N=3)
 - Neutralität (N=4)

- **Anwendung**
 - Feedback für die Transferphase des Ma-TOP
 - Quantifizierung des Prüferverhaltens

Fragestellung

- Skalen in Empirie nachvollziehbar?
- Beobachterübereinstimmung
- Reliabilität der Skalen

Methode

- **Stichprobe**
 - 20 Prüfer
 - Einverständnis zur Beobachtung von allen Prüfern

- **Material**
 - Ma-FOP (3 Skalen mit jeweils 3 bzw. 4 Likert-Items in Schulnotenskalierung)

Methode

- **Durchführung**
 - 4 OSCE mit 4 bzw. 6 bzw. 9 Stationen (BP Chirurgie, klin. Praxis, KDP) im Zeitraum April bis Juli 2014
 - 2 Beobachter
 - Beobachtung mit Ma-FOP ohne Feedback
 - 2 Situationen je Prüfer
 - Beide Beobachter gleichzeitig
 - = 40 Situationen

Methode

■ **Statistische Analyse**

- empirische Replikation der Skalen:
Mokken-Analyse (R)
 - Skala: $H \geq 0,5^7$
- Beobachterübereinstimmung:
Rangkorrelation nach Spearman (SPSS)
 - Interpretation: $r=0,1$ (klein), $r=0,3$ (moderat), $r=0,5$ (groß)⁸
- Reliabilität:
Cronbachs α (SPSS)
 - Interpretation: $\alpha \geq 0,7$ (akzeptabel) $\alpha \geq 0,8$ (gut)⁹

Ergebnisse

Item	Skala
Formale Aspekte	
- Prüfungszeit	
- Umsetzen der Checklisteninstruktion	
- Ausfüllen der Checkliste	1
- Abstand Prüfungssituation	1
Aufmerksamkeit	
- Handlung während Prüfung sichtbar	2
- Handlung während Prüfungszeit hörbar	2
- Vollständige Konzentration während Prüfungszeit	2
Neutralität	
- Verbale Reaktion	1
- Nonverbale Reaktion	1
- Keine Zwischenfragen	1
- Sachliches Zurückführen*	

Anmerkungen: Skalen ermittelt mit Hilfe der Mokken-Analyse unter der Bedingung: $H \geq 0,5$; * Filterfrage, mit zu wenig Daten um sie in die Mokken-Analyse einzufügen

Ergebnisse

Skala / Item	H	Interrater- Reliabilität ⁺	Reliabilität [§]	
			B1	B2
Skala 1: Prüferverhalten	0,77	0,93***	0,81	0,83
Skala 2: Ablenkung	0,64	0,88***	0,70	0,66

Anmerkung: H: Mokken-Homogenität der Skala, + Rangkorrelation nach Spearman, § Cronbachs Alpha, B1: Beobachter 1, B2: Beobachter*** signifikant mit $p \leq 0,001$

Diskussion



- Zwei Skalen (Prüferverhalten und Ablenkung)
- Sehr hohe Beobachterübereinstimmung
- Akzeptable bzw. gute Reliabilität



- Fehlendes externes Validitätskriterium

Ausblick

- Objektive und zuverlässige Messung des Prüferverhaltens möglich
- Ansatz zur Überprüfung der Wirksamkeit von OSCE-Prüfer-Trainings

Referenzen

1. Chesser A, Cameron H, Evans P, Cleland J, Boursicot K, Mires G. Sources of variation in performance on a shared OSCE station across four UK medical schools. *Med Educ.* 2009; 43: 526-532.
2. Khan KZ, Ramachandran S, Gaunt K, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: an historical and theoretical perspective. *Med Teach.* 2013; 35:e1437-e1446.
3. Newble DI, Hoare J, Sheldrake PF. The selection and training of examiners for clinical examinations. *Med Educ.* 1980; 14:345–349.
4. Pell G, Homer M, Roberts TE. Assessor training: Its effects on criterion based assessment in a medical context. *Int J Res Method Educ.* 2008; 31:143–154.
5. Van Der Vleuten CPM, Van Luyk SJ, Van Ballegooijen AMJ, Swansons DB. Training and experience of examiners. *Med Educ.* 1989; 23:290–296.
6. Nühse K, Braun B, Kaden J, Peters Y, Schüttpelz-Brauns K. Das Mannheimer Training für OSCE-Prüfer – Entwicklung eines Blended Learning Konzepts. In: Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). Hamburg, 25.-27.09.2014. Düsseldorf: German Medical Science GMS Publishing House; 2014. DocV321.
7. Krause B. Zur Konstruktion von messtheoretisch begründeten Ratingskalen. In: Krause B, Hrsg. ZeE-Publikationen. Reihe Empirische Evaluationsmethoden, Band 12. Berlin: ZeE-Verlag;. 2007.
8. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
9. Brosius F. *SPSS 19.* Heidelberg: mitp; 2011.

Vielen Dank für Ihre Aufmerksamkeit!



Kontakt: katrin.schuettpelz-brauns@medma.uni-heidelberg.de