

# Assessment in medical education

Subha Ramani, MBBS, MMEd, MPH  
Brigham and Women's Hospital &  
Harvard Medical School  
[sramani@bwh.harvard.edu](mailto:sramani@bwh.harvard.edu)



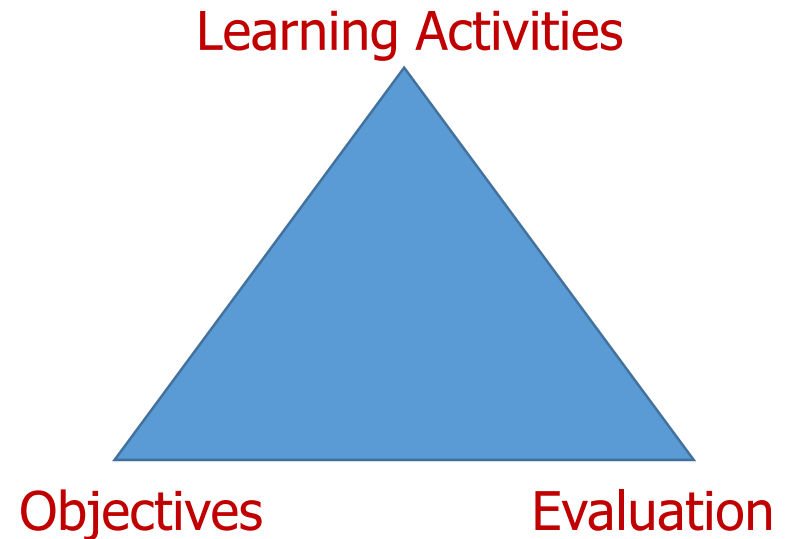
# Goals

- Reflect on key principles of learner assessment in medical education
- Discuss the levels of Miller's pyramid and assessment strategies to target various learning outcomes
- Reflect on methods to assess skills and clinical performance

# **General principles**

# Purpose of Assessment

- Communicate educational goals
- Identify areas of learner deficiency
- Determine effectiveness of course or curriculum
- Determine readiness to practise
- Increase learner self-reflection



# Exercise

- Reflect on the kinds of assessments you use when assessing your learners in the work setting
  - What is the purpose of these assessments?
  - How did you choose the method / items?
  - How did you decide on a rating scale / checklist?

