



UniversitätsKlinikum Heidelberg

OSCE

Objective Structured Clinical Examination

**18th Graz Conference on Medical Education 2014
Salzburg, Austria, 03.04.2014**

**OA Priv.-Doz. Dr. med. Christoph Nikendei, MME
University of Heidelberg
Department for General Internal Medicine and Psychosomatics**

Agenda

OSCE

OSCE

Agenda

OSCE

OSCE

Agenda

Assessment of clinical competence – why and how?

OSCE

OSCE

Agenda

Assessment of clinical competence – why and how?

OSCE – history

Agenda

Assessment of clinical competence – why and how?

OSCE – history

OSCE – what is the idea behind it?

Agenda

Assessment of clinical competence – why and how?

OSCE – history

OSCE – what is the idea behind it?

OSCE – why is it objective?

Agenda

Assessment of clinical competence – why and how?

OSCE – history

OSCE – what is the idea behind it?

OSCE – why is it objective?

OSCE – design decisions before getting started

Agenda

Assessment of clinical competence – why and how?

OSCE – history

OSCE – what is the idea behind it?

OSCE – why is it objective?

OSCE – design decisions before getting started

Pre-OSCE quality assurance

Agenda

Assessment of clinical competence – why and how?

OSCE – history

OSCE – what is the idea behind it?

OSCE – why is it objective?

OSCE – design decisions before getting started

Pre-OSCE quality assurance

Post-OSCE quality assurance

Agenda

Assessment of clinical competence – why and how?

OSCE – history

OSCE – what is the idea behind it?

OSCE – why is it objective?

OSCE – design decisions before getting started

Pre-OSCE quality assurance

Post-OSCE quality assurance

OSCE – aspects of reliability

Agenda

Assessment of clinical competence – why and how?

OSCE – history

OSCE – what is the idea behind it?

OSCE – why is it objective?

OSCE – design decisions before getting started

Pre-OSCE quality assurance

Post-OSCE quality assurance

OSCE – aspects of reliability

OSCE – aspects of validity

Agenda

Assessment of clinical competence – why and how?

OSCE – history

OSCE – what is the idea behind it?

OSCE – why is it objective?

OSCE – design decisions before getting started

Pre-OSCE quality assurance

Post-OSCE quality assurance

OSCE – aspects of reliability

OSCE – aspects of validity

OSCE – a tool for medical education research

Agenda

Assessment of clinical competence – why and how?

OSCE – history

OSCE – what is the idea behind it?

OSCE – why is it objective?

OSCE – design decisions before getting started

Pre-OSCE quality assurance

Post-OSCE quality assurance

OSCE – aspects of reliability

OSCE – aspects of validity

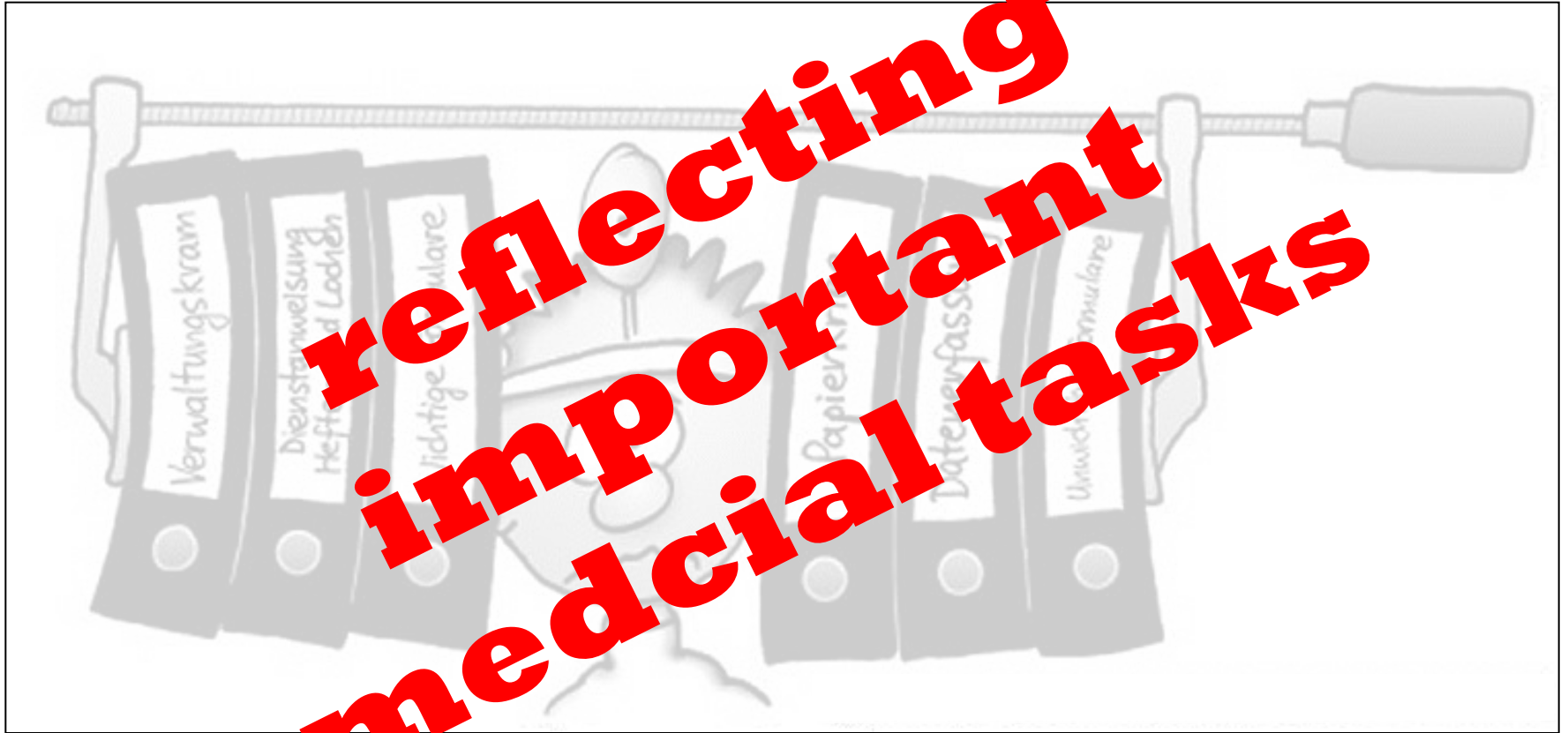
OSCE – a tool for medical education research

