



Universitätsklinikum  
Hamburg-Eppendorf

AG Auswahlverfahren

# Wie am besten? - Situational Judgment Test: eine Methode, drei Herangehensweisen

Anja Bath, Johanna Hissbach, Mirjana Knorr,  
und Wolfgang Hampe

Institut für Biochemie und Molekulare Zellbiologie  
Zentrum für Experimentelle Medizin

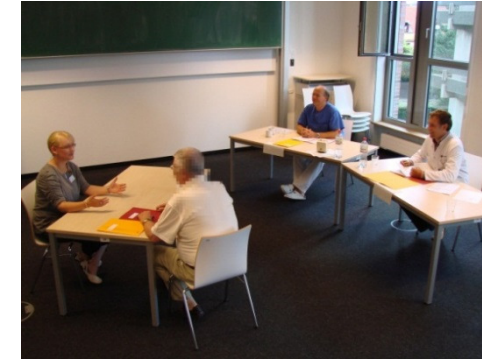
25.09.-27.09.2014  
GMA - Jahrestagung in Hamburg 2014



etwa 1200 Bewerber  
(plus Notengleiche)

**HAM-Nat**

Zulassung für  
Rangplätze 1-115



Rangplätze 116-315

**HAM-Int (MMI)**

**Situational Judgment Test**

2011 papierbasiert/single choice

2012 videobasiert/freie  
Antworten

2013 videobasiert/Fragebogen  
mit Likert-Skala



- Darstellung einer berufsbezogenen Situation mit mehreren Handlungsalternativen
- Situations- und Antwortdarstellung:
  - Schriftlich vs. video-/multimediabasiert
- Frageformat:
  - Unterscheidung in „Should do“ und „Would do“-Ansatz
  - Auswahl:
    - Beste Alternative und/oder schlechteste Alternative
    - Beste und zweitbeste Alternative
    - Ratingskala: Einschätzung jeder Alternative



## Papierbasierterer SJT mit Auswahl der besten

**Antwort** (N=194, Kooperation mit der Medizinischen Universität Graz)

- 12 Situationen mit 5 möglichen Antworten
- Bewertung: 1 Punkt/richtige Antwort (Expertenrating)

Pro: Durchführbarkeit, objektives und einfaches Bewertungssystem

Contra:

- Schwierigkeiten:  $p = .82-.99$
- Trennschärfe:  $r = .03-.34$
- Reliabilität: Cronbach's  $\alpha = .44$

Fazit: Antwortformat muss mehr Varianz erzeugen



## **videobasierter SJT mit freien Antworten**

(N=192, Kooperation mit Maastricht Universität)

- 4 Situationen mit 2-3 Fragen pro Situation
- Ethische Entscheidungen vs. Verhaltensbeobachtung
- Bewertung: a priori definierter Antwortschlüssel
  - Ethische Bewertung: Gesamtqualität der Argumentation einschätzen  
(0-3 Punkte)
  - Verhaltensbeurteilung: genannte Verhaltensweisen abzählen  
(0-9 Punkte)
- 2 unabhängige trainierte Rater



Pro: Korrelation mit MMI Ergebnis:  $r = .26^{**}$

Contra:

- Inter-Rater-Übereinstimmung auf Antwortlevel:
  - $ICCs_{(ethisch)} = .28-.64$
  - $ICCs_{(Verhalten)} = .47-.87$
- sehr zeitaufwendig und teuer (Videoproduktion, Rater-training, Auswertung)

Fazit:

- videobasierte Darstellung wird als ansprechender empfunden
- ein standardisierteres und weniger zeitaufwendiges Bewertungssystem ist nötig

## Videobasierter SJT mit einem Fragebogen (N=198)

- eine interpersonelle Konfliktsituation
- Fragebogen orientiert sich am TACT-Konzept (Ajzen & Fishbein, 1980)
  - 4 **t**argets,
  - 6 – 8 **a**ctions per target
  - **c**ontext und **t**ime  
wurden konstant gehalten
- Bewertungssystem:  
(Mittelwert<sub>Expert</sub> – Antwort<sub>Bewerber</sub>)





### Pro:

- Durchführbarkeit: Vorbereitung, Ausführung, Auswertung
- Deskriptive Statistik Schwierigkeiten:  $p = .0-.43$   
Streuung:  $SD = .47-1.24$

### Contra:

- Reliabilität: Anteil erklärter Varianz:

Ziele als "retest" <sub>(4)</sub> :	0%
Handlungen als "retest" <sub>(28)</sub> :	0%
Experten als "retest" <sub>(38)</sub> :	42%
- Validität:  
Korrelation mit MMI Ergebnis:  $r = .13$





Fazit:

Fragebogen mit standardisierter Bewertung

aber

Vollständige Matrix: jedes Ziel mit jeder Handlung  
kombinieren

Offene Fragen:

- Wie viele Situationen sind nötig?
- Wie müssen die Situationen konzipiert sein?
- Sind Expertenratings ein guter Vergleichsmaßstab?
- Welches Maß ist gut geeignet, um die Abweichungen zwischen Experten und Bewerbern zu messen?