Test Blueprinting I:
Selecting an Assessment Method

Test Blueprinting I:
Selecting an Assessment Method

_____

Test Blueprinting I: Selecting an Assessment Method

Objectives

By the end of this lesson, you should be able to:

- o Recognize the role of **learning objectives** in deciding what knowledge and skills to assess
- o Classify learning objectives according to **skill domain** (cognitive, affective, and psychomotor) and **level of learning** (recognition/recall or application/critical thinking)
- o Consider a wide range of assessment methods and select a method that is optimal for the skill being assessed

Introduction

Consider yourself in the following scenario: You teach an introductory course in clinical diagnosis and want to ensure that your first-year students have mastered auscultation. The previous instructor for the course left you a bank of 70 multiple-choice questions (MCQs) and 20 short-answer questions on auscultation. How would you assess student mastery of auscultation? What assessment method would you use?

The choice of the optimal assessment method will depend on the type of knowledge or skill being assessed and the purpose of the assessment itself. **Factual knowledge** about auscultation might be effectively tested with MCQs or other written formats. But when it comes to assessment of the student's **procedural skills**, such as properly placing the stethoscope or accurately identifying aortic stenosis when heard, will a set of MCQs still be sufficient? Does this change if the exam has a passing standard and high stakes for the students who fail? In this lesson, you will learn to use a simple framework to ensure the method of assessment aligns with the type and purpose of the assessment.

Role of Learning Objectives

When developing an assessment, the first step is to identify what you want to assess; these are the **learning objectives** associated with the test. These are specific, detailed sets of knowledge and skills you expect students to know and be able to do after completing a course. Some examples of objectives that relate to cardiac auscultation are shown below, and you may be able to come up with others:

- List three causes of aortic stenosis.
- Offer to and properly drape a female patient prior to auscultation.

- Compare and contrast characteristics of a functional systolic murmur and a pathologic murmur.

- Correctly identify common systolic murmurs on an audio recording.

- Explain the maneuver for differentiating aortic stenosis from hypertrophic obstructive cardiomyopathy.

- Perform the maneuver to differentiate aortic stenosis from hypertrophic obstructive cardiomyopathy.

- Select the most cost-effective method to evaluate a patient with a suspected heart murmur.

- Position the patient and transducer and perform an echocardiogram to visualize the aortic valve in the short axis view.

The skills described in the objectives are all important. However, keep in mind that these skills are fundamentally different from each other. Some require **knowing** while others require **doing**. Cognitive psychologists have come up with multiple frameworks for categorizing skills. You may have heard of Miller's Pyramid, which nicely separates the **knowing** from the **doing** (Miller, 1990). It consists of four stages: knows, knows how, shows, and does. This framework can serve as a useful method to develop and organize a list of learning objectives for a course.

Figure 1. Miller's Pyramid
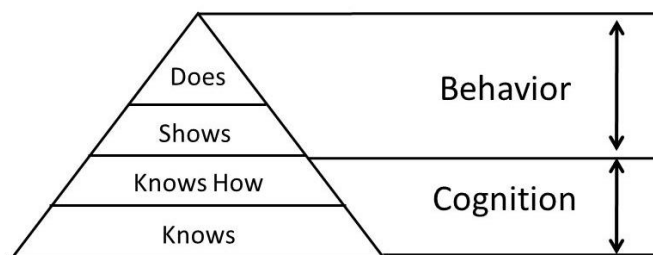
Miller GE. The Assessment of Clinical Skills / Competence / Performance, Academic Medicine, 65:9, S63-S67.

_____

Another useful framework is Bloom's Taxonomy (Bloom, 1956). It comprises three skill domains:

- **Cognitive**: knowledge and cognitive skills
- **Affective**: attitudes, feelings/emotions, and interpersonal skills
- **Psychomotor**: sensory perception and fine and gross motor skills

Some learning objectives may involve a single domain. For example, the learning objective of "Recognize that the Valsalva maneuver intensifies murmur associated with hypertrophic obstructive cardiomyopathy" is primarily cognitive. Other learning objectives may involve all three. For example, the learning objective of "Proper placement of the stethoscope requires knowing where to put it and why" covers knowing the placement (cognitive), asking permission from the patient (affective), and then actually placing it (psychomotor).

Exercise: Objectives and Bloom's Taxonomy

For each learning objective, indicate the domain in which it is __most__ associated.

Learning Objective	Cognitive	Affective	Psychomotor
List three causes of aortic stenosis.			
Offer to and properly drape a female patient prior to auscultation.			
Correctly identify common systolic murmurs on an audio recording.			
Explain the maneuver for differentiating aortic stenosis from hypertrophic obstructive cardiomyopathy.			
Perform the maneuver to differentiate aortic stenosis from hypertrophic obstructive cardiomyopathy.			
Position the patient and transducer, and perform an echocardiogram to visualize the aortic valve in the short axis view.			