# Shifts in Assessment

From assessing to learn what learners do not know



To assessing to learn what learners understand

From end-of-term assessments by teachers



To learners engaged in ongoing assessment of their work and others

From judgmental feedback that may harm learner motivation



To behavioral feedback that empowers and motivates learners



**ASSESSMENT OF  
LEARNING**



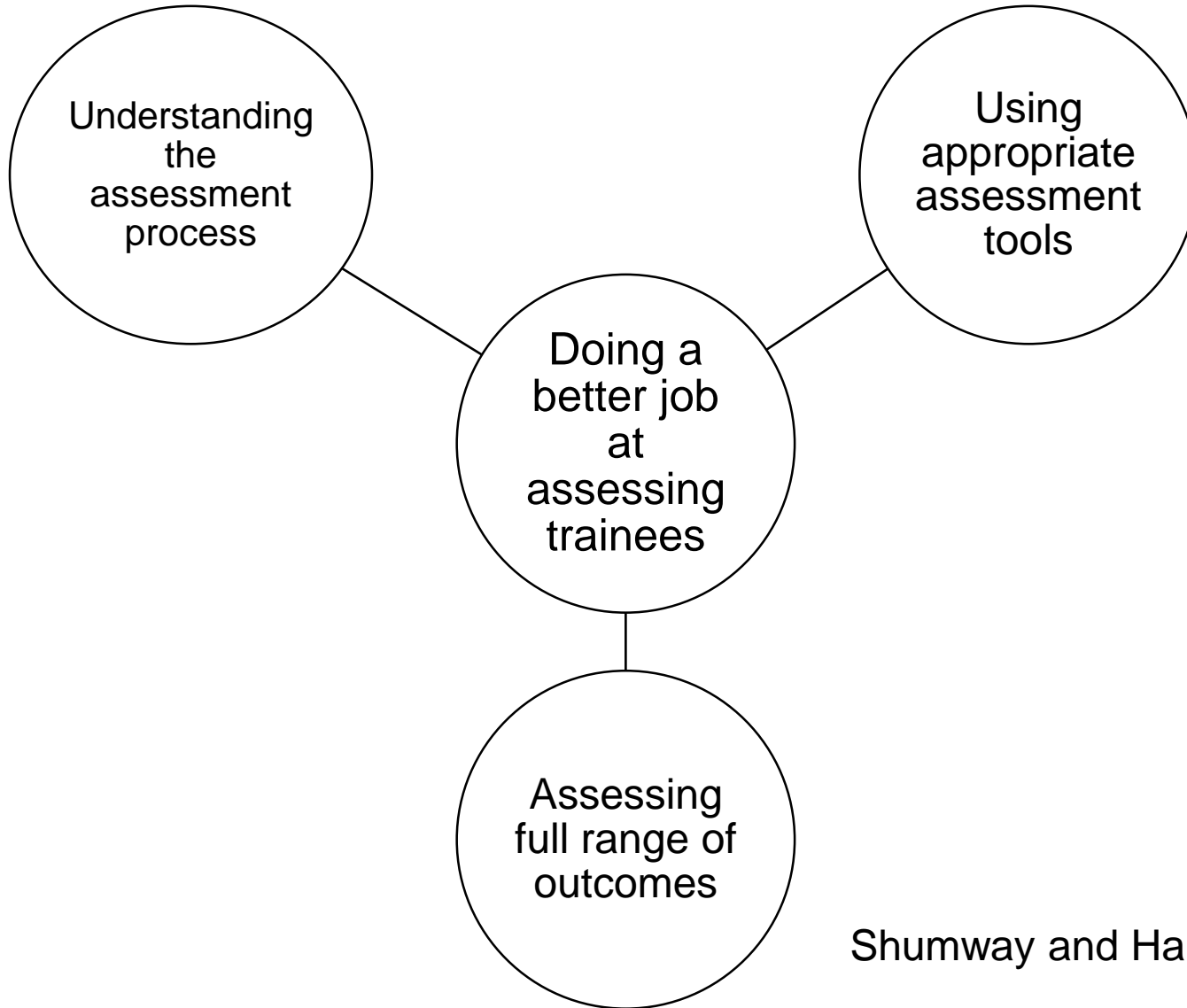
**ASSESSMENT FOR  
LEARNING**

# Things to consider



1. How does the assessment scheme deal with all the curricular outcomes?
2. How do individual assessments contribute to overall curricular outcomes?
3. How have the targeted outcomes been met in individual assessments?
4. Are trainees clear about what is expected in the assessment?
5. Are examiners trained to carry out the assessment consistently?
6. Do assessors have clear guidelines for calibrating assessments?

# Multidimensional model for WBA



Shumway and Harden 2003



# Key features of assessment methods

- Reproducible (reliable), valid, feasible, fair, and beneficial to learning (van der Vleuten 1996)
- Content and form of assessments need to be aligned with their purpose and desired outcomes
- Broad sampling is needed to achieve an accurate representation of ability (multiple biopsies)
- Constructed according to clearly defined standards and derived using systematic and credible methods
- Consider stakeholders

# Criteria for good assessment

- Validity
- Reliability
- Feasibility
- Educational effect
- Acceptability

# **Assessment: Why, What, Who, How**

# Types

Use	Individual (Faculty / Learners)	Program
Formative	<p>Evaluation of individual learner or faculty member used to help individual improve performance:</p> <ul style="list-style-type: none"><li>▪ Identify areas for improvement</li><li>▪ Specific suggestions for improvement</li></ul>	<p>Evaluation of a program that is used to improve program performance:</p> <ul style="list-style-type: none"><li>▪ Identification of areas for improvement</li><li>▪ Specific suggestions for improvement</li></ul>
Summative	<p>Evaluation used for judgments or decisions about the individual:</p> <ul style="list-style-type: none"><li>▪ Verification of achievement</li><li>▪ Motivation of individual to maintain or improve performance</li><li>▪ Certification of performance for others</li><li>▪ Grades</li><li>▪ Promotion</li></ul>	<p>Evaluation of a program used for judgment / decisions re the program or its developers:</p> <ul style="list-style-type: none"><li>▪ Judgment re success, efficacy</li><li>▪ Allocation of resources</li><li>▪ Motivation/recruitment of faculty and learners</li><li>▪ Satisfying outside requirements</li><li>▪ Dissemination: presentations</li></ul>