OSCE – future prospects and challenges

Assessment of clinical competence



Assessment of clinical competence



Assessment of clinical competence



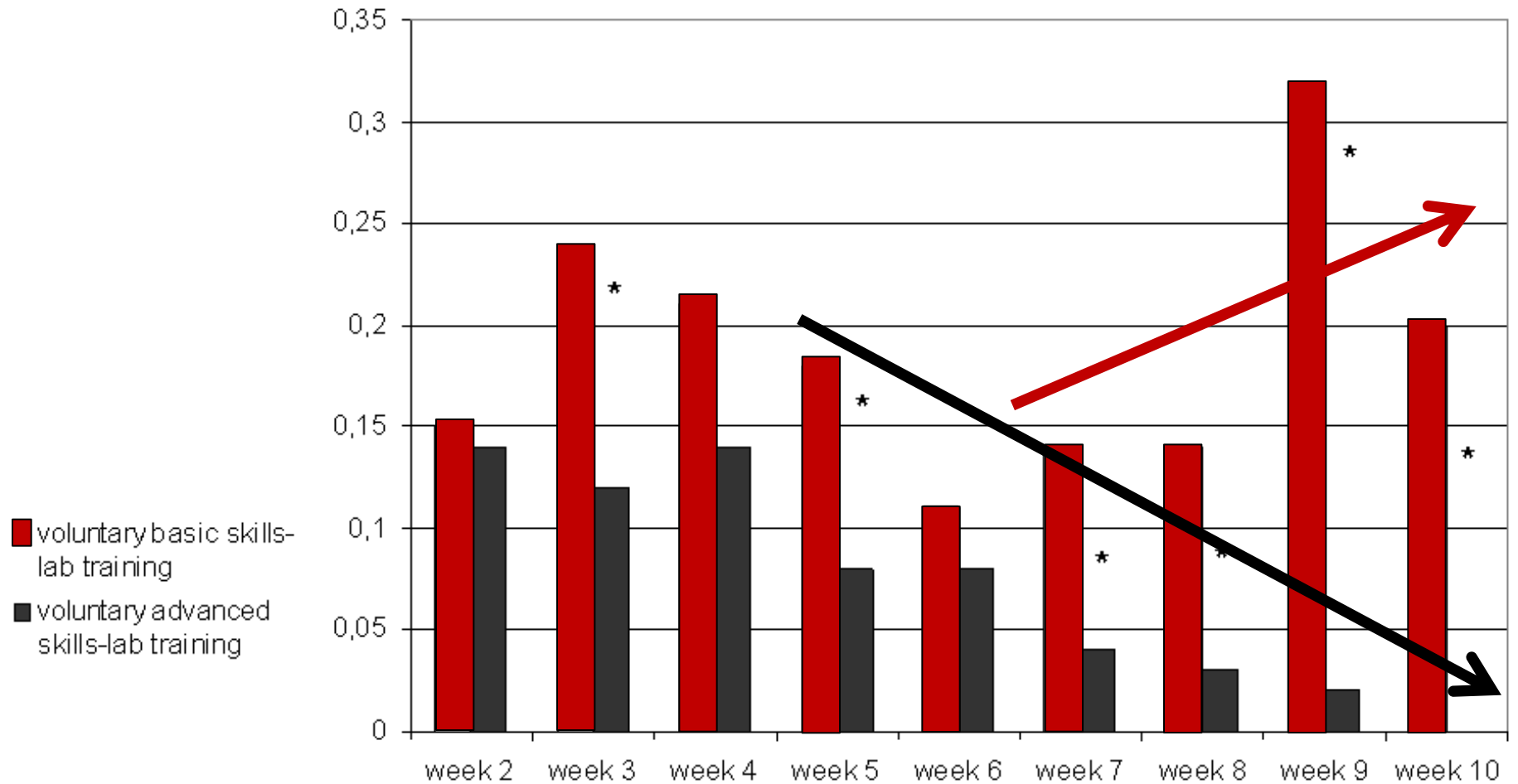
Assessment of clinical competence



Assessment of clinical competence



Assessment of clinical competence



OSCE - history



OSCE - history



Medical Education

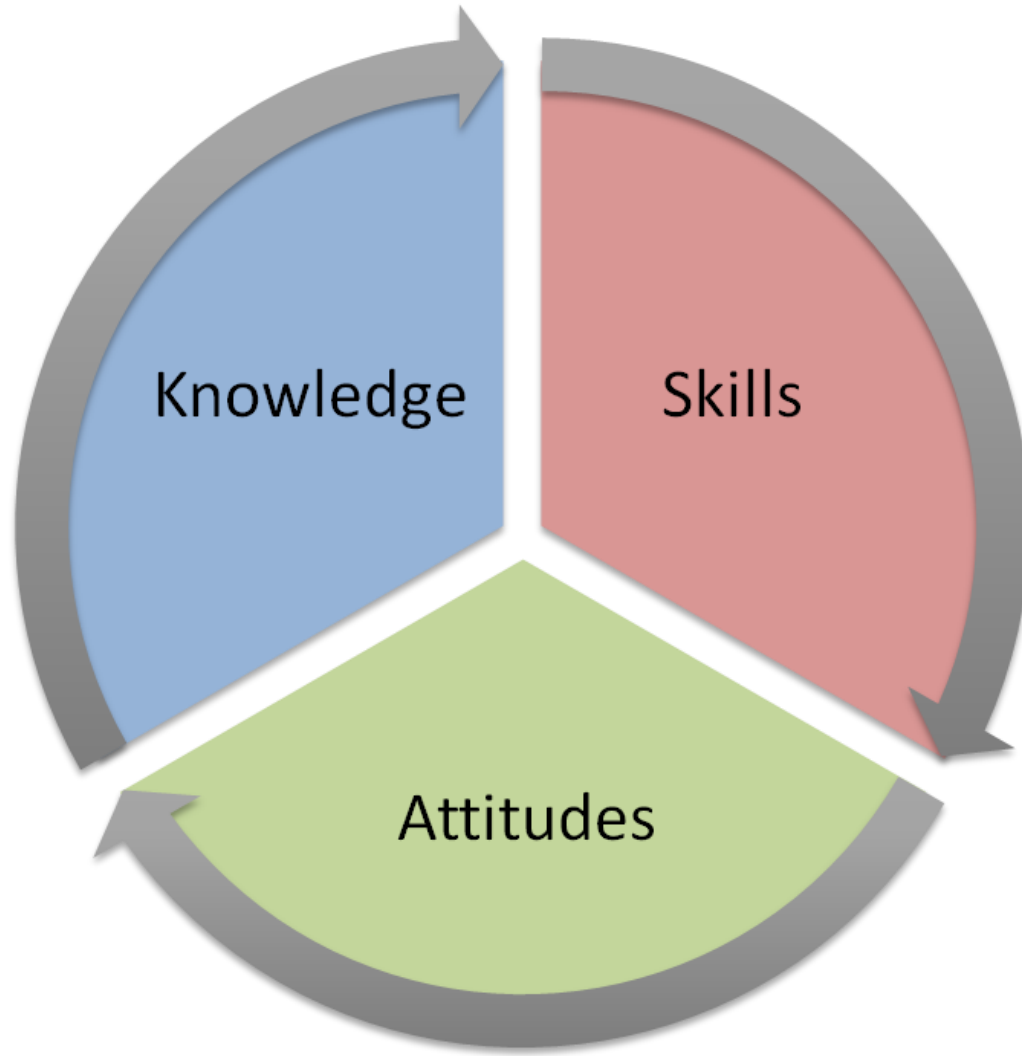
Assessment of Clinical Competence using Objective Structured Examination

R. McG. HARDEN, MARY STEVENSON, W. WILSON DOWNIE, G. M. WILSON

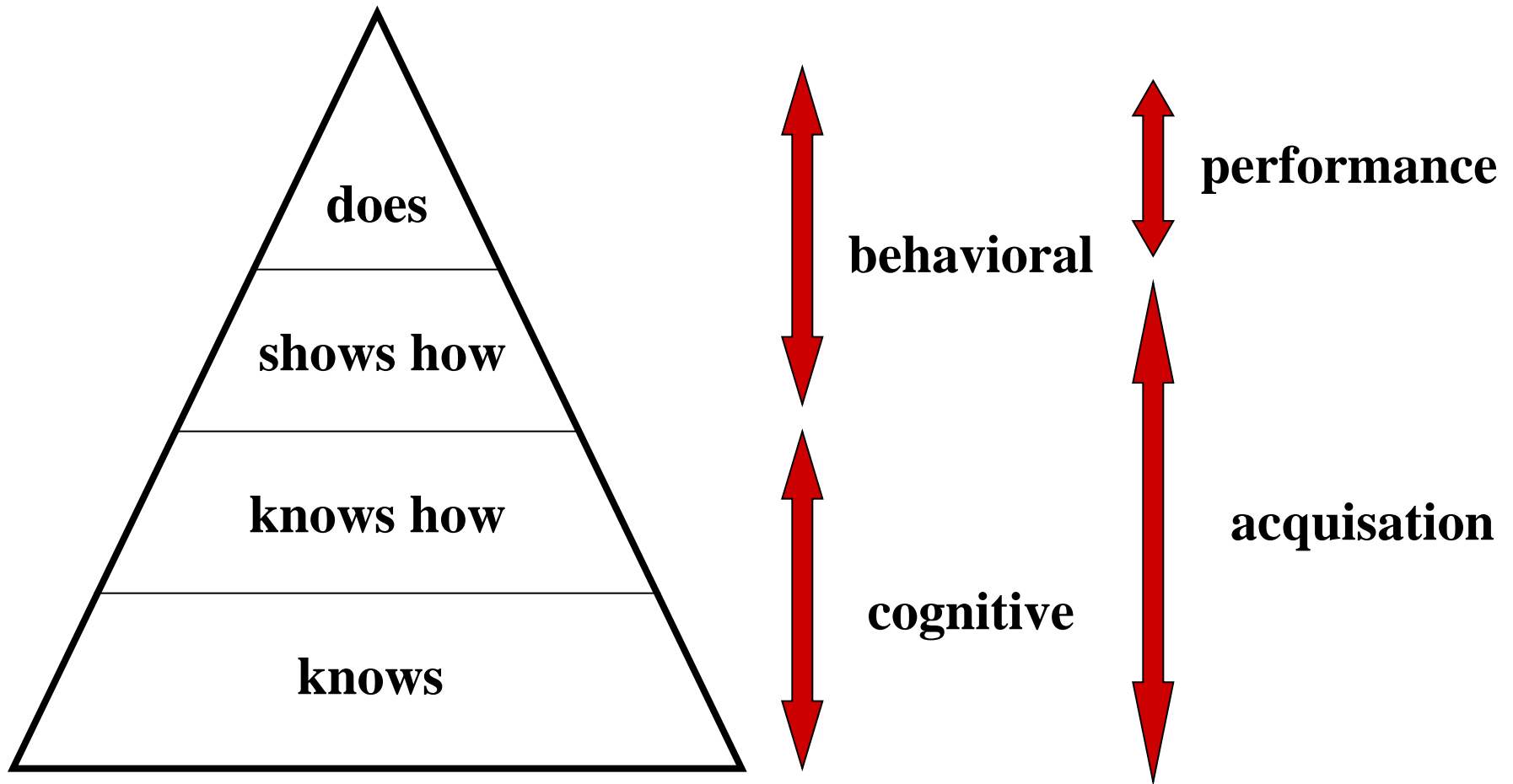
OSCE – what is the idea behind it?



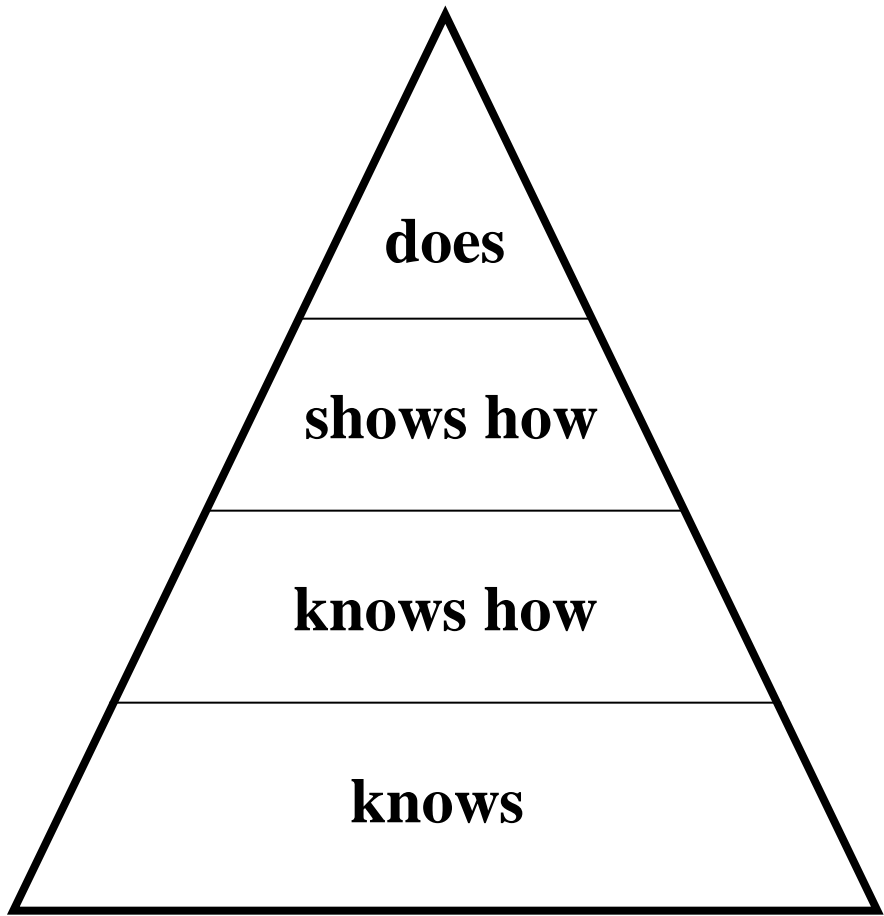
OSCE – what is the idea behind it?



OSCE – what is the idea behind it?



OSCE – what is the idea behind it?



