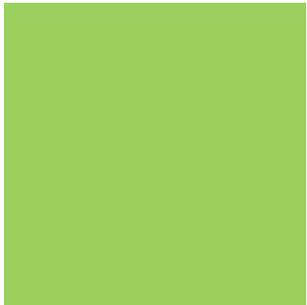
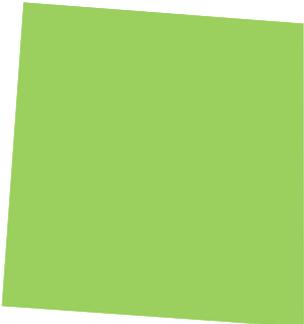


Ist ein Situational Judgement Test zur Messung kognitiver Anteile der Gesprächskompetenz praktikabel für den Einsatz im Medizinstudium?

Claudia Kiessling, Martin R. Fischer



Was möchte ich Ihnen heute präsentieren?

- Theoretischer Hintergrund
- ProfKom Projekt
- Testentwicklung
- Interne Konsistenz der Tests
- Diskussion & Schlussfolgerungen



Kann man Kommunikation schriftlich testen?

Wie sind nicht die ersten, die es versucht haben...

Humphris. Communication skills knowledge, understanding and OSCE performance in medical trainees: a multivariate prospective study using structural equation modelling. 2002

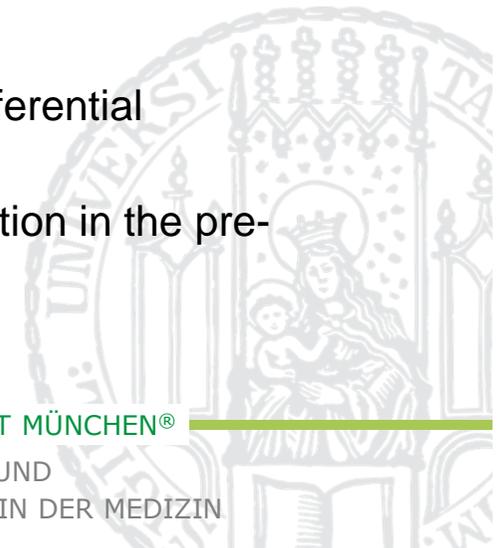
Van Dalen et al. Predicting communication skills with a paper-and-pencil test. 2002

Hulsman et al. Assessment of medical communication skills by computer: assessment method and student experiences. 2004

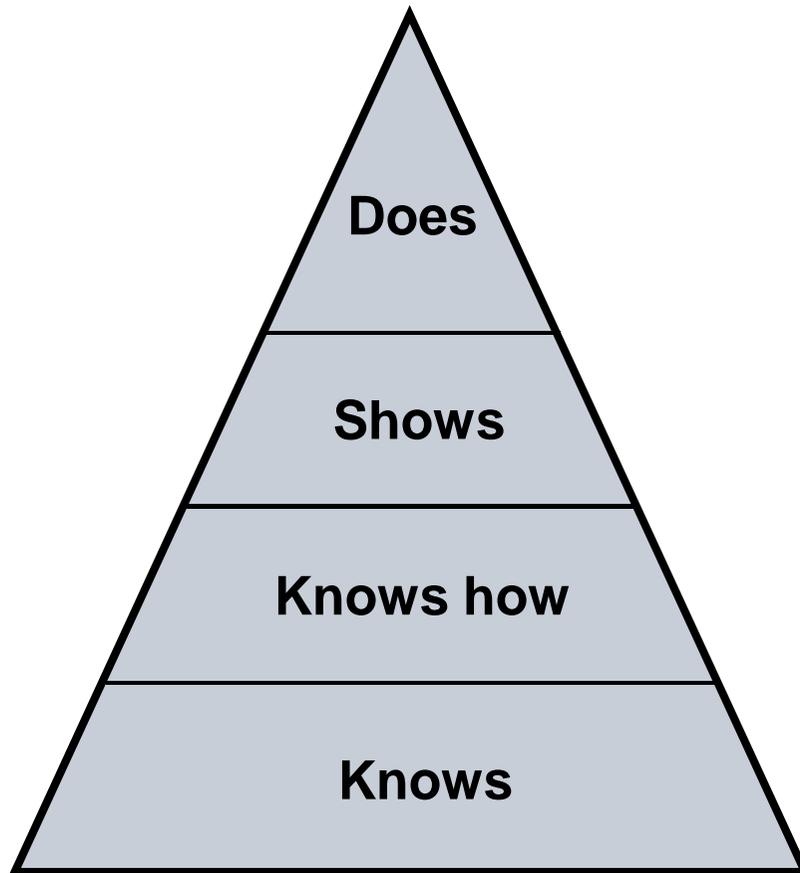
Kim et al. Comparing narrative and multiple choice formats in online communication skills assessment. 2009