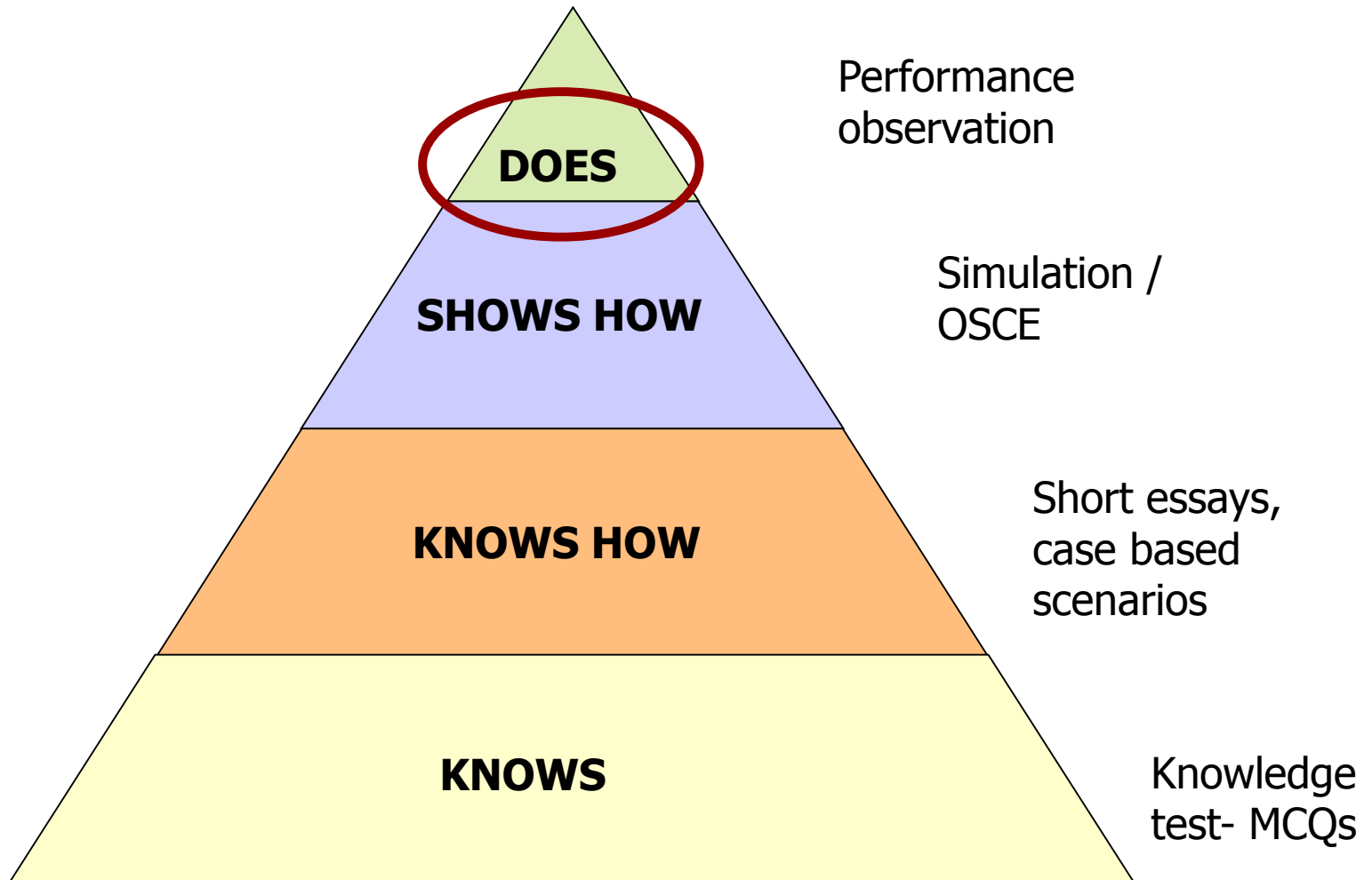
# Exercise 1

**Pick one assessment that is formative and one that is summative in your work environment. Compare and contrast the two.**

	Summative	Formative
Purpose	?	?
Outcome	?	?
Effect on learner	?	?