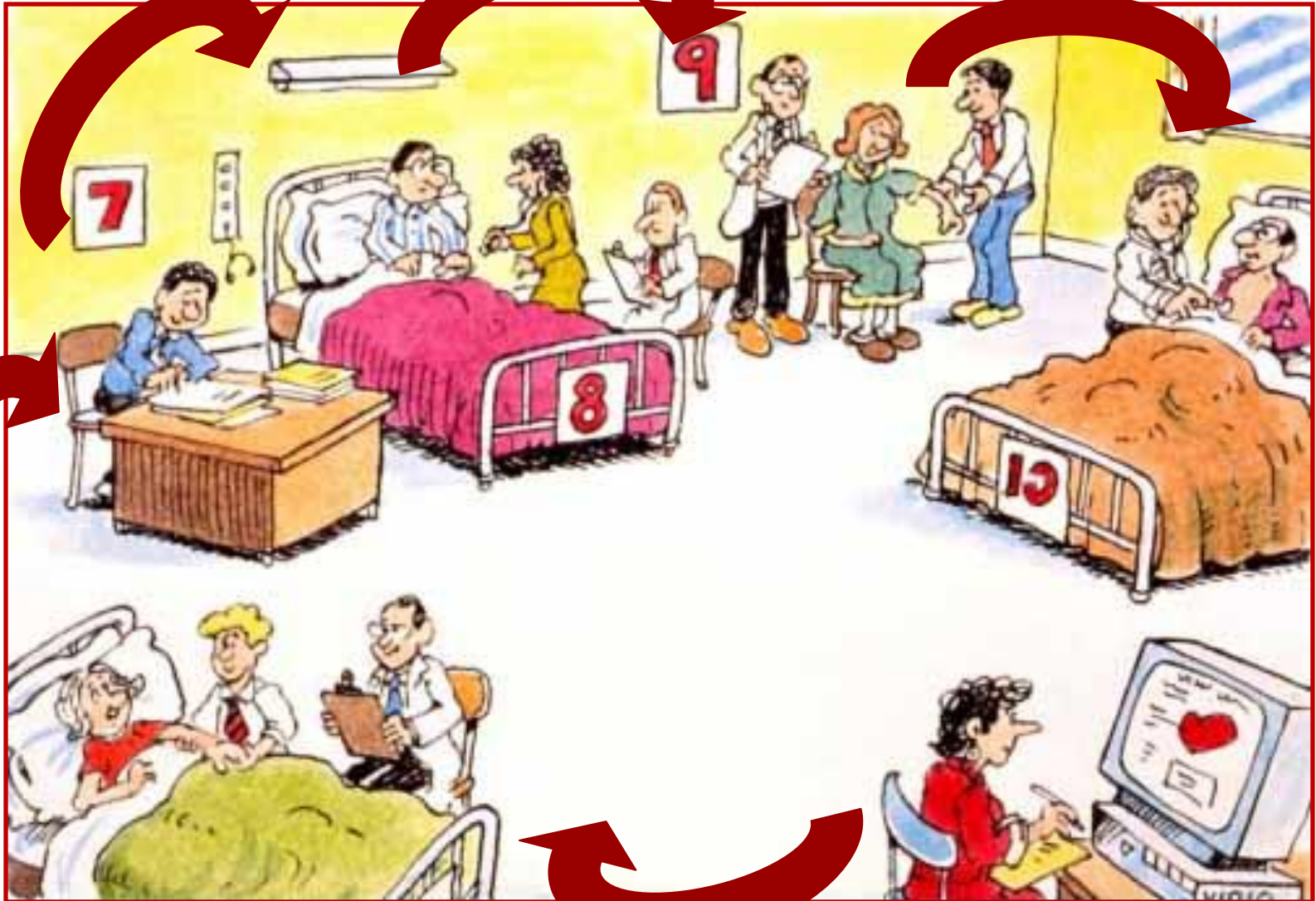
**performance
assessment in vivo**

**performance
assessment in vitro**

**clinical context
based tests**

**factual
tests**

OSCE – what is the idea behind it?



OSCE – why is it objective?

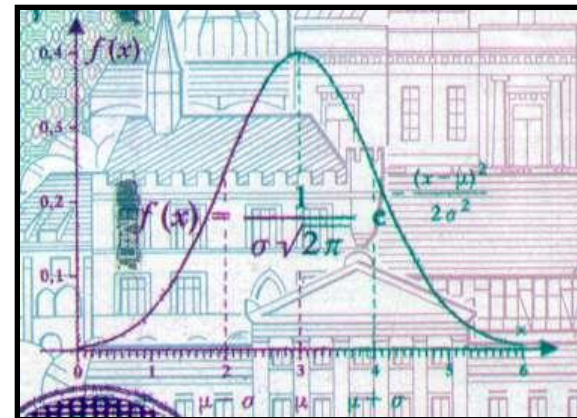
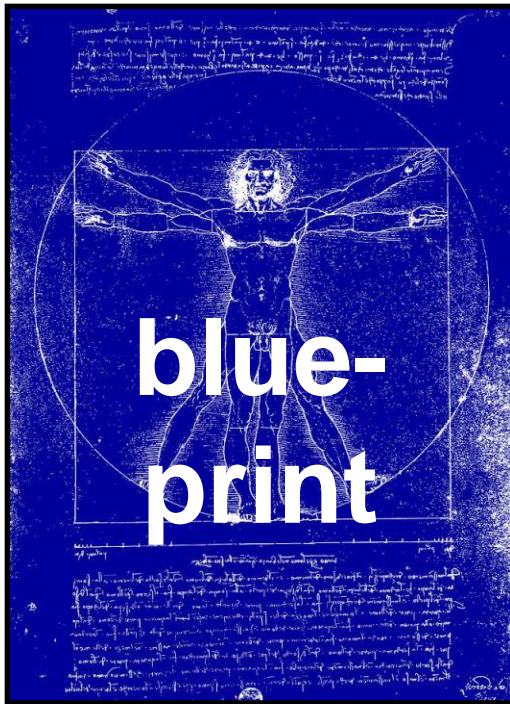
OSCE

OSCE

OSCE – why is it objective?



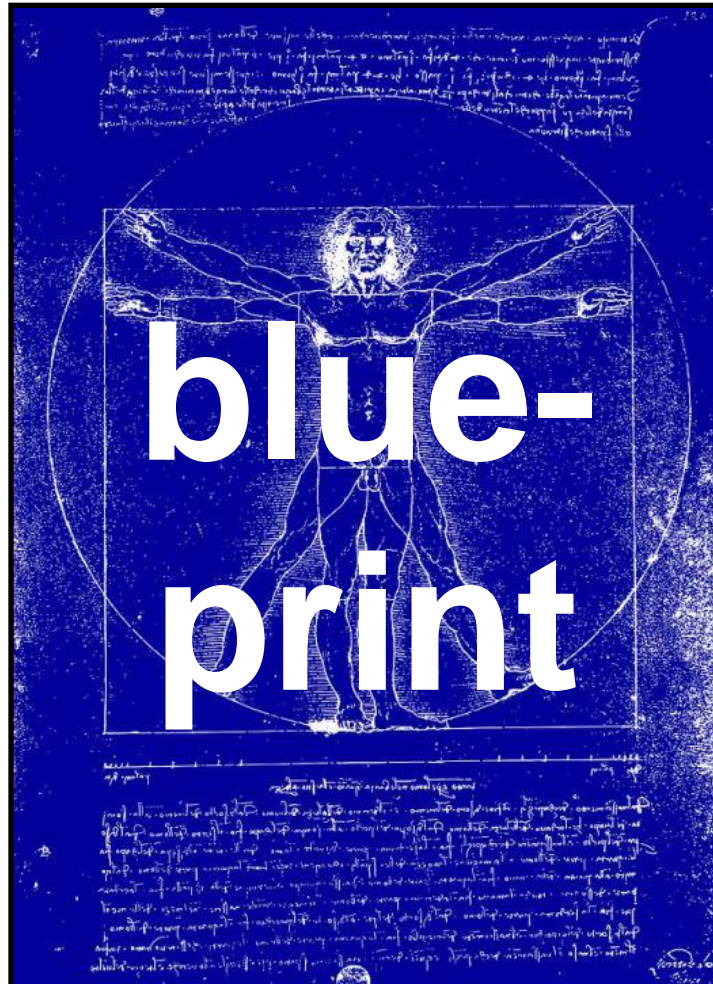
OSCE – why is it objective?



OSCE – design decisions



OSCE – design decisions blueprint

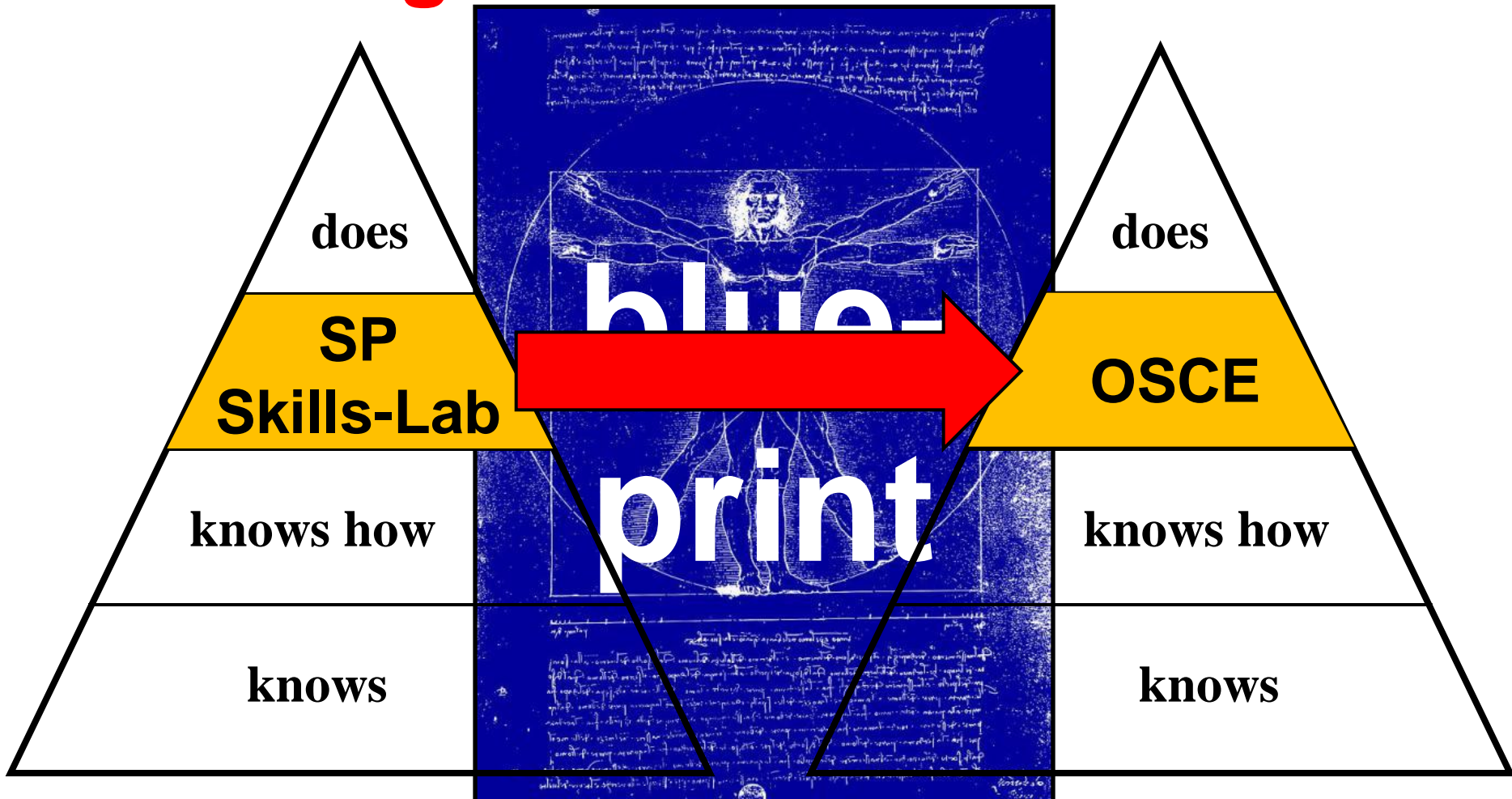


OSCE – design decisions

Teaching

blueprint

Assessment



Biggs, SRHE (1999)

