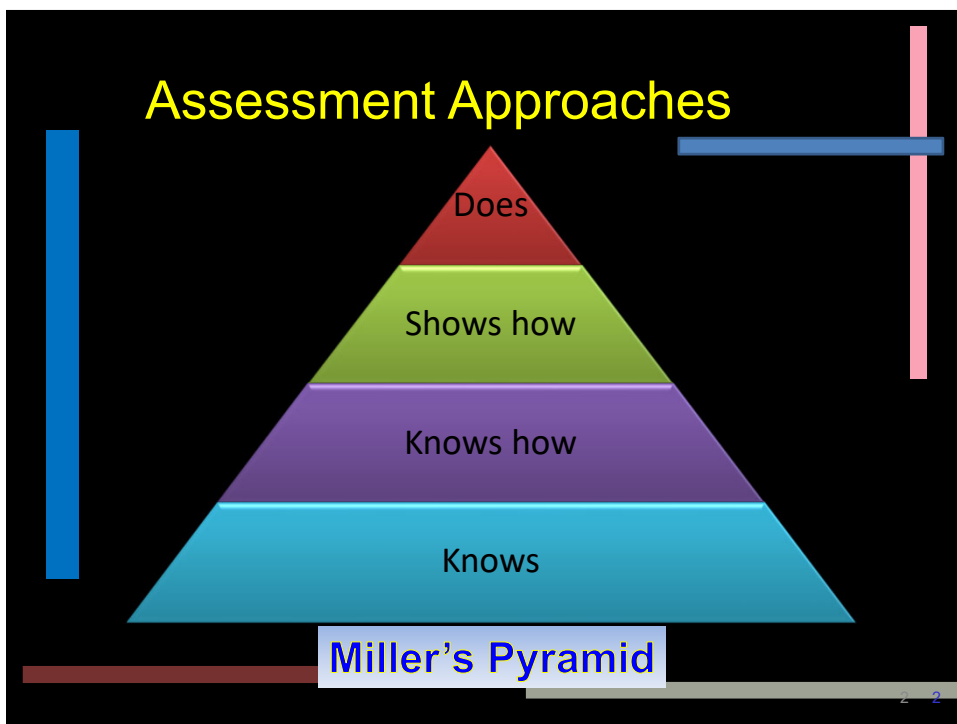


How to Choose Assessment Methods

นพ. เชิดศักดิ์ ไอร่มณีรัตน์
ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล
มหาวิทยาลัยมหิดล



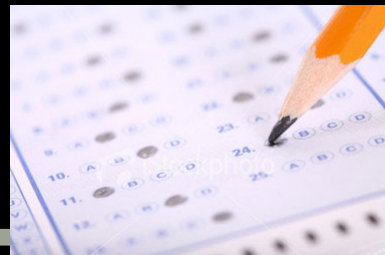
Group Activity

- ยกตัวอย่างวิธีการประเมินในแต่ละระดับของ Miller's pyramid

(5 นาที)

Multiple-Choice Questions

- Selected Response Exam
 - True/False
 - Simple True/False items
 - Multiple true/false items (K-type)
 - One best response
 - Standard MCQ
 - Extended matching items



Multiple-Choice Questions

- Advantages
 - Objective scoring
 - High internal consistency reliability
 - Strong research evidence to support its validity
 - Efficiency in testing and scoring

Multiple-Choice Questions

- Limitations
 - Cueing of correct answer
 - Random guessing
 - Testing of trivial knowledge
 - Difficulty of development of good MCQ items
 - Unable to assess psychomotor and other non-cognitive abilities

Constructed Response Items

- Constructed response items ask examinees to create responses rather than select answers from lists of possible answers.

Comparison

	Selected Response	Constructed Response
Measured construct	Concrete knowledge, basic interpretation, some applications	Complex cognitive ability: problem solving, interpretation, decision making
Item construction	Simple	Complex
Cost of scoring	Low	Expensive
Type of scoring	Objective	Subjective
Rater effects	No effect	Significant factor
Reliability	High	Low

Adapted from Table 3.2 In Haladyna TM, Developing and validating multiple-choice Test items, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004.

CR: Strengths

- Examinees' responses are non-cued: more authentic
- Able to measure higher-order cognitive tasks: application, analysis, synthesis, and evaluation
- Motivation for clinical learning

CR: Limitations

- Difficult to develop and score
- Inefficient exam format
- Expensive
- Subjectivity
- Low reliability
- Construct underrepresentation
- Cannot assess affective or psychomotor abilities

Traditional essay questions

- Long essay examinations
 - An exam is consist of a few open-ended essay questions, each requires lengthy written responses from examinees
- Short essay examinations
 - An exam is consist of many open-ended essay questions, each requires short written answer consisting of a sentence or two

