



Kompetenzzentrum für Prüfungen in der Medizin Baden-Württemberg



Medizinische Fakultät Heidelberg



Bundesministerium
für Bildung
und Forschung

MERLIN

Medical Education Research -
Lehrforschung im Netz BW

Numerische Analyse von Prüfereffekten in OSCEs

Andreas Möltner,
Jörn Heid, Stefan Wagener, Jana Jünger

GMA 2017 Münster



Hintergrund

Prüfereffekte:

Einfluss auf die Leistungsbewertung

- Beobachtungsfehler
- Beurteilung der Beobachtungen
- Strukturierung der Prüfung, Beeinflussung der Prüfungskandidaten

Zeitunabhängige und –abhängige Effekte
(z. B. unterschiedliche Strenge bzw.
Ermüdungs- oder Kontrasteffekte)



Hintergrund

Relevanz:

Masterplan 2020: OSCEs oder OSCE-ähnliche Stationsprüfungen in den Staatsexamina der Human- und Zahnmedizin.

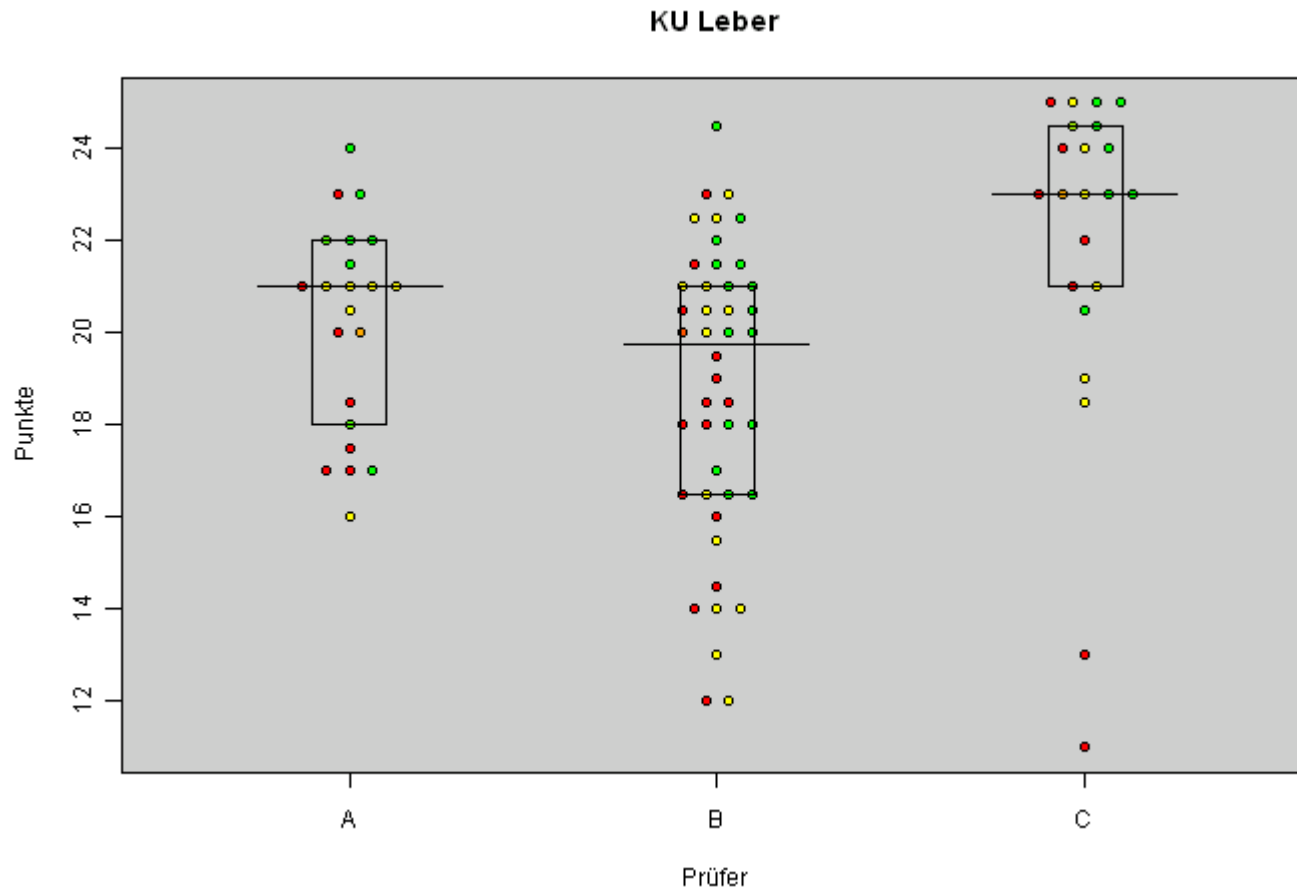
Problem:

Sicherung der Objektivität der Prüfungen



Deskriptive Analyse

Unterschiedliche Strenge, Skalennutzung und Trennschärfe von Prüfern





Ziel: Numerische Bestimmung der Effekte

Graphische Darstellung ist hilfreich als illustratives Feedback an Prüfer, um Konsequenzen der Effekte auf Reliabilität und Genauigkeit der Prüfung abzuschätzen, ist eine quantitative Analyse erforderlich.

Weiterhin sind zur Beurteilung von Prüferleistungen (Akzeptanz als Prüfer?) adäquate Kennwerte erforderlich.



Generalisierbarkeitstheorie

Verfahren zur Analyse des Einflusses von Faktoren („Facetten“) auf die Prüfungsergebnisse.

Aber:

Modellierung additiver Einflüsse möglich, nicht aber z. B. der Skalenausnutzung.



Modell der Analyse I

$$X_{s,t,r} = b_t^T b_r^R (S_s + \varepsilon_t^T + \varepsilon_r^R) + a_t^T + a_r^R$$

$X_{s,t,r}$ Punktwert des Studierenden s bei Aufgabe (Task) r und Prüfer (Rater) r .

S_s „Fähigkeit“ des Studierenden s
 a_t^T, a_r^R Schwierigkeit der Aufgabe t bzw. Strenge des Prüfers r



Modell der Analyse II

$$X_{s,t,r} = b_t^T b_r^R (S_s + \varepsilon_t^T + \varepsilon_r^R) + a_t^T + a_r^R$$

b_t^T, b_r^R

Skalenausnutzung durch
Aufgabe t bzw. durch den
Prüfer r

$\varepsilon_t^T, \varepsilon_r^R$

Unsystematischer Messfehler
durch Aufgabe t bzw. durch
Prüfer r mit Standardabweichung
 σ_t^T bzw. σ_r^R



Modell der Analyse III

Identifizierbarkeit der Parameter:

Parameterschätzungen beruhen auf den mittleren vergebenen Punktzahlen (a_t^T, a_r^R) und den Varianzen und Kovarianzen der vergebenen Punkte $(b_t^T, b_r^R, \sigma_t^T, \sigma_r^R)$ an gleichen Studierenden.



Modell der Analyse III

Identifizierbarkeit der Parameter:

Minimales Design, in welchem alle Parameter schätzbar sind (bis auf Normierungskonstanten)

Parcours 1

T1, R1

T2, R2

T3, R3

Parcours 2

T1, R2

T2, R3

T3, R1



Kontrolle der Prüfer

Mögliches Vorgehen:

Kalibrierung der Stationen: Bestimmung der Stationsparameter $(a_t^T, b_t^T, \sigma_t^T)$ mit hinreichend vielen Prüfern.

Bestimmung der Prüferparameter $(a_r^R, b_r^R, \sigma_r^R)$ anhand der kalibrierten Stationen (Videoaufzeichnungen?).



Was ich Ihnen vorenthalten musste

Beispiele: Parameterschätzungen an empirischen Daten

Auswirkungen von Prüfereffekten auf die Reliabilität und Genauigkeit der Leistungsbeurteilung

Abschätzungen über die Zahl der erforderlichen Datenzahlen für Kalibrierung und Prüferbeurteilungen



Zusammenfassung

Die Modellierung erlaubt:

- Quantitative Kenngrößen für die Prüfereffekte Strengung, Skalennutzung und Trennschärfe
- Abschätzungen für die Auswirkungen der Prüferinflüsse auf die Messzuverlässigkeit und –genauigkeit
- Kalibrierung von Stationen
- Identifikation auffälliger Prüfer