Miller, Acad Med (1990)

Wass & Jones, Lancet (2001)

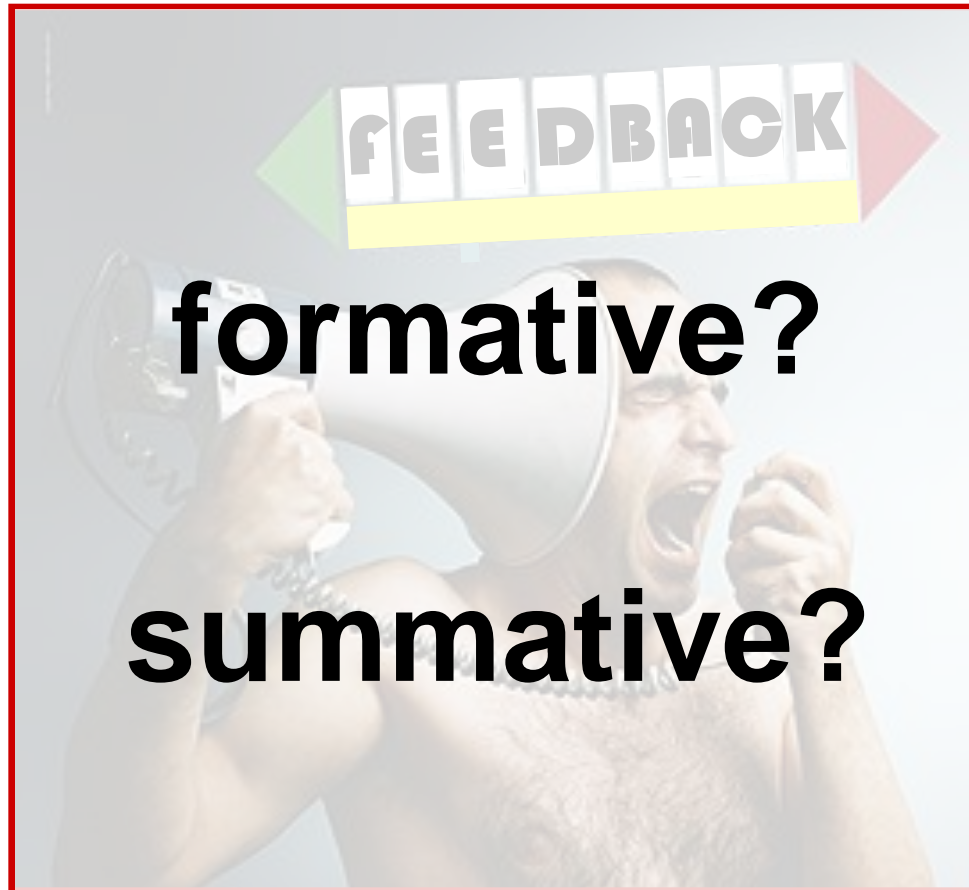
OSCE – design decisions

formative / summative



OSCE – design decisions

formative / summative



OSCE – design decisions

compensatory / non-compensatory

knock-out stations?



compensatory?

Newble, Med Educ (2004)
Muijtens & von der Vleuten, Eur J Dent Educ (2009)
Chesser & Heyes, Med Educ (2004)
Friedman Ben-David, AMEE (2000)

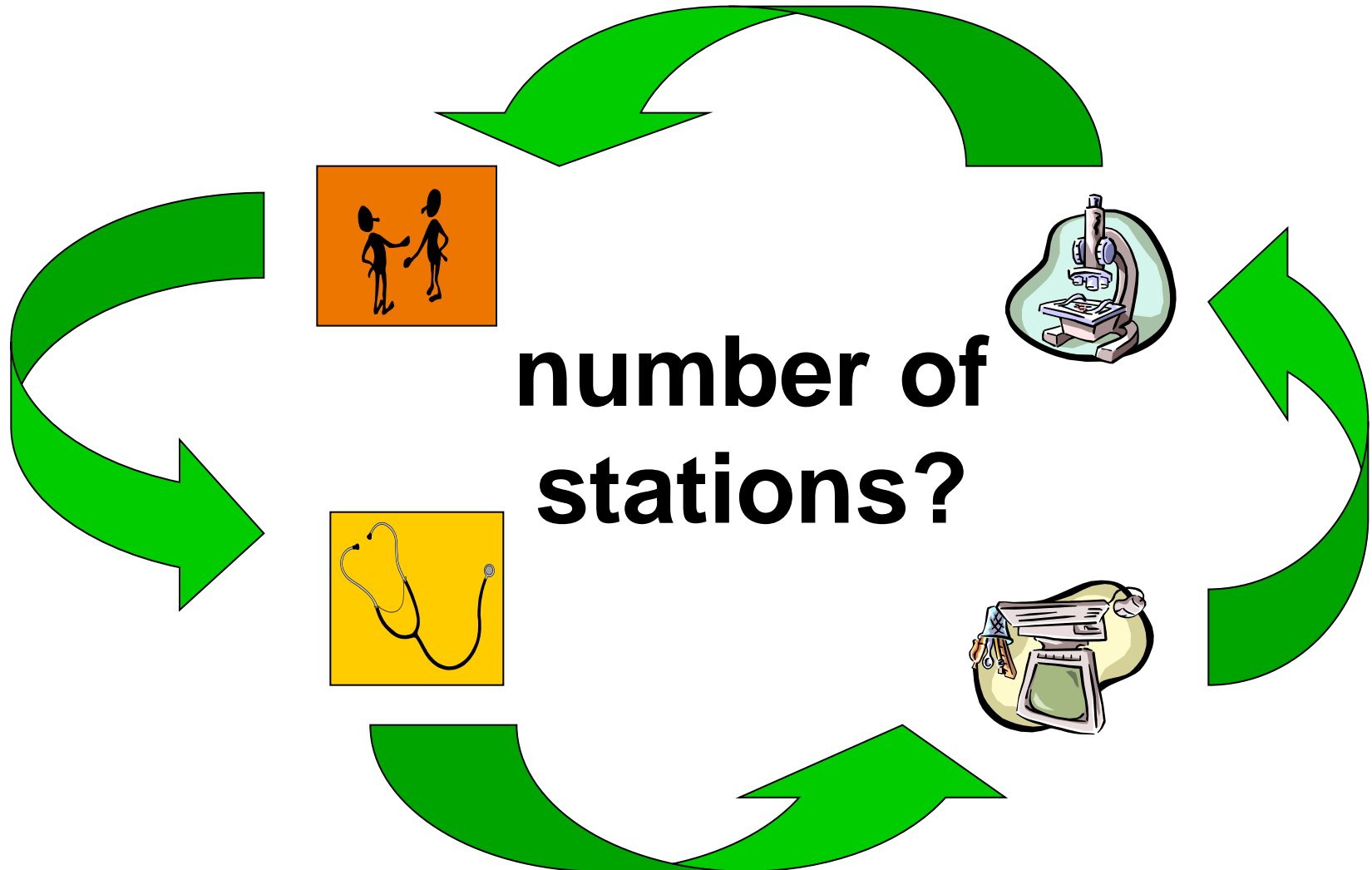
OSCE – design decisions

feedback at single stations



OSCE – design decisions

number of stations



OSCE – design decisions

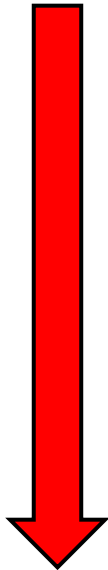
number of stations

- **University of Dundee, Scotland:**
25 – 35 stations à 4½ minutes (Davis, 2003)
 - **Medical Council, Canada:**
20 stationens à 10 minutes (Reznick et al., 1996)
 - **Havard Medical School, USA:**
16 stationens à 9 minutes (Hamann et al., 2002)
 - **ACGME & ABMS, USA:**
14 – 18 stationens à 10 – 15 minutes (Swanson et al., 2002)
- 

OSCE – design decisions

number of stations

Test Length Hours	One examiner per station	number of stations
1	0.34	4
2	0.51	8
4	0.67	12
6	0.76	16
8	0.81	20



OSCE – design decisions

checklists / global rating scales

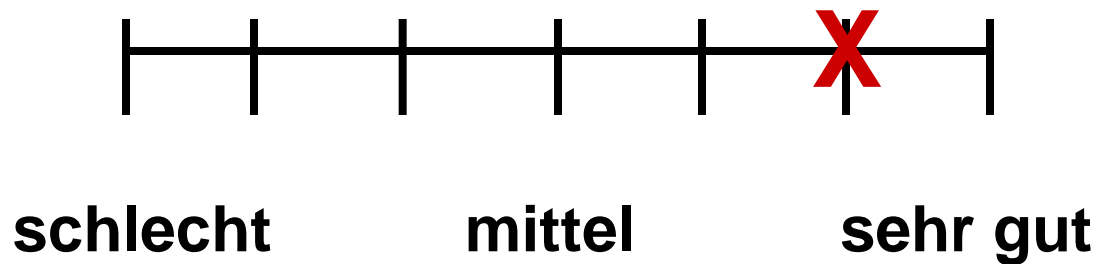
Checkliste

- ✓
- ✓
- ✓
- ✓
- ✓
- ✓
- ✓
- -

→ 6 Punkte

Globale Bewertung

→ 6 Punkte



OSCE – design decisions

checklists / global rating scales

checklist vs. global rating scale

OSCE – design decisions

checklists / global rating scales

checklist vs. global rating scale

- higher interrater-reliability for checklists when assessing procedural technical skills (0.9 vs. 0.7)

Lund & Nikendei, Plos One (2012)

OSCE – design decisions

checklists / global rating scales

checklist vs. global rating scale

- higher interrater-reliability for checklists when assessing procedural technical skills (0.9 vs. 0.7)
- global rating scale better captures interactive competencies

Lund & Nikendei, Plos One (2012)

Van der Vleuten & de Graaff, Med Educ (1991)
Noraman & de Graaff, Med Educ (1991)