Baribeau et al. Using an objective structured video exam to identify differential understanding of communication skills. 2012

Karabilgin et al. Assessing medical student competency in communication in the pre-clinical phase: Objective structured video exam and SP exam. 2012.



Was kann man schriftlich testen?



Stimulus: Kontext-reich (Patienten-Szenarien)
Antwort: offen oder MC; Papier & Bleistift oder
Computerbasiert

Stimulus: Kontext-arm (Fakten)
Antwort: offen oder MC; Papier & Bleistift oder
Computerbasiert

Miller G. The assessment of clinical skills/Competence/Performance 1990

Van der Vleuten: The assessment of professional competence: building blocks for theory development 2010

Wie kann man schriftlich testen?

Kontext-reich

- Key-Feature Ansatz

- Situational judgement Test

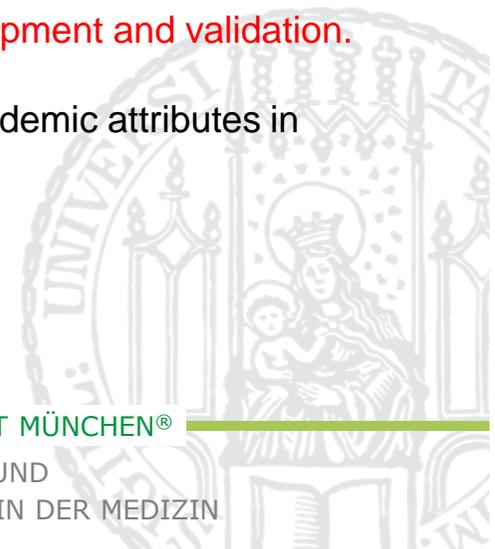
Strahan et al. Predicting Performance on a Situational Judgement Test: The Role of Communication Skills, Listening Skills, and Expertise. 2005

McDaniel et al. Situational Judgment Tests, response instructions, and validity: a meta-analysis. 2007

Schubert et al. A situational judgement test of professional behaviour: development and validation. 2008

Patterson et al. Evaluations of situational judgement tests to assess non-academic attributes in selection. Med Educ 2012

- Script Concordance Test



ProfKom – Professionalisierung von zukünftigen Ärztinnen, Ärzten und Lehrkräften im Bereich der Kommunikationskompetenz

TU München: School of Education

Prof. Dr. Manfred Prenzel, Dr. Johannes Bauer, Dr. Martin Gartmeier, Anne Noll,

Christian-Albrechts-Universität zu Kiel: Dekanat der Medizinischen Fakultät - Arbeitsgruppe Medizindidaktik

Dr. Gudrun Karsten, Grit Möller, Tobias Hoppe-Seyler

Klinikum der Universität München: Lehrstuhl für Didaktik und Ausbildungsforschung in der Medizin

Prof. Dr. Martin Fischer, Dr. Claudia Kiessling

Förderung: BMBF



Zielsetzung

... verschiedene Kombinationen von Komponenten eines Kommunikationstrainings zu professionellen Gesprächsführungskompetenzen im Medizin- und im Lehramtsstudium zu vergleichen.

