

Assessment and evaluation workshop

In postgraduate family medicine
training programmes

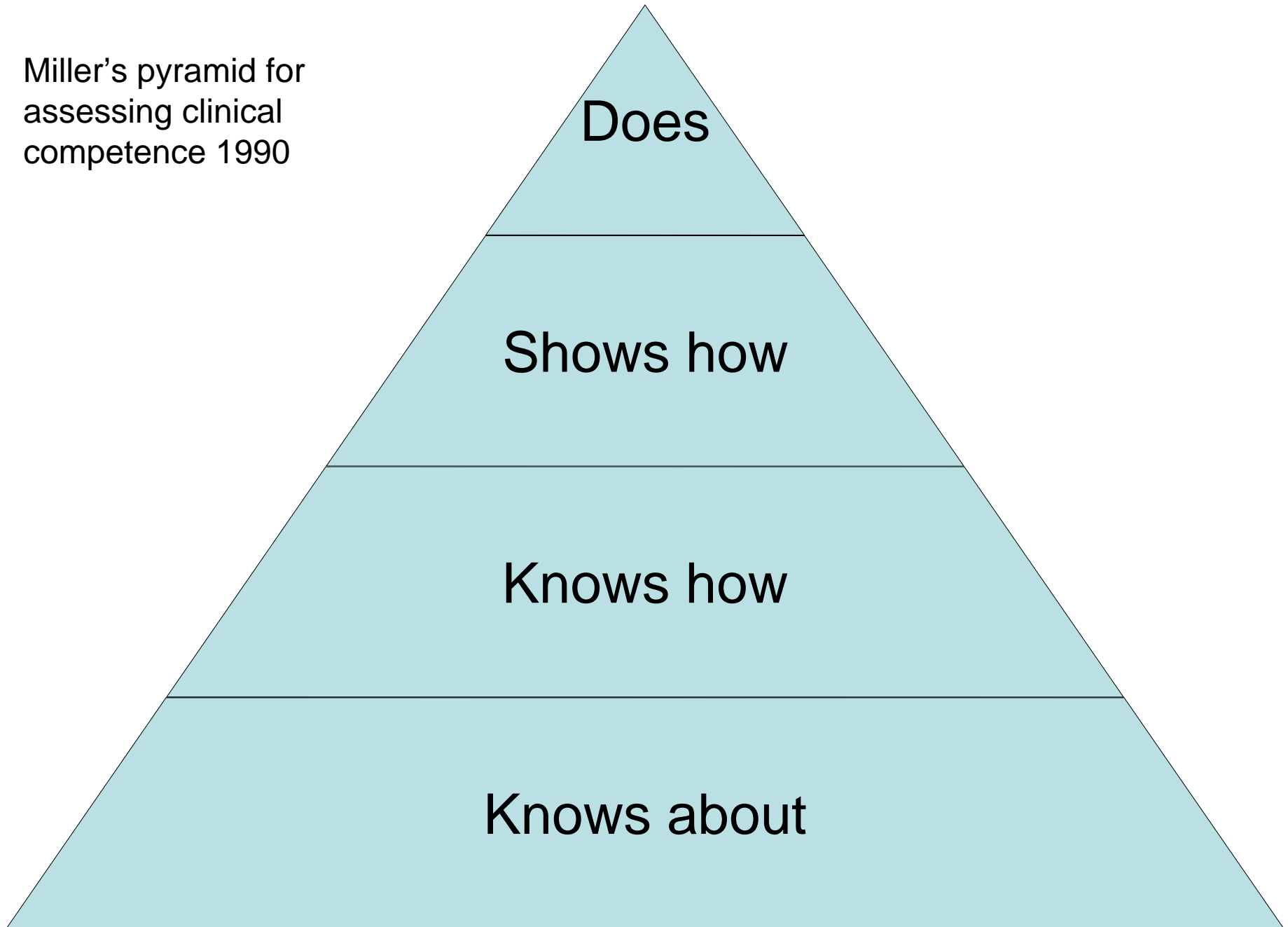
Assessment

The key issues in assessment