OSCE – design decisions

checklists / global rating scales

checklist vs. global rating scale

- higher interrater-reliability for checklists when assessing procedural technical skills (0.9 vs. 0.7)
Lund & Nikendei, Plos One (2012)
- global rating scale better captures interactive competencies
Van der Vleuten & de Graaff, Med Educ (1991)
Noraman & de Graaff, Med Educ (1991)
- combination of both formats is suggested
Newble, Med Educ (2004)

OSCE – design decisions

checklists / global rating scales

checklist vs. global rating scale

- higher interrater-reliability for checklists when assessing procedural technical skills (0.9 vs. 0.7)
Lund & Nikendei, Plos One (2012)
- global rating scale better captures interactive competencies
Van der Vleuten & de Graaff, Med Educ (1991)
Noraman & de Graaff, Med Educ (1991)
- combination of both formats is suggested
Newble, Med Educ (2004)
- high number of items decreases reliability
Schmidt & Boshuizen, Acad Med (1998)
Wilkinson & Egan, Acad Med (2003)

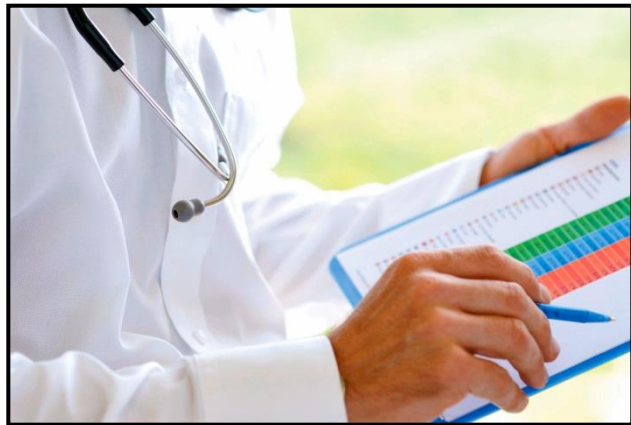
OSCE – design decisions

raters



OSCE – design decisions

raters

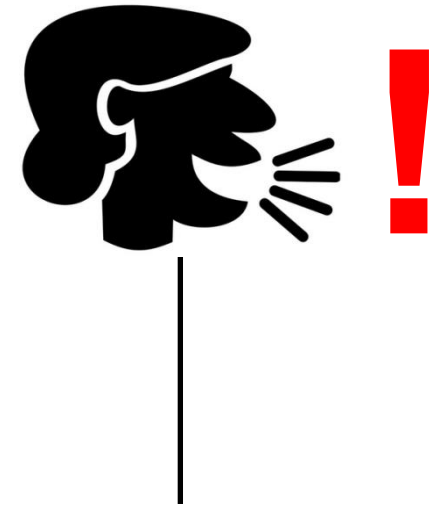


— **0.78; <.001**
|



OSCE – design decisions

raters



OSCE – design decisions

stations and raters

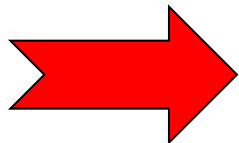
Test Length Hours	One examiner per station		Two examiners per station
1	0.34	↔	0.33
2	0.51	↔	0.50
4	0.67	↔	0.67
6	0.76	↔	0.75
8	0.81	↔	0.80

OSCE – design decisions

stations and raters

Review including 39 studies

- alpha across stations 0.66
- alpha within stations 0.78
- higher number of stations
higher number of raters



higher reliability

OSCE – design decisions

costs

n = 145

4 stations

4 minutes per station

personnel costs

per student: 86 €

Pre-OSCE quality assurance



Pre-OSCE quality assurance review of checklists

Name der OSCE-Station:	Autor:
	Reviewer:

Inhaltliche Kriterien:			
Schwierigkeitsgrad der OSCE-Station insgesamt	<input type="checkbox"/> leicht	<input type="checkbox"/> mittel	<input type="checkbox"/> schwer
Schwierigkeitsgrad der Unterfragen	<input type="checkbox"/> leicht	<input type="checkbox"/> mittel	<input type="checkbox"/> schwer
Fachliche Relevanz des Themas für die Zielgruppe	<input type="checkbox"/> vorhanden	<input type="checkbox"/> bedingt vorhanden	<input type="checkbox"/> nicht vorhanden
Anwendungsbezug	<input type="checkbox"/> hoch	<input type="checkbox"/> mittel	<input type="checkbox"/> gering
Klinische Fallvignette vorhanden	<input type="checkbox"/> ja	<input type="checkbox"/> nein	
An dieser Station werden folgende Teile geprüft:	Bitte Prozent-Angaben:		
Kommunikative Fähigkeiten:	%		
Praktische Fähigkeiten:	%		
Entscheidungs-Wissen:	%		
Fakten-Wissen:	%		
Wieviel Prozent der Aufgaben könnten auch schriftlich geprüft werden?	%		

Pre-OSCE quality assurance review of checklists

Name der OSCE-Station:	Autor:
	Reviewer:

Inhaltliche Kriterien:			
Schwierigkeitsgrad der OSCE-Station insgesamt	<input type="checkbox"/> leicht	<input type="checkbox"/> mittel	<input type="checkbox"/> schwer
Schwierigkeitsgrad der Unterfragen	<input type="checkbox"/> leicht	<input type="checkbox"/> mittel	<input type="checkbox"/> schwer
Fachliche Relevanz des Themas für die Zielgruppe	<input type="checkbox"/> vorhanden	<input type="checkbox"/> bedingt vorhanden	<input type="checkbox"/> nicht vorhanden
Anwendungsbezug	<input type="checkbox"/> hoch	<input type="checkbox"/> mittel	<input type="checkbox"/> gering
Klinische Fallvignette vorhanden	<input type="checkbox"/> ja	<input type="checkbox"/> nein	
An dieser Station werden folgende Teile geprüft:	Bitte Prozent-Angaben:		
Kommunikative Fähigkeiten:	%		
Praktische Fähigkeiten:	%		
Entscheidungs-Wissen:	%		
Fakten-Wissen:	%		
Wieviel Prozent der Aufgaben könnten auch schriftlich geprüft werden?	%		

Pre-OSCE quality assurance review of checklists

Name der OSCE-Station:	Autor:
	Reviewer:

Inhaltliche Kriterien:	
Schwierigkeitsgrad der OSCE-Station insgesamt	<input type="checkbox"/> leicht <input type="checkbox"/> mittel <input type="checkbox"/> schwer
Schwierigkeitsgrad der Unterfragen	<input type="checkbox"/> leicht <input type="checkbox"/> mittel <input type="checkbox"/> schwer
Fachliche Relevanz des Themas für die Zielgruppe	<input type="checkbox"/> vorhanden <input type="checkbox"/> bedingt vorhanden <input type="checkbox"/> nicht vorhanden
Anwendungsbezug	<input type="checkbox"/> hoch <input type="checkbox"/> mittel <input type="checkbox"/> gering
Klinische Fallvignette vorhanden	<input type="checkbox"/> ja <input type="checkbox"/> nein
An dieser Station werden folgende Teile geprüft:	Bitte Prozent-Angaben:
Kommunikative Fähigkeiten:	%
Praktische Fähigkeiten:	%
Entscheidungs-Wissen:	%
Fakten-Wissen:	%
Wieviel Prozent der Aufgaben könnten auch schriftlich geprüft werden?	%

Pre-OSCE quality assurance review of checklists

Name der OSCE-Station:	Autor:
	Reviewer:

Inhaltliche Kriterien:			
Schwierigkeitsgrad der OSCE-Station insgesamt	<input type="checkbox"/> leicht	<input type="checkbox"/> mittel	<input type="checkbox"/> schwer
Schwierigkeitsgrad der Unterfragen	<input type="checkbox"/> leicht	<input type="checkbox"/> mittel	<input type="checkbox"/> schwer
Fachliche Relevanz des Themas für die Zielgruppe	<input type="checkbox"/> vorhanden	<input type="checkbox"/> bedingt vorhanden	<input type="checkbox"/> nicht vorhanden
Anwendungsbezug	<input type="checkbox"/> hoch	<input type="checkbox"/> mittel	<input type="checkbox"/> gering
Klinische Fallvignette vorhanden	<input type="checkbox"/> ja	<input type="checkbox"/> nein	
An dieser Station werden folgende Teile geprüft:	Bitte Prozent-Angaben:		
Kommunikative Fähigkeiten:	%		
Praktische Fähigkeiten:	%		
Entscheidungs-Wissen:	%		
Fakten-Wissen:	%		
Wieviel Prozent der Aufgaben könnten auch schriftlich geprüft werden?	%		

Pre-OSCE quality assurance

review of checklists

Formale Kriterien:			
Eindeutigkeit der Aufgabenstellung für den Prüfling	<input type="checkbox"/> eindeutig	<input type="checkbox"/> verbesserungswürdig	<input type="checkbox"/> nicht eindeutig
Komplexität der Aufgabe	<input type="checkbox"/> hoch	<input type="checkbox"/> angemessen	<input type="checkbox"/> nicht angemessen
Zeitvorgabe (5 Minuten) zum Lösen der Aufgabe:	<input type="checkbox"/> angemessen	<input type="checkbox"/> eher knapp	<input type="checkbox"/> nicht ausreichend
Homogenität der Lösungs-/Antwortmöglichkeiten	<input type="checkbox"/> angemessen	<input type="checkbox"/> eher nicht angemessen	
Bewertungs-Checkliste:			
Aufteilung der Punkte	<input type="checkbox"/> sinnvoll	<input type="checkbox"/> verbesserungswürdig	<input type="checkbox"/> eher nicht sinnvoll
Klarheit der Kriterien zur Punktevergabe	<input type="checkbox"/> eindeutig	<input type="checkbox"/> verbesserungswürdig	<input type="checkbox"/> nicht eindeutig
Kommentare:			

Gesamteinschätzung der Station
1= sehr gut, 5 = mangelhaft

1 2 3 4 5

Pre-OSCE quality assurance

review of checklists

Formale Kriterien:			
Eindeutigkeit der Aufgabenstellung für den Prüfling	<input type="checkbox"/> eindeutig	<input type="checkbox"/> verbesserungswürdig	<input type="checkbox"/> nicht eindeutig
Komplexität der Aufgabe	<input type="checkbox"/> hoch	<input type="checkbox"/> angemessen	<input type="checkbox"/> nicht angemessen
Zeitvorgabe (5 Minuten) zum Lösen der Aufgabe:	<input type="checkbox"/> angemessen	<input type="checkbox"/> eher knapp	<input type="checkbox"/> nicht ausreichend
Homogenität der Lösungs-/Antwortmöglichkeiten	<input type="checkbox"/> angemessen	<input type="checkbox"/> eher nicht angemessen	
Bewertungs-Checkliste:			
Aufteilung der Punkte	<input type="checkbox"/> sinnvoll	<input type="checkbox"/> verbesserungswürdig	<input type="checkbox"/> eher nicht sinnvoll
Klarheit der Kriterien zur Punktevergabe	<input type="checkbox"/> eindeutig	<input type="checkbox"/> verbesserungswürdig	<input type="checkbox"/> nicht eindeutig
Kommentare:			