Dabei stehen die Fragen im Vordergrund,

- a) ob es gelingt, auf Basis von Kommunikationstheorien ein paralleles Kommunikationstraining für beide Studiengänge zu erstellen,
- b) welches Lehr-Lernmodell sich in Bezug auf die Kompetenz professionelle Gesprächssituationen zu analysieren und selbst zu gestalten als besonders günstig erweist.

Beratungsanlass: Partizipative Entscheidungsfindung



Studiendesign: Lehramt und Medizin

geplant: jeweils 4x24 Probanden (Medizin real: 72)

<p>Bedingung 1 5 Stunden</p> <p>Bearbeitung von drei e-learning Modulen:</p> <ul style="list-style-type: none"> • MMG • PEF • Übungsvideo <p>Ggf. zusätzliche Aufgaben, und/oder Videos</p>	<p>Bedingung 2 5 Stunden</p> <p>Kurze theoretische Einführung (MMG, PEF, Feedback)</p> <p>Rollenspiele in 4er Gruppen mit Videofeedback</p>	<p>Bedingung 3 5 Stunden</p> <p>Kombination: 2,5 h e-learning 2,5 h Rollenspiele</p>	<p>Kontrollgruppe</p> <p>Schriftlicher Test 30 Minuten</p> <p>2 Gespräche mit SP je 10 Minuten</p> <p>5 Stunden</p> <p>Kombination: 2,5 Stunden e-learning 2,5 h Rollenspiele</p>
<p>Schriftlicher Test 30 Minuten</p> <p>2 Gespräche mit SP je 10 Minuten</p>	<p>Schriftlicher Test 30 Minuten</p> <p>2 Gespräche mit SP je 10 Minuten</p>	<p>Schriftlicher Test 30 Minuten</p> <p>2 Gespräche mit SP je 10 Minuten</p>	

Unsere Schritte der Testentwicklung

- ↓ Erstellung von Ausbildungszielen
- ↓ Erstellung eines Tests-Blueprints
- ↓ Entwicklung des Medizin-Tests (Papier-&-Bleistift)
- ↓ Validierung des Medizin-Tests mit 7 Ärztinnen & Ärzten
- ↓ Überarbeitung des Medizin-Tests
- ↓ Erstellung eines strukturgleichen Lehramts-Tests (Papier-&-Bleistift)
- ↓ Testlauf mit Lehramtsstudierenden
- ↓ Technische Umsetzung in CASUS
- ↓ Überprüfung mit Studierenden



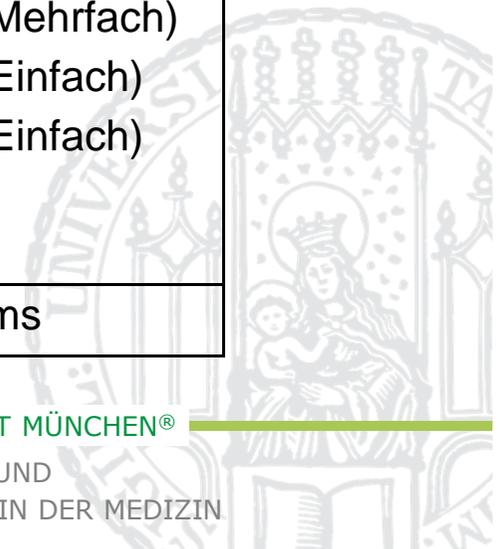
Wie sah unser Test aus?