# **Assessment methods**

# Workplace assessment



# Assessment of Skills: Shows How- OSCE

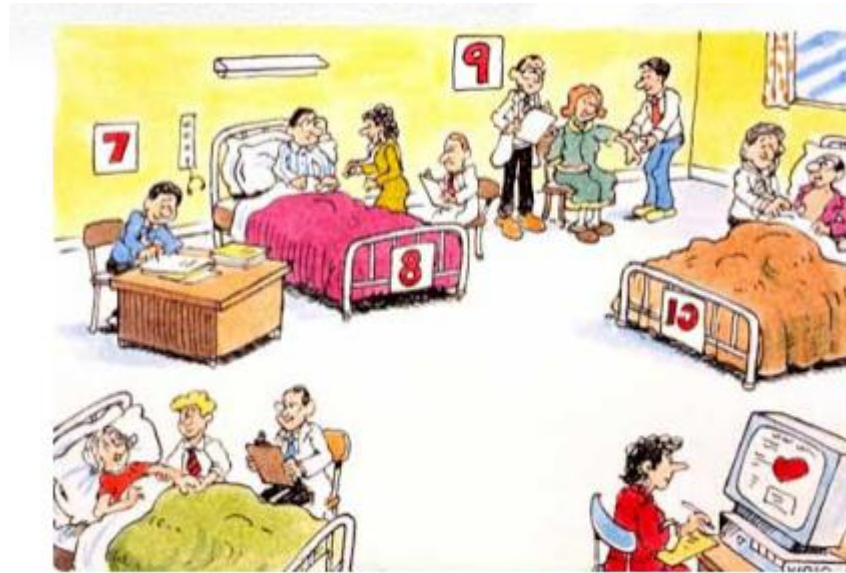


Image Credit: Ronald Harden (r.m.harden@dundee.ac.uk)



# History

- Described in 1975
- Developed in response to criticism of the long and short case assessment
- Multiple, short stations
- Assesses specific skills
- Uses standardized checklists

# Sources of error

- Item
  - Checklist not anchored or anchors not clear
  - Item-specific variance
- SP
  - Portrays case incorrectly
  - Variable portrayal of case by different SPs
- Raters
  - Systematic rater error
  - Ratings based on irrelevant traits
  - Inconsistent ratings
- Case
  - Case specificity
  - Case task ambiguous

# OSCE To-Do

- Preparation and planning
  - Lead person
  - Team
  - Administrator
- Exam details
  - Blueprint
  - Length, number of stations
  - Skills to be assessed
  - Case writers
  - Pilot
- Scoring
  - Checklist
  - Global
- Examiners
  - Faculty development
- Patients
  - Recruitment
  - training
- Running the OSCE
  - The circuit
  - Rest
  - Space and equipment
  - Briefings
  - Quarantine

# **Assessment in the Workplace: DOES**

# Exercise

- Reflect on the kinds of assessments you have seen used in workplace assessment
  - What is the purpose of these assessments?
  - What are the challenges?
  - How should the items / domains be selected?
  - What types of assessment methods exist?

# Dreyfus Developmental Model

- **Novice**
  - Rule driven, cannot prioritize information, cannot see big picture
- **Advanced beginner**
  - Able to sort through rules, see relevance, starts to see big picture
- **Competent**
  - Sees patterns, uses analytic reasoning and pattern recognition
- **Proficient**
  - Extrapolates from known to unknown problems, comfortable with ambiguity
- **Expert**
  - Knows when pattern does not fit the expected, notices the unexpected, action and intuition
- **Master**
  - Goes beyond big picture into culture and context of each situation, emotional engagement, reflects in and on action

What do we mean by a “Developmental Approach to Assessment?”

# Developmental Approach

- For trainees:
  - Formative
  - Knowledge of current performance
  - Helps with goal setting
- For assessors:
  - Knowledge of curricular goals
  - Snap shot of trainee performance at a given level of training
  - Behaviours clearly observable
  - Helps with feedback
- For training programmes:
  - Review of trainee performance and reporting to the accrediting organizations
  - Determine whether trainees overall are progressing



# ACGME Outcomes

1. Patient care
2. Medical knowledge
3. Interpersonal and communication skills
4. Professionalism
5. Systems based practice
6. Practice based learning and improvement

## MEDICAL KNOWLEDGE

Displays broad fund of knowledge in basic and clinical science

1 Never 2 3 4 5 6 7 8 9 Always N/A

Demonstrates analytic clinical reasoning skills

1 Never 2 3 4 5 6 7 8 9 Always N/A

Demonstrates ability to synthesize and accurately summarize complex patients in oral presentations