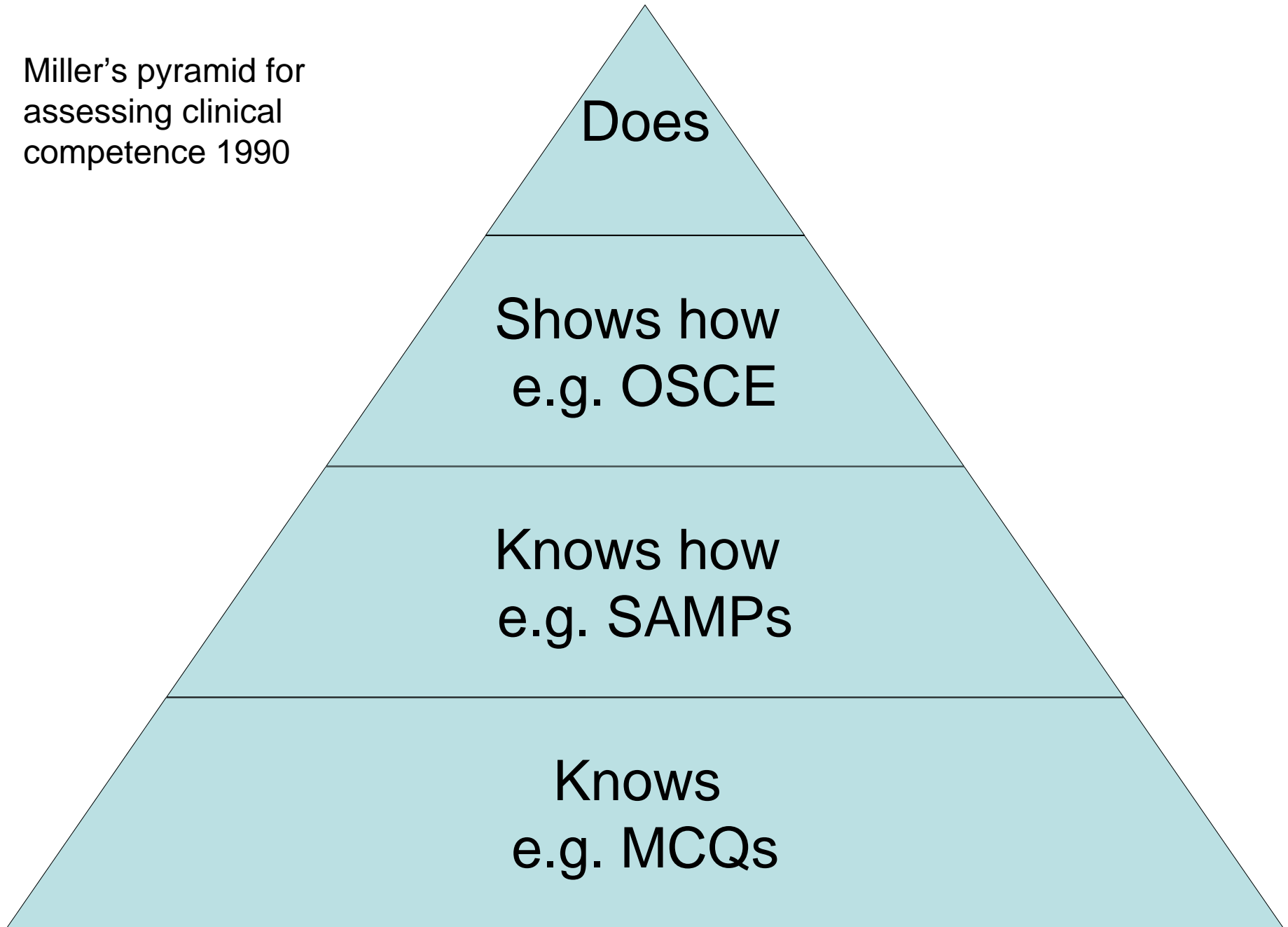
- Outcome-based
- Formative vs. summative
- Reliability
- Validity
- Acceptable
- Practical
- Educational impact