30 Minuten Testzeit

7 Fragen: Deklaratives Wissen

8 Vignetten mit 1-2 MC-Fragen (15 Fragen): Prozedurales Wissen

	Struktur	Problemlösung	Beziehung
Deklarativ	D3 (Einfach) D4 (Reihenfolge) D5 (Mehrfach)	D2 (Mehrfach) D6 (Mehrfach)	D1 (Mehrfach) D7 (Zuordnung)
Prozedural	1.2 (Einfach) 3.2 (Einfach) 5.1 (Einfach) 6.2 (Einfach) 8.2 (Einfach)	1.1 (Einfach) 2.1 (Einfach) 4.1 (Einfach) 6.1 (Einfach) 7.2 (Einfach) 8.1 (Einfach)	3.1 (Einfach) 4.2 (Mehrfach) 5.2 (Einfach) 7.1 (Einfach)
Anzahl	8 Items	8 Items	6 Items



Ergebnisse Medizin

Item	MW (SD)	Item-Test-K.
1. D1 Wertschätzung	0,76 (0,25)	,482
2. D2 aktive Zuhören	0,88 (0,16)	,407
3. D3 PEF Schritt	0,74 (0,44)	,292
4. D4 PEF Reihenfolge	0,85 (0,17)	,508
5. D5 Überleitungen	0,74 (0,23)	,352
6. D6 Problemdefinition	0,67 (0,14)	,243
7. D7 Umgang mit Gefühlen	0,55 (0,26)	,266
8. Frage 1.1 Patientensicht	0,67 (0,47)	,014
9. Frage 1.2 PEF Schritt	0,89 (0,32)	,091
10. Frage 2.1 klare Info	0,89 (0,32)	,127
11. Frage 3.1 Unterstützung an	0,64 (0,48)	,301
12. Frage 3.2 PEF Schritt	0,82 (0,39)	,115
13. Frage 4.1 Patientensicht	0,69 (0,46)	,168
14. Frage 4.2 Aktivierung	0,61 (0,15)	,063
15. Frage 5.1 zusammenfassen	0,32 (0,47)	,072
16. Frage 5.2 einbeziehen	0,96 (0,20)	,063
17. Frage 6.1 PEF Schritt	0,51 (0,50)	,200
18. Frage 6.2 zusammenfassen	0,82 (0,39)	,296
19. Frage 7.1 Emotion ansprechen	0,94 (0,23)	-,058
20. Frage 7.2 Gesprächsende	0,49 (0,50)	,245
21. Frage 8.1 Patientensicht	0,68 (0,47)	,386
22. Frage 8.2 Gesprächsende	0,40 (0,49)	,123

72 Studierende

Pro Frage 1 Punkt = 22 Punkte max.

MW = 15,5 (70%) (SD 2,59)

Gesamttest: $\alpha = .590$

Subskalen

7 D Fragen: $\alpha = .678$

15 KF Fragen: $\alpha = .410$

Subskalen

Struktur $\alpha = .376$

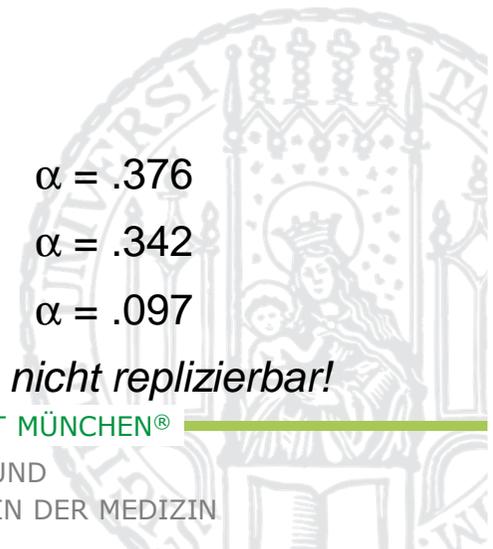
Problem $\alpha = .342$

Beziehung $\alpha = .097$

Faktorenanalytisch nicht replizierbar!