Gesamteinschätzung der Station 1= sehr gut, 5 = mangelhaft	1	2	3	4	5
--	---	---	---	---	---

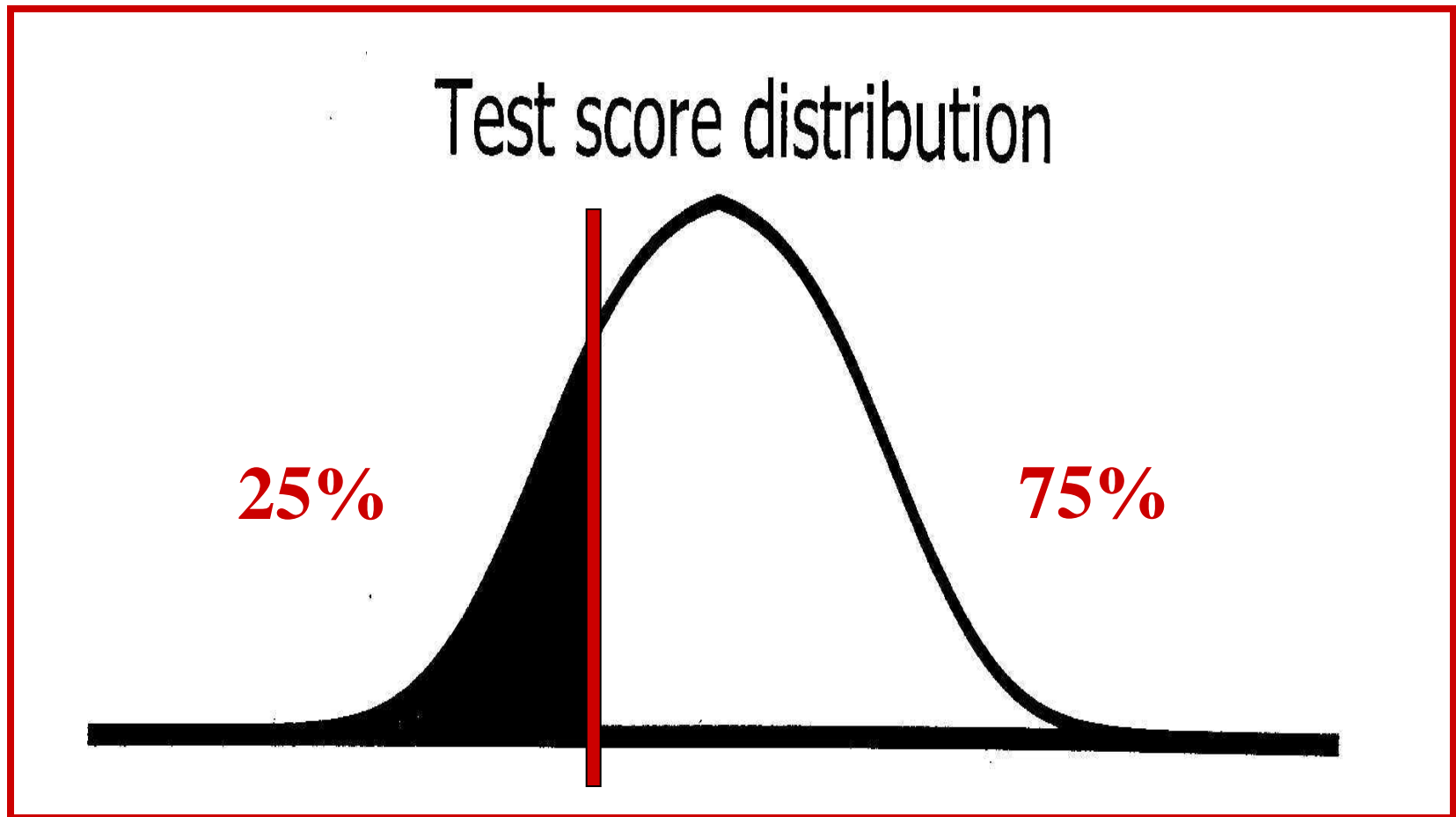
Pre-OSCE quality assurance standard setting

Different methods of standard setting

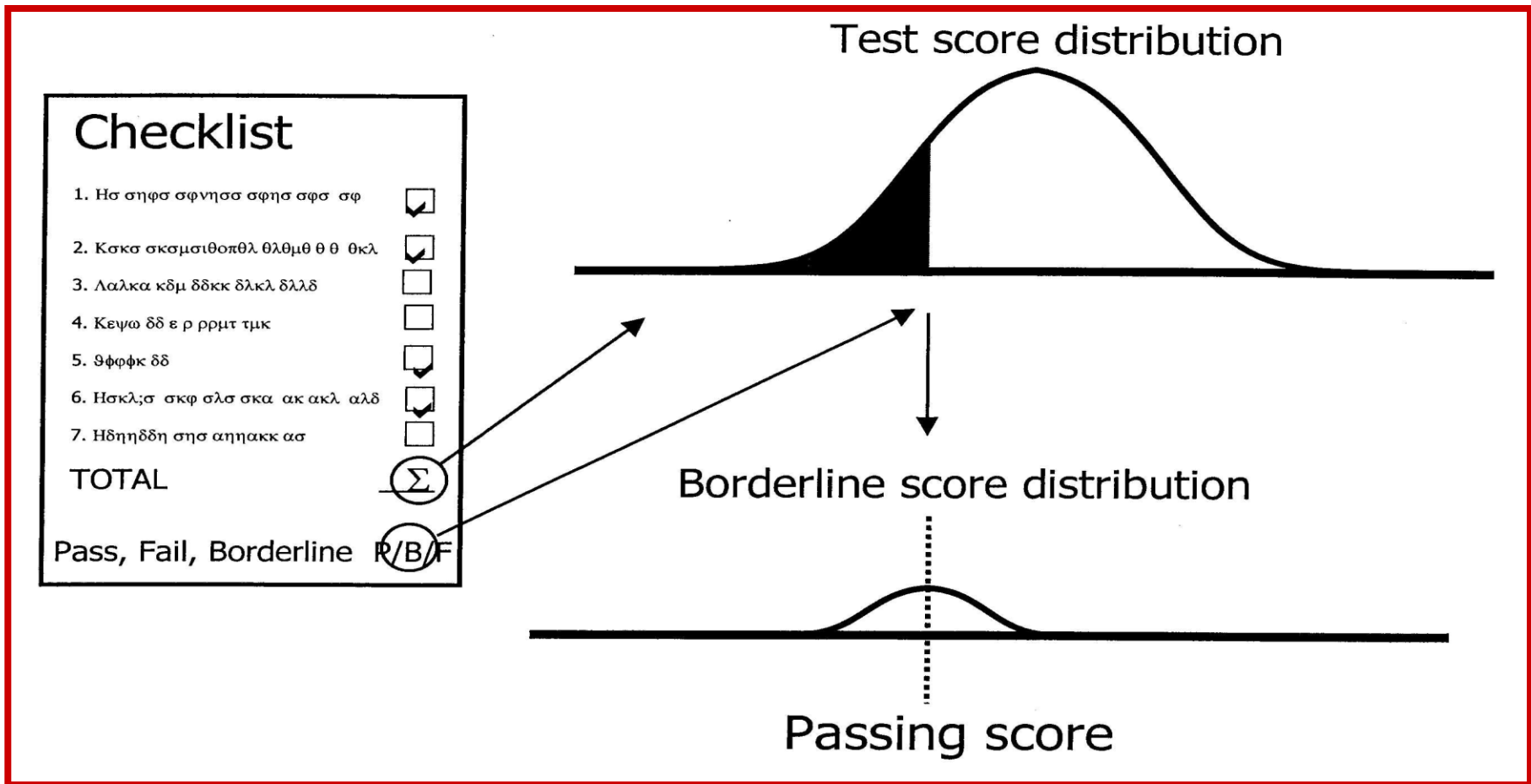
- **holistic method**
- **relative, norm-orientated methods**
- **absolute, criteria-orientated methods**