Comparison

	Long Essay	Short Essay
Content coverage	Narrow	Broad
Item development	Easy	Difficult
Scoring guideline development	Very difficult	Easier
Students' answers	Infinite possibilities	More focused scope
Reliability	Very low	Low
Time used	More	Less
Good use	Assessment of complex cognitive abilities: analysis, synthesis, evaluation, and presentation of ideas	Assessment of simplified, structured problems with limited answers

Shortcomings

- Tendency to assess recall knowledge
- Different from actual clinical practice

13

Modified Essay Question

- คำถามบรรยายชนิดที่มีการประยุกต์ให้เสมือนการแก้ปัญหาผู้ป่วยในชีวิตจริง ซึ่งแพทย์จะไม่มีข้อมูลทั้งหมดตั้งแต่เริ่มเห็นผู้ป่วย แต่แพทย์ต้องค่อยๆสืบค้นหาข้อมูลเพิ่มเติมและตัดสินใจทำการวิเคราะห์และแก้ปัญหาไปที่ละขั้นตอน
- การแก้ปัญหาของผู้ป่วยรายหนึ่งๆประกอบไปด้วยหลายขั้นตอน เมื่อทำแต่ละขั้นตอนแล้ว ไม่สามารถย้อนกลับไปแก้ไขสิ่งที่ได้ทำไปก่อนหน้านี้ได้

CRQ

- **Constructed Response Question**

- **Adaptation of MEQ**

- หนึ่งในข้อที่มีหลายขั้นตอนในการแก้ปัญหา แต่ให้ตอบคำถามทุกข้อในช่วงเวลาเดียวกัน
 - ข้อดี: ลดปริมาณงานให้เจ้าหน้าที่คุมสอบ
 - ข้อจำกัด: มีปัญหาในการประเมินความเข้าใจ และคิดคะแนน

15

Oral Examination

An examination conducted by spoken communication, focusing on the assessment of abilities/skills that cannot be evaluated by paper examination (e.g., communication, problem solving, decision making)

Advantages

- Can be used to assess things that other assessment methods cannot address (e.g., decision making, communication)
 - “ Direct personal contact allows assessment of appearance, manner, personality, alertness, confidence, honesty, self-awareness and other aspects of values, and attitudes (Cox, 1982). ”
- An examiner can adjust the questions based on candidates' responses

Limitations

- Expensive
- Time consuming
- Difficult to administer
- Limited content validity
- Low reliability
 - Require subjective assessment
 - Small number of cases

OSCE

- Objective Structured Clinical Examination
- Assessment of clinical skills
 - History taking
 - Physical examination
 - Communication skills
 - Procedural skills
 - Interpretation of medical investigations
 - Ordering of medical treatment

Assessment

19

OSCE

- Advantages
 - Can assess clinical skills, technical skills, communication skills
 - Standardization of cases, observations
 - Supporting research evidence

Medical Competencies

OSCE

- Limitations
 - Expensive
 - Time consuming
 - Difficult to administer
 - Many potential sources of CIV: SPs, raters, cases, scoring sheets
 - Construct underrepresentation

Medical Competencies

Performance Ratings

Ratings of learners' performance based on observing real-life practice by attending faculty members

Competence and Performance

- Competence = The capacity of a person to perform a defined task (Maximal ability)
- Performance = The actual act in carrying out or execute the duty (Typical ability)

Performance Ratings

- Advantages
 - Typical performance assessment
 - Motivation for clinical learning
 - Inexpensive