- The *reliability* of a test describes the degree to which the test consistently measures what it is supposed to measure
- The *validity* of a test is a measure of the degree to which the test actually measures what it is supposed to measure

Miller's pyramid for
assessing clinical
competence 1990



Miller's pyramid for
assessing clinical
competence 1990



Planning

Domains for assessment

- A: Knowledge (factual, evolving, evidence-based)
- B: Application of knowledge (decision-making, problem-solving)
- C: Communication skills
- D: Values and attitudes
- E: Clinical and procedural skills
- F: Research skills
- G: Teaching and facilitation of learning skills
- H: Management and administration skills

Modules	Primary focus	Secondary focus
Continuous tasks	B, H	A, C, D
Final assignments	B, H	A, D
MCQ exam	A, H	B
Examination		
SOO Part 1	C	A, B, D
OSCE Part 1	E	A, B, C, D
Oral Part 1	D	B, C
Presentation Part 2	G	F
Clinical Part 2	E	B, A, D
Research thesis	F	

Work-based assessment

Basis of judgement

- Clinical outcomes
- Process of care
- Volume of care

Data collection

- Medical records
- Logbooks
- Observation
 - Direct
 - Peer rating
 - Patient rating

Skills-based assessment

- Long case and oral
- Short case and oral
- Objective Structured Clinical Examination (OSCE)
- OSCE adaptations

Written assessment

- True / false questions
- Single best option MCQ
- Multiple option MCQ
- “Key feature” questions
- Matching questions
- Short answer open questions
- Essays

Evaluation

“Evaluation should be designed at the start of developing a curriculum, not added as an afterthought”

Jill Morrison BMJ 2003 326 385

The purpose of evaluation is to...

- Ensure teaching is meeting students' learning needs
- Ensure teaching is linked to the learning outcomes
- Identify areas where teaching can be improved
- Inform the allocation of resources
- Provide feedback to teachers
- Support promotion of teachers
- Facilitate revision and development of the curriculum

Levels of focus in evaluation

1. Learners reactions
2. Change in attitudes and perceptions
3. Gain in knowledge and skills
4. Change in professional behaviour
5. Change in organisational practice
6. Benefits to patients

Levels of formality

1. National requirements e.g. HPCSA
2. Institutional standard requirements e.g. external examiner report, student results
3. Programme's ad hoc feedback e.g. interviews, surveys

How do you do this in your family
medicine programme now?

What tips would you give others
on how to do this well?

OSCE

Ingredients of the OSCE

- Student demonstrates a skill (shows how) on model or standard/real patient
- Student is assessed across a range of problems (reliability)
- Topics are aligned with curriculum / outcomes (validity)
- 8-20 stations lasting 5-10 minutes
- Checklist
- Global rating scale

Evaluation of stations

- Is the scenario adequately explained?
- Is the task appropriate to the level of training being assessed?
- Is a skill actually being tested?
- Are the items specific enough and observable?
- Are marks allocated appropriately?

Review the examples of
stations

What issues affect the reliability
and validity of the OSCE?

Reliability and validity

Reliability:

- Too few stations
- Too little time
- Checklist items too easy or too hard
- Unreliable patients
- Unreliable examiners
- Administrative problems i.e. noise

Validity

- Are the topics important / relevant to the learning outcomes?
- Have the topics been taught in the curriculum?
- Are the stations standardised?

