Case Western Reserve University
School of Medicine, Cleveland, Ohio, USA

Assessments that Matter: Beyond Knowledge & Recall of Factual Information

Klara K. Papp, PhD & Amy Wilson-Defosse, PhD

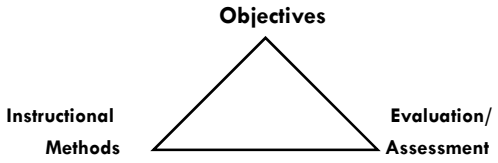
Objectives

- Identify top reasons for using tests in the medical curriculum assessing basic biomedical sciences
- Consider how to better align course goals and objectives with instructional methods and assessments
- Classify test questions according to cognitive domain
- Describe indicators of quality

Why Test?

- Assess understanding of key concepts
- Assess application of concepts taught
- Assess problem solving, analytical reasoning, critical thinking
- Identify students who need remediation
- Inform decisions about the curriculum
- Identify students who should not become doctors
- Assess retention & integration of material

Alignment



How is your alignment?

- What goals do your assessments communicate to students?
- Indicate how, if at all, you might change assessments to achieve greater alignment with course goals?

CWRU's re-alignment

Pre-WR2

- Exams developed on a lecture-by-lecture basis where each faculty submitted 2-3 MCQ-items per hour—lacked integration and perspective
- Minimal investment of faculty time
- Students 'binged and purged'
- Challenged questions that weren't covered in lecture
- No effort to assess material covered earlier

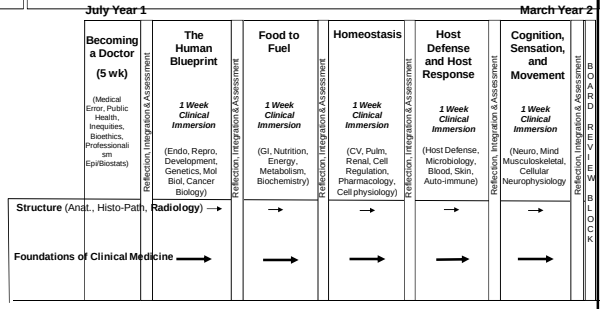
Western Reserve$_2$ Curriculum

Western Reserve2 Curriculum

Year I	Year II	Year III	Year IV
Foundations of Medicine and Health (20 months, including vacation)		**Core Clinical Rotations** **IQ+ Program** (48 weeks, flexible scheduling)	
		Research & Scholarship (16 week block + electives, flexible scheduling)	
			Advanced Clinical and Scientific Studies Areas of Concentration Electives (10 months, flexible scheduling)

WR$_2$ Foundations of Medicine and Health

WR2 Foundations of Medicine and Health



Assessment in WR$_2$: Competencies

WR2 Competencies

1. Medical Knowledge
2. Clinical Mastery
3. Interpersonal and communication skills
4. Professionalism
5. Practice based learning and improvement
6. Systems based practice
7. Leadership and Health Advocacy
8. Life-long learning and personal development
9. Research & Scholarship

Guiding Principles

- Assessment drives learning
- Facilitate continuous rather than episodic learning
- Promotes supportive and collaborative relationship among students
- Provides frequent formative feedback
- Provides periodic judgments of students progress based on multiple methods of assessment both qualitative and quantitative

Early Discussions

- Multiple methods of assessment
- Provide opportunities for frequent self-assessment
- Assess not just medical knowledge but also interpersonal skills, qualities of mind, professionalism
- Assess writing skills, coherency, use of language

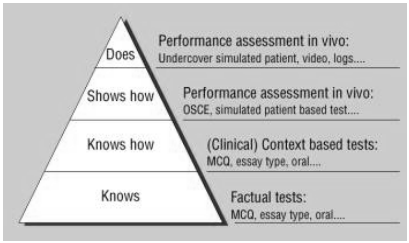
Assessment Tools

On-going Self-Assessment	***Assessment by Faculty***	***Boards Preparation***
Weekly **MCQs** - (**Formative**)	**IQ Group** Assessment **FCM** Assessment	NBME **cumulative** achievement tests (**Formative**)
Weekly Synthesis **Essays** **(Formative)**	End of Block Summative Synthesis **Essay** Exam	
Learning Objectives (**Formative**)	Structure **Practical Exam**	
Professional Learning Plans (PLP) (Formative)	End of Year **Portfolios**	

Design Features

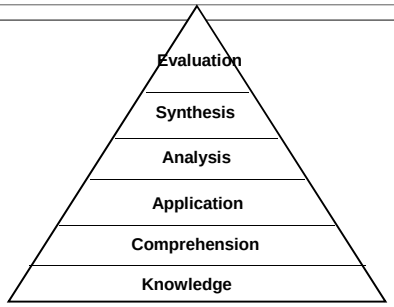
- Formative Assessments –Throughout the Course
  - Weekly 2 Essay Questions
- Summative Assessments—End of Course
  - Summative Synthesis Essay Questions (SSEQs)

Miller's Triangle[1]



[1]GE Miller. The assessment of clinical skills/competence/performance. Acad Med 1990; 65: 63S-67S.