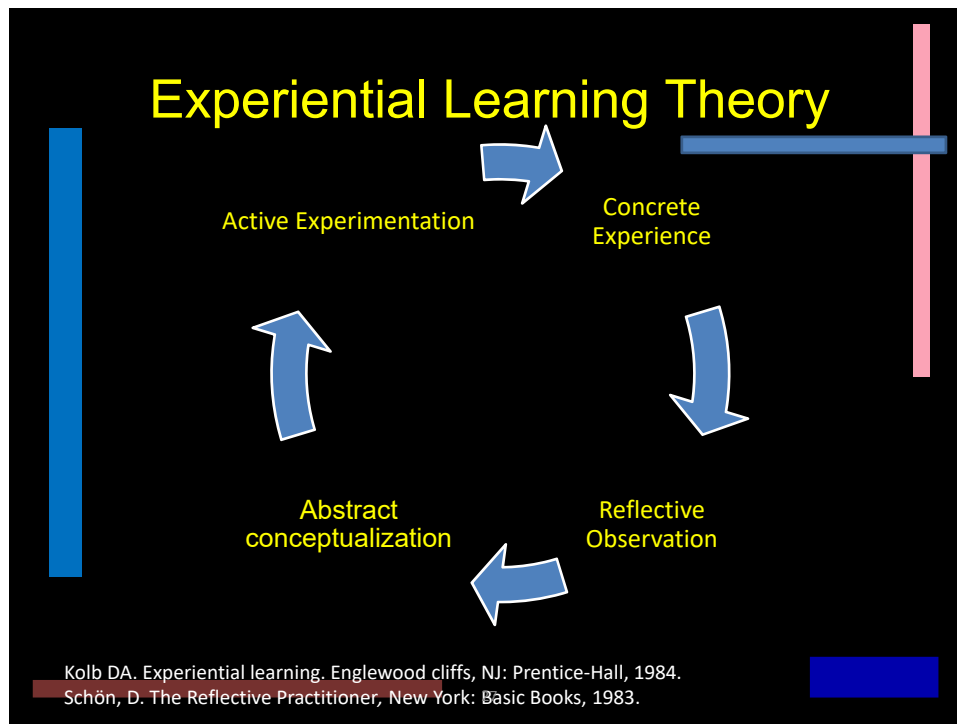
Performance Ratings

- Disadvantages
 - Subjective ratings
 - Unstructured settings
 - Adequacy of observation
 - Low reliability

Portfolio

- A systematic collection of student work and related material that depicts a student's activities, accomplishments, and achievements in one or more school subjects. The collection should include evidence of student reflection and self-evaluation, guidelines for selecting the portfolio contents, and criteria for judging the quality of the work.

Venn JJ..Assessing students with special needs, 2nd ed.. Upper Saddle River, NJ: Merrill., 2000



- ## Reflective Thinking
- Recognizing the nature of the problem
 - Responding based on comparing similarities and differences between possible solutions
 - Examining the intended and unintended consequences of the solution selected
- Ross D. Programmatic structures for the preparation of reflective teachers. In Clift R, Houston W, Pugach M, editors. Encouraging reflective practice in education. Teachers College Press, NY 1990.

Examples of Portfolio Contents

- Case reports
- Research projects
- Student's reflection on learning experience
- Critical-incident reports
- Personal memo on difficulties or successes during their training

Advantages of Portfolio

- Assess and promote critical thinking
- Encourage students to become accountable and responsible for their learning
- Initiation a discussion between student and mentor
- Facilitate reflection
- Accommodate diverse learning styles
- Monitor students' progress over time
- Assessment in real-life settings

Advantages of Portfolio (2)

- Use multiple methods of assessment
- Take into account multiple assessors
- Integrate learning and assessment
- Promote creativity and problem solving
- Can be used for both formative and summative
- Can be used to assess attitudes and personal development
- Provide vital information for student diagnosis

Davis MH, Ponnamparuma GG. Portfolio assessment. JVME 2005; 32: 279 – 83.

Disadvantages of Portfolios

- For summative assessment, students may be reluctant to reveal weaknesses.
- Privacy and confidentiality of information on portfolio
- Difficulty in verification of the materials (plagiarism?)
- Workload (students, teachers)
- Low inter-rater reliability

Davis MH, Ponnamparuma GG. Portfolio assessment. JVME 2005; 32: 279 – 83.

Workplace-based Assessment

- A number of assessment methods, suitable for providing feedback based on observation of trainee performance in the workplace.
 - Mini-clinical Evaluation Exercise (mini-CEX)
 - Clinical Encounter Card (CEC)
 - Blinded Patient Encounter (BPE)
 - Direct Observation of Procedural Skills (DOPS)
 - Procedure-based Assessment (PBA)
 - Case-based Discussion (CbD)
 - Multisource Feedback (MSF)

WPBA: Characteristics

- เป็นการประเมินที่ให้นักศึกษาเป็นผู้เริ่มต้น
- นักศึกษาสามารถขอให้อาจารย์ประเมินได้ตลอดช่วงเวลาปฏิบัติงาน
- ประเมินซ้ำได้ โดยคิดคะแนนจากครั้งที่นักศึกษาทำคะแนนดีที่สุด
- จุดมุ่งหมายสำคัญคือการเปิดโอกาสให้อาจารย์ได้ **feedback**
- เมื่อประเมินเสร็จแล้ว นักศึกษารับผิดชอบนำคะแนนที่ได้ไปรายงาน

WPBA: Strengths

- Validity: assessment of “does” level
- Identify students in needs of support early
- Provide feedback
- Create a nurturing culture
- Samples widely in many workplaces
- Utilize a number of assessors

General Medical Council. Workplace based assessment: A guide for implementation, April 2010.

WPBA: Limitations

- Low reliability
- Can be opportunistic
- Trainees may delay or avoid assessment
- Learner dependent and vulnerable
- Require time and training
- Bias due to the interaction between trainers and trainees

General Medical Council. Workplace based assessment: A guide for implementation, April 2010.

หลักในการเลือกวิธีประเมินผล

1. คำนึงถึงผลลัพธ์ที่ต้องการวัดว่าเป็นความรู้ ความสามารถในระดับใดของ Miller's pyramid
2. คำนึงถึงหลักในการประเมินผลที่ดี (criteria for good assessment)
3. เลือกวิธีการที่บรรลุวัตถุประสงค์ได้ด้วยความประหยัดทรัพยากร

Questions & Comments

Cherdsak.ira@mahidol.ac.th