Setting standards

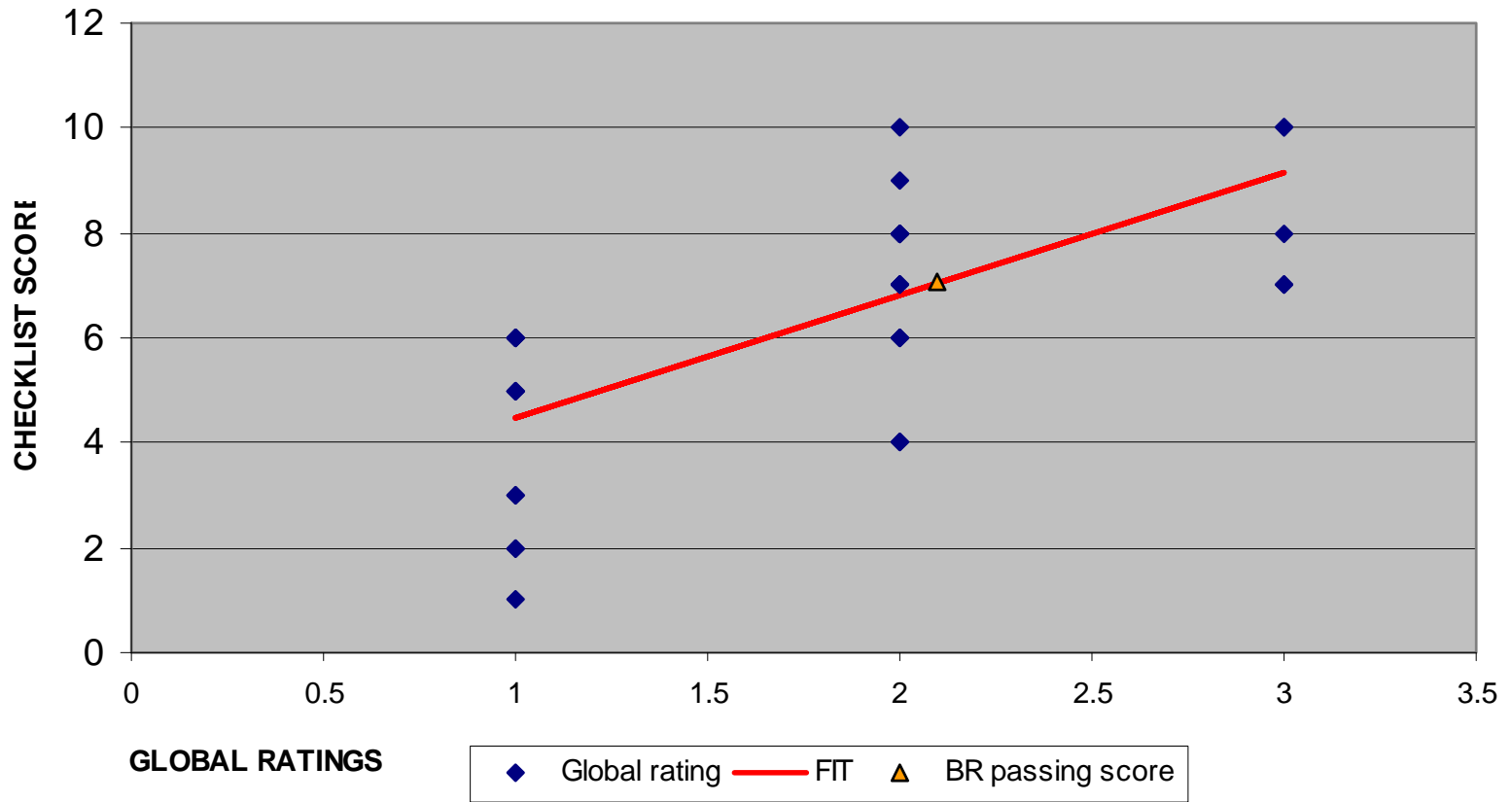
Angoff procedure

- Judges (≥ 12) imagine the performance of a borderline candidate
- Judges record the chance of this candidate performing each item
- Judges discuss and reach consensus
- Results in item score, station pass mark, OSCE pass mark

Borderline regression method

- At each station during OSCE judge completes checklist and global rating
- After exam checklist score and rating score are regressed to determine the station pass mark for the borderline candidate
- Results in station pass mark, OSCE pass mark

CHECKLIST VS GLOBAL RATINGS GRAPH: STATION 2



Other practical issues

- Preparation of stations, patients, examiners
- Logistic load high
- Cost
- Limitations: Fragmented tasks, reward thoroughness not experience, not all skills can be included