Taxonomy of the Cognitive Domain*



*Bloom B, et al. Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. NewYork: Toronto: Longmans, Green, 1956.

Beyond Knowledge & Comprehension

- Knowledge and comprehension
- Application
- Analysis
- Synthesis
- Evaluation

Identify the cognitive level

- What area is supplied with blood by the posterior inferior cerebellar artery?

Cited from: Swanson & Case. Assessment in basic science instruction: Directions for practice and research. Adv Hlth Sciences Educ 1997; 2: 71-84.

1. Knowledge & comprehension
2. Application
3. Analysis
4. Synthesis
5. Evaluation

Identify the cognitive level

A 62-year-old man develops left-sided limb ataxia, Horner's syndrome, nystangmus, & loss of appreciation of facial pain and temperature sensations. What may be causing this patient's problem?

Cited from: Swanson & Case. Assessment in basic science instruction: Directions for practice and research. Adv Hlth Sciences Educ 1997; 2: 71-84.

1. Knowledge & comprehension
2. Application
3. Analysis
4. Synthesis
5. Evaluation

Identify the cognitive level

Decreased delivery of blood to the kidneys, e.g., by renal arterial stenosis, results in a significant systemic increase of which of the following hormones?

A antidiuretic hormone
B atrial natriuretic factor
C endothelin
D rennin

1. Knowledge & comprehension
2. Application
3. Analysis
4. Synthesis
5. Evaluation

Clinical Scenario:

Patricia Moran is a 26-year-old woman who has been treated for Graves' disease with propylthiouracil (PTU) in the past. She stopped taking the pill about a year ago.

Since then she has lost 20 pounds and says that she is unusually nervous. She is "hot and sweaty" all the time, has hand tremors, and heart palpitations. Her periods are also scanty and she can go 2-3 months in between periods. She has noticed an increased frequency of bowel movements.

On physical exam, she has a rapid heart rate of 120. Her skin is warm and moist, and feels unusually smooth and velvety. She has a fine tremor in her hands. She also has bilateral exophthalmos (eyes protrude from the orbits). Her thyroid gland is large and there is a thyroid bruit (rushing of blood when auscultating the thyroid with a stethoscope). Her deep tendon reflex (DTR) relaxation phase is very brisk.

Example taken from Block 2: Summative Synthesis Essay Questions

Identify the cognitive level

- Explain what has gone awry with this patient's thyroid gland and why her eyes are bulging forward.

1. Knowledge & comprehension
2. Application
3. Analysis
4. Synthesis
5. Evaluation

Identify the cognitive level

- Using your knowledge of the normal physiologic effects of thyroid hormone on the body, explain how thyroid hormone causes the signs and symptoms noted above.

1. Knowledge & comprehension
2. Application
3. Analysis
4. Synthesis
5. Evaluation

Indicators of Quality

- Validity
- Reliability

Reliability

- The extent to which a test provides the same results with repeated measurement
  - Stability (across measures)
    - Test-retest reliability
  - Consistency (within a measure)
    - Internal consistency (Cronbach's alpha)

Validity

Extent to which the test measures the concepts it was designed to measure.

- Distinct from reliability
- Measure may be reliable but not valid

Types of Validity

- Face Validity
- Content Validity
- Criterion Validity
- Convergent Validity
- Discriminant Validity
- Construct Validity

Face Validity

On the "face" of it, does the test assess the desired construct?

- For example, does a test of cancer genetics "seem" related and relevant to the discipline?
- Fairly weak form of evidence of validity

Content Validity

Does the measure capture all of the relevant 'domains' of the subject?

- For example, to measure RA, we need to measure joint count, morning stiffness, grip strength—all must be included for good content validity
- May be difficult to assess if construct is not well-defined (e.g., analytical thinking, professionalism)

Criterion Validity

- How does the test perform compared to an independent 'gold standard'?

Predictive Validity

- How does the measure predict what it theoretically should be able to predict?
  - Achievement in the internal medicine core clerkship should predict how well a student will do in an acting internship or during 1st year of residency.
  - Use of USMLE Step 2 CK and CS to predict future success

Concurrent Validity

How well can the measure distinguish between groups that it theoretically should be able to distinguish between.

- Measure to assess clinical skills should be able to distinguish between those who perform well clinically and those who are having trouble.
- Create a new clinical skills exam. Administer it to students with high and low clerkship grades—how well does it discriminate?

Convergent Validity

How similar is the new measure to other existing measures which assess the same construct.

- Compare USMLE Step 1 Scores to Comprehensive Tests of Basic Science

Discriminant Validity

Examines the degree to which the test differs (diverges) from other measures of different constructs

- Compare scores on USMLE Step 1 Scores with USMLE Step 2 CS. Low correlations between the two would suggest that they are each measuring different constructs.

Summary

- Assessments must be aligned with instructional goals and objectives and with teaching methods
- The methods of assessment used drive student learning
- "Learning is a change in the ***quality of one's thinking*** rather than a change in the quantity of one's knowledge."
- Reliable and valid assessments require investments of faculty time

Comments? Questions?

Thank you!