1 Never 2 3 4 5 6 7 8 9 Always N/A

## PATIENT CARE

Makes diagnostic/therapeutic decisions based on best available evidence and patient preference

1 Never 2 3 4 5 6 7 8 9 Always N/A

Performs comprehensive and time efficient medical interviews and physical exams

1 Never 2 3 4 5 6 7 8 9 Always N/A

Demonstrates clinical judgment in sorting out major from minor issues

1 Never 2 3 4 5 6 7 8 9 Always N/A

# Patient Care

## 1. Gathers and synthesizes essential and accurate information to define each patient's clinical problem(s). (PC1)

Critical Deficiencies			Ready for unsupervised practice	Aspirational
Does not collect accurate historical data	Inconsistently able to acquire accurate historical information in an organized fashion	Consistently acquires accurate and relevant histories from patients	Acquires accurate histories from patients in an efficient, prioritized, and hypothesis-driven fashion	Obtains relevant historical subtleties, including sensitive information that informs the differential diagnosis
Does not use physical exam to confirm history	Does not perform an appropriately thorough physical exam or misses key physical exam findings	Seeks and obtains data from secondary sources when needed	Performs accurate physical exams that are targeted to the patient's complaints	Identifies subtle or unusual physical exam findings
Relies exclusively on documentation of others to generate own database or differential diagnosis	Does not seek or is overly reliant on secondary data	Consistently performs accurate and appropriately thorough physical exams	Synthesizes data to generate a prioritized differential diagnosis and problem list	Efficiently utilizes all sources of secondary data to inform differential diagnosis
Fails to recognize patient's central clinical problems	Inconsistently recognizes patients' central clinical problem or develops limited differential diagnoses	Uses collected data to define a patient's central clinical problem(s)	Effectively uses history and physical examination skills to minimize the need for further diagnostic testing	Role models and teaches the effective use of history and physical examination skills to minimize the need for further diagnostic testing
Fails to recognize potentially life threatening problems				

## **1. Patient Care: History, Physical, Clinical reasoning for presenting cardiac problems, examples listed in Table 1 (PC 1).**

- Does not gather pertinent data; unable to recognize patients' central cardiac problem
- Gathers and synthesizes clinical data; demonstrates clinical reasoning for simple clinical presentations
- Consistently gathers and synthesizes data; demonstrates clinical reasoning for complex clinical presentations with guidance
- Consistently gathers and synthesizes data; demonstrates independent clinical reasoning for complex clinical presentations
- Models and teaches the art of clinical data collection and clinical reasoning in complex presentations
- Unable to evaluate

## **2. Patient Care: Appropriate patient management including urgent/emergent issues (PC2).**

- Unable to outline an appropriate management plan for common and/or urgent cardiac problems
- Provides appropriate management for simple clinical problems; with guidance initiates management of emergencies
- With guidance provides appropriate management of complex problems; initiates management of emergencies independently
- Initiates appropriate management of complex problems and emergencies independently; modifies plans based on additional data or changes in clinical course
- Models and teaches appropriate management of simple, complex, and emergent problems making necessary modifications as appropriate
- Unable to evaluate

# Goals of direct observation

- Make a judgement on a learner's performance of a specific task in the clinical/workplace setting
- Provide guidance to the learner through effective feedback to improve future performance of the task

## Rate this trainee exercise

- Poor
- Fair
- Good
- Very
- Excellent

# Assessment methods

- Faculty ratings
- Chart stimulated recall
- Peer ratings
- Self assessment
- Staff / team evaluations
- Direct observation
- Patient surveys
- Quality ratings
- Portfolios

**Essential elements:**

Feedback to learners

Faculty development

# Challenges

- Impression management
- Observer skill
- Subjectivity: Chain of events
  - Learner behavior
  - Faculty observation
  - Faculty interpretation
  - Recording of data
  - Judging of data



# Threats

- Content underrepresentation
- Construct-irrelevant variance
  - Rater errors- halo, leniency / severity, central tendency, idiosyncratic
  - Examinee bluffing
  - Extraneous characteristics

# Summary

1. Cover broad range of clinical situations
2. Multiple Assessors
3. Short and focused instruments
4. Separate formative from summative
5. Recorded promptly
6. Direct observations critical
7. Group review and consensus for summative assessment
8. Educate Assessors
9. Assessors should observe specific behaviors
10. Anchor ratings