Pre-OSCE quality assurance

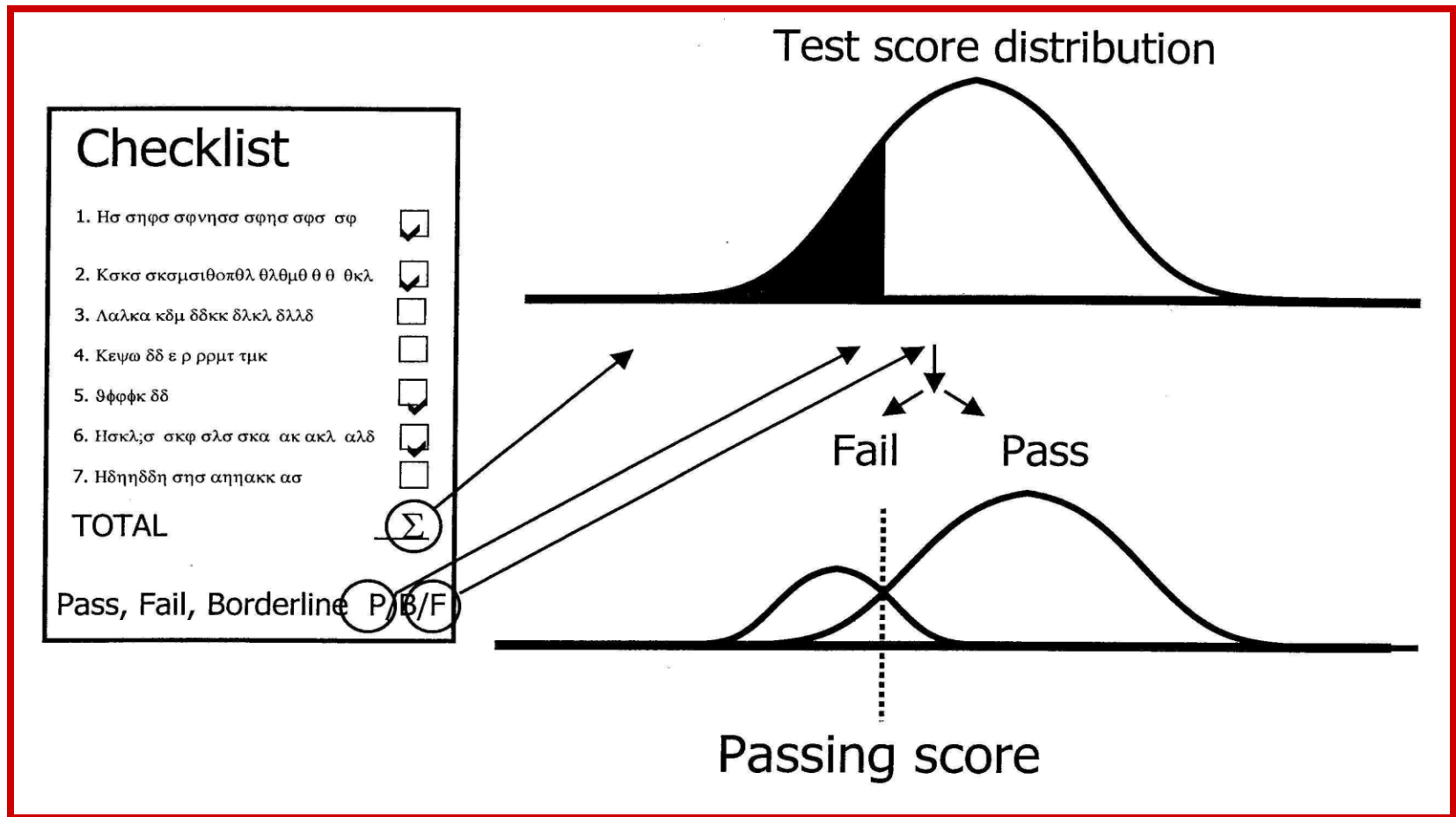
holistic method



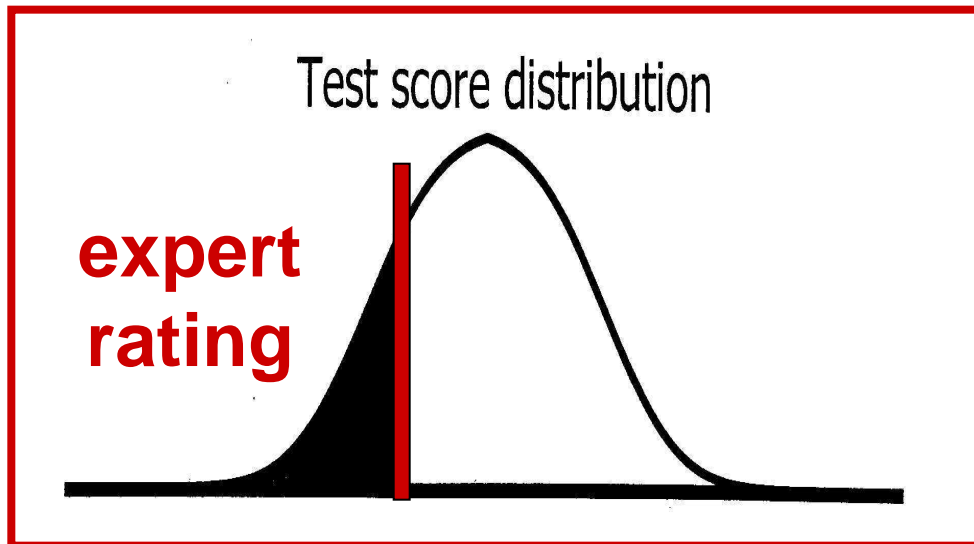
Pre-OSCE quality assurance borderline method



Pre-OSCE quality assurance contrast method



Pre-OSCE quality assurance modified Angoff procedure



- expert committee
- discussion about the „borderline-candidate“
- rating of items
- discussion within the expert team about discrepancies

Pre-OSCE quality assurance

rater training

Agenda

- 1) Short communication on OSCE**
 - use of checklists**
 - rating**
 - prompting**
 - feedback**
- 2) Video presentation of OSCE station**
- 3) Discussion on named aspects**

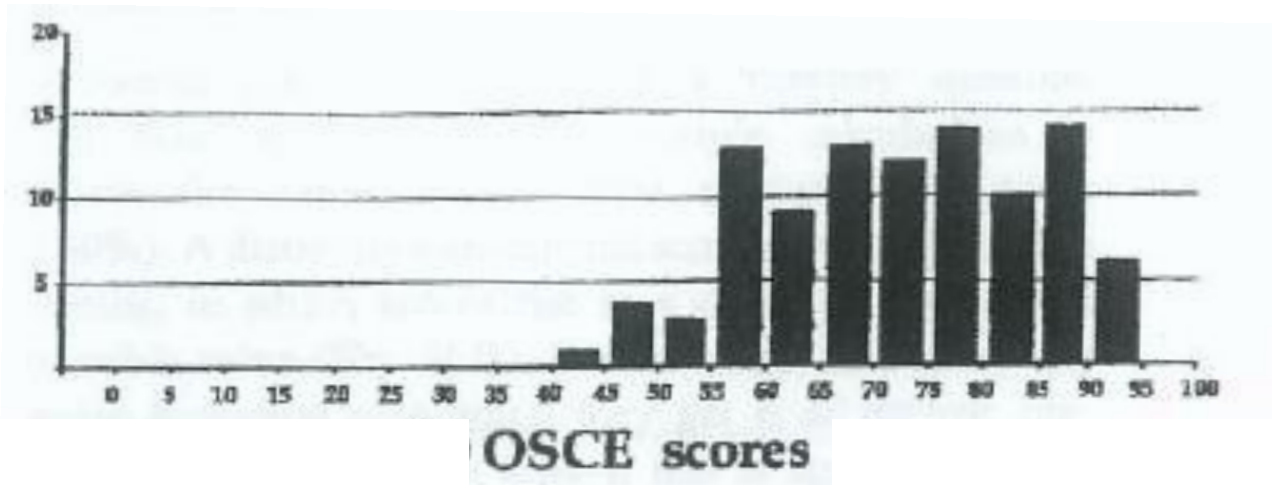
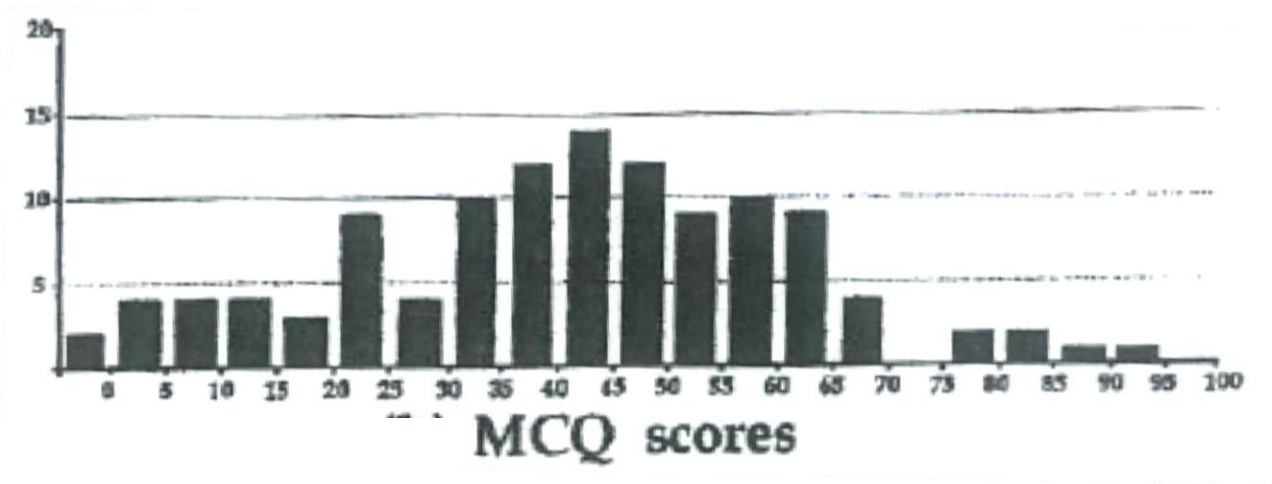
Post-OSCE quality assurance



Post-OSCE quality assurance analyses

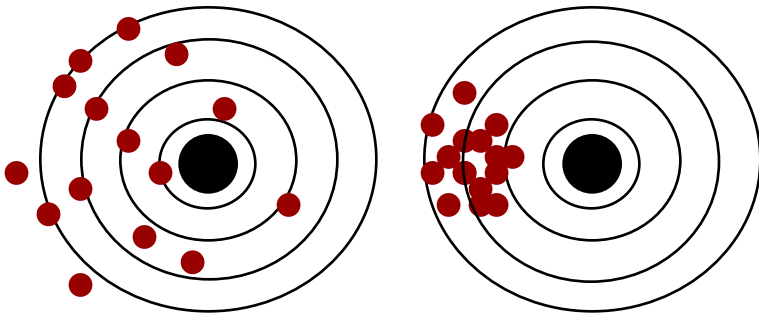
	difficulty	selectivity
ECG interpretation	0,817	0,457
endocrinology	0,803	0,390
physical examination	0,932	0,256
communication	0,784	0,324
pulmonology	0,812	0,433
rheumatology	0,904	0,418
general internal medicine	0,742	0,345
nephrology	0,747	0,425
psychosomatics	0,795	0,331
haematology/gastroenterology	0,853	0,273
– lymph nodes	0,885	0,463
– icterus	0,765	0,315
– ulcer	0,800	0,811

Post-OSCE quality assurance score distribution



OSCE – aspects of reliability

Reliabilität

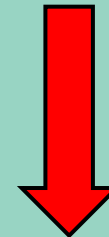


OSCE – aspects of reliability

→ high number of stations

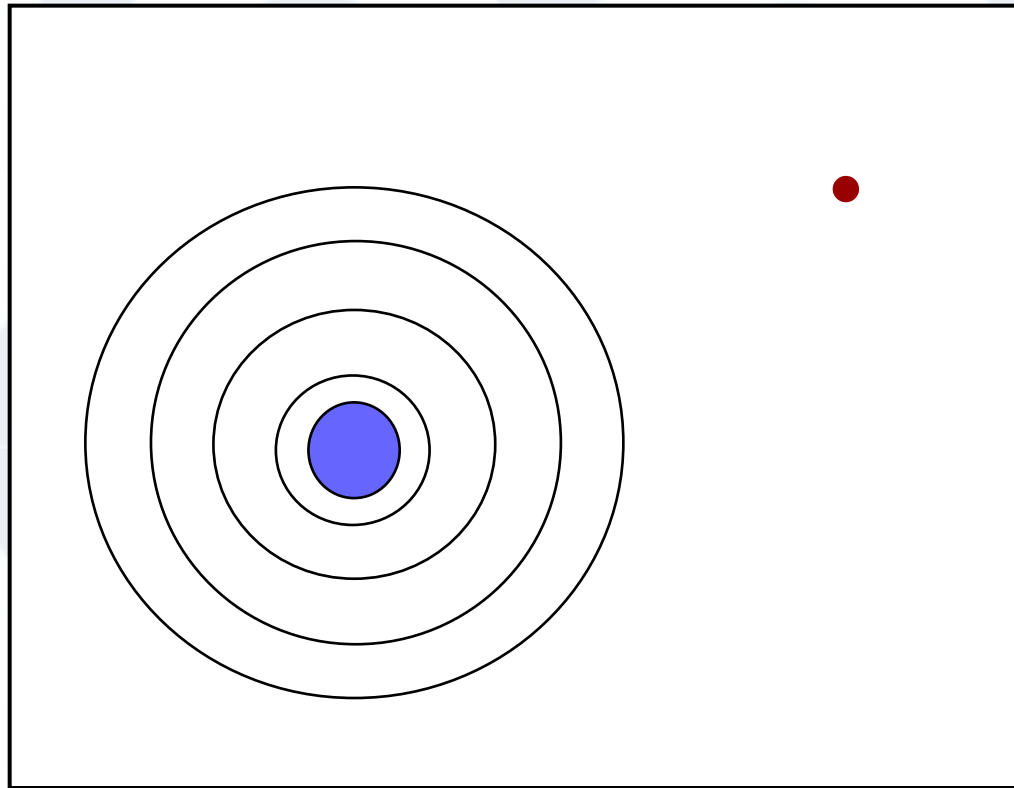
→ pre-review of stations

Semester	Reliabilität (Cronbachs α)
Wintersemester 2005/2006	0,772
Sommersemester 2006	0,715
Wintersemester 2006/2007	0,839
Sommersemester 2007	0,829