KLINIKUM DER UNIVERSITÄT MÜNCHEN®

LEHRSTUHL FÜR DIDAKTIK UND
AUSBILDUNGSFORSCHUNG IN DER MEDIZIN



Ergebnisse Lehramt

Item	MW (SD)	Item-Test-K.
1. D1 Wertschätzung	0,63 (0,34)	,256
2. D2 aktive Zuhören	0,75 (0,26)	,238
3. D3 PEF Schritt	0,74 (0,44)	,173
4. D4 PEF Reihenfolge	0,74 (0,20)	,253
5. D5 Überleitungen	0,73 (0,20)	,269
6. D6 Problemdefinition	0,74 (0,19)	,135
7. D7 Umgang mit Gefühlen	0,58 (0,23)	-,042
8. Frage 1.1 Patientensicht	0,65 (0,48)	,025
9. Frage 1.2 PEF Schritt	0,84 (0,37)	,295
10. Frage 2.1 klare Info	0,80 (0,40)	,101
11. Frage 3.1 Unterstützung an	0,60 (0,49)	,072
12. Frage 3.2 PEF Schritt	0,54 (0,50)	,098
13. Frage 4.1 Patientensicht	0,78 (0,42)	,211
14. Frage 4.2 Aktivierung	0,80 (0,22)	,121
15. Frage 5.1 zusammenfassen	0,06 (0,24)	,010
16. Frage 5.2 einbeziehen	0,18 (0,38)	,173
17. Frage 6.1 PEF Schritt	0,45 (0,50)	-,088
18. Frage 6.2 zusammenfassen	0,14 (0,34)	,041
19. Frage 7.1 Emotion ansprechen	0,49 (0,50)	,170
20. Frage 7.2 Gesprächsende	0,29 (0,46)	-,027
21. Frage 8.1 Patientensicht	0,83 (0,37)	,129
22. Frage 8.2 Gesprächsende	0,43 (0,50)	-,087

96 Studierende

Pro Frage 1 Punkt = 22 Punkte max.

MW = 12,8 (58%) SD 2,23

Gesamttest: $\alpha = .368$

Subskalen

7 D Fragen: $\alpha = .625$

15 KF Fragen: $\alpha = .269$



Was nun?

- ↓ Kritische Items (Fragen & Antworten) identifiziert
- ↓ Konstruktion der Vignette, Frage und Antworten diskutiert

Lehramt

- ↓ 1 Item komplett eliminiert da unklare Formulierung der richtigen Antwort (6.2.)
- ↓ 4 x unklare Disktraktoren als richtig gewertet (D7, 6.1, 7.2, 8.2)

Medizin

- ↓ 1 x unklare Disktraktoren als richtig gewertet (D7)



Bessere interne Konsistenz?

Lehramt

- α (ges.) = ,447
2 Items weiterhin negativ korreliert (6.1, 8.2)
zusätzlich ein weiteres Item negativ korreliert (1.1)
- Nach Elimination der 3 Items: $\alpha = ,566$
- 64/33 Fragen wären notwendig für ein $\alpha = ,700$ (~ 90-45' Testzeit)*

Medizin

- α (ges.) = .583
3 Items: Item-Test-Korrelation < ,100 bzw. negativ
- Nach Elimination der 3 Items: $\alpha = ,625$
- 37/27 Fragen wären notwendig für ein $\alpha = ,700$ (~ 50-40' Testzeit)*

**Spearman-Brown-Formel*



Diskussion & Schlussfolgerungen

- ! Testzeit zu kurz
- ! Validierung mit Ärzten hat Test wahrscheinlich verbessert
- ! Deklarative Fragen führen zu höherem α ; aber Ärzte z.T. nicht begeistert davon („irrelevant“)
- ! Antwortalternativen der SJT Fragen schwierig zu konstruieren, aber realitätsnäher



Diskussion & Schlussfolgerungen

- ? Antwortformate: Rating/Ranking statt Einfachauswahl?
- ? Gewichtetes Scoring angemessener als 0:1?
- ? Sensitiv genug um Unterschiede zwischen Bedingungen zu messen?
- ? Bessere Ergebnisse mit Videos statt Textbasierten Vignetten?
- ? Abgrenzung Key Feature vs SJT