→ combination of different assessment formats

OSCE – aspects of validity



OSCE – aspects of validity

face / content validity

assessment in simulated „real-life-situation“

→ face validity

OSCE – aspects of validity

face / content validity

assessment in simulated „real-life-situation“

→ face validity

learning goals → assessment goals → blueprint

→ content validity

OSCE – aspects of validity

predictive validity

OSCE predictor for later clinical performance

→ predictive validity

Campos-Outcait & Gordon, Fam Med (1999)
Graham & Anderson, J Dent Educ (2013)

OSCE – aspects of validity

predictive validity

OSCE predictor for later clinical performance

→ predictive validity

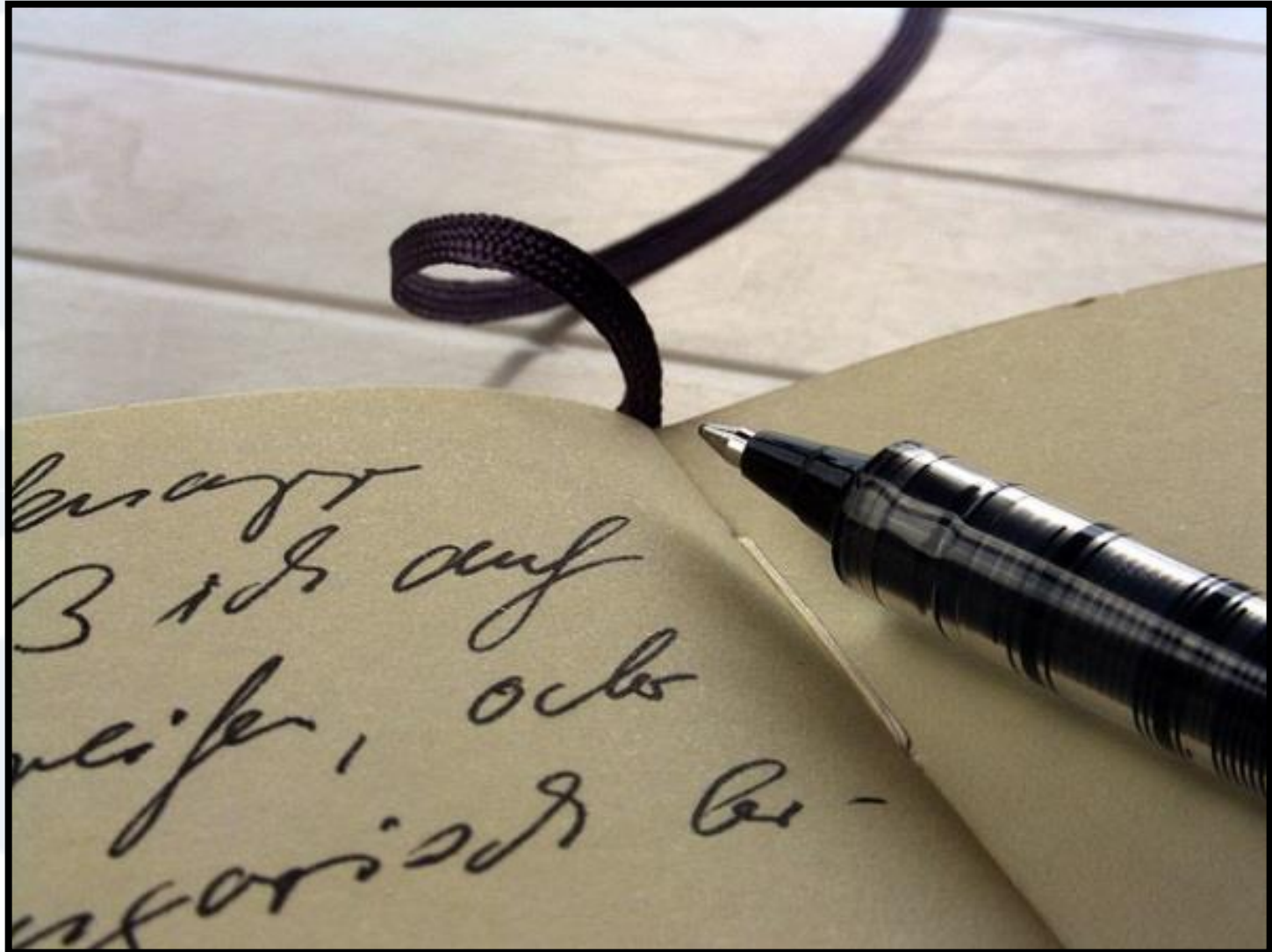
Campos-Outcait & Gordon, Fam Med (1999)
Graham & Anderson, J Dent Educ (2013)

high correlation with assessment of clinical
emplacements, expert ratings and other
rating instruments

→ concurrent validity

Hodges & McNaughton, Acad Med (1998)
Lieberman & Lang, Acad Radiol (2001)
Scheffer & Ortwein, Adv Health Sci Theory Prac (2007)

OSCE – a tool for research



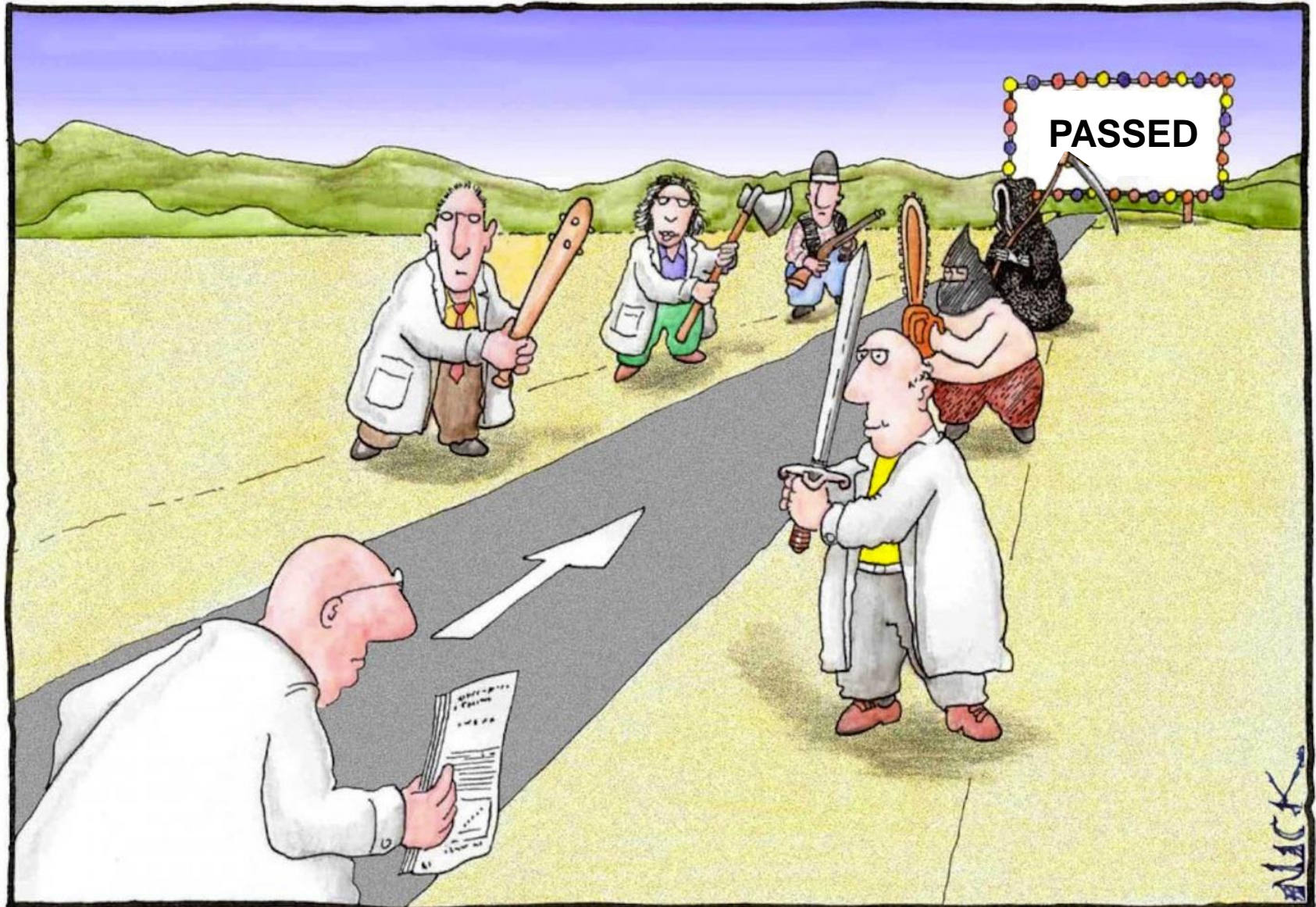
OSCE – a tool for research reporting checklist



**Patricio M, Juliao M, Fareleira F, Young M,
Norman G, Vaz Carneiro A.
A comprehensive checklist for reporting the
use of OSCEs.
Medical Teacher, 2009.**

- problem I:** missing data on feasibility, reliability, validity
- problem II:** inconsistent data (lack of standardisation, structure; statistics)

OSCE – future challenges



OSCE – future challenges



Recommendations and guidelines for medical assessment



Initiatives for assessment guidelines on European level



Item Management Systems (IMS) providing reviewed checklists



International licensing examination (e.g. USMLE)



Interdisciplinary 17 station OSCE as final state examination since 2001



THANK YOU VERY MUCH

FOR YOUR ATTENTION!

christoph.nikendei@med.uni-heidelberg.de