_____

Exercise: Learning Objectives and Bloom's Taxonomy—Cognitive Domain

Because so much of medicine—and so much of assessment—involves the cognitive domain, let's explore that further. Bloom proposed six levels within the cognitive domain, where each successive level implies greater depth of learning. These levels are Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. These levels were later updated using new terms and a slightly revised order (Anderson, Krathwohl, Airasian, Cruikshank, Mayer, Pintrich, Raths, & Wittrock, 2000); these are Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating. For most assessment purposes, the lowest two levels (Knowledge/Comprehension or Remembering/Understanding) can be grouped together as a general "Recognition and Recall" set of tasks, and the remaining four can be grouped together as "Application and Critical Thinking."

For each learning objective, indicate the group in which you think it is **<u>most</u>** associated.

| Learning Objective	Recognition and Recall	Application and Critical Thinking
List three causes of aortic stenosis.		
Correctly identify common systolic murmurs on an audio recording.		
Select the most cost-effective method to evaluate a patient with a suspected heart murmur.		
Recognize that the P wave represents atrial depolarization.		
List the five stages of the cardiac cycle.		
Explain the position of the AV and semilunar valves at each stage of the cardiac cycle, the status of the ventricles and atria, the pressure differences, and the sounds heard.		

Assessment Methods

So far in this lesson, the focus has been on the **purpose** of assessment and identifying what skills you want to assess. However, there are many different methods of assessment, for example:

- o MCQs
- o Short Answer
- o Essay Exam
- o Objectives Structured Clinical Examination (OSCE)
- o Standardized Patient
- o Mini-Clinical Evaluation Exercise (Min-CEX)
- o Oral Exam
- o Case-Based Discussion
- o Chart Stimulated Recall
- o Portfolios
- o Chart Audit
- o Simulators (Harvey, SimMan)
- o Patient Surveys
- o 360-degree Feedback
- o Direct Observation

The following graphic shows some of these methods ordered on the continua of fidelity (how similar the assessment format is to the actual work-related task) and reliability (how replicable the test scores are across assessments).

Figure 2. Continuum of Some Assessment Methods.

Continuum of Some Assessment Methods
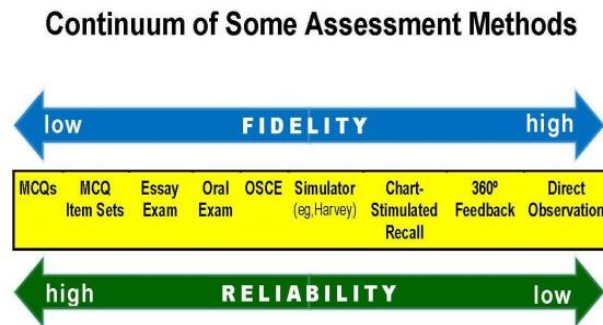
There are a number of differences between the assessment methods as we move along the continuum:

1. The methods on the left better assess **knowing**, while the methods on the right better assess **doing**.
2. The methods toward the left tend to be quick and efficient to develop and score, while the methods toward the right tend to be more costly and/or time-consuming.
3. The methods toward the right are more realistic, because they reproduce the natural environment with greater **fidelity** and are more likely to elicit the clinical skills of most interest for assessment. Unfortunately, this fidelity comes with a trade-off, as these also tend to produce **less reliable** scores.

The continuum of assessment methods can also be grouped into three broad categories:

- **Written exams**: for example, MCQs, essays
- **Simulations**: for example, standardized patients, OSCEs, and oral exams (these methods simulate the cognitive, affective, or psychomotor skills required in the practice setting)
- **Workplace assessments**: for example, 360-degree feedback (evaluations that take place during actual practice)

Putting It All Together

When developing an assessment method for a given skill or set of knowledge, the purpose of the test and the type of skills to be assessed are the two most important factors. The method should follow naturally from these aspects of the assessment. The following table summarizes some topics from this lesson and provides a guide for selecting an assessment method appropriate to a particular skill. It identifies the skills that can be assessed by several different methods and lists the strengths and limitations of each method.

Table 1. Assessment Formats: Matching Method with Skill Domain.

	Assessment Method*	Skill Domain	Advantages	Disadvantages
Written Tests	Multiple-Choice Questions (MCQs) • Standard MCQs; true-false; extended matching • Clinical vignettes; case studies; item sets; SJT – situation judgment test • Other "selected response" formats	Cognitive: Yes. Recognition of concepts; application of knowledge and critical thinking; cognitive procedural skills (e.g., computation skill). Affective: Partial. Knowledge of how to respond or appropriate steps to take. Psychomotor: No. Can only assess knowledge of procedures and related skills.	Ease of development; cost-effective Objective scoring; highly reliable scores Broad content coverage; efficient; test many topics in little time	Limited to cognitive domain Recognition and memory emphasized Encourage "MCQ study habits" Irrelevant skills (e.g., reading speed, testwiseness)
	Short Answer / Essay • Stand-alone questions and clinical vignettes; case studies and item sets • Computerized and other low fidelity)text-based) simulations	Same as MCQs, plus: depth of knowledge; critical thinking; organization and writing skills.	Assesses recall (not just recognition); gets at critical thinking skills Moderate reliability if detailed scoring rubrics or multiple items/graders are used	Subjective scoring; lower reliability Limited content coverage in a given time period; case specificity Time and effort to develop scoring rubrics and grade exams Irrelevant skills (e.g. writing ability)
Simulations	Oral exam Case-Based Discussion: Chart-Stimulated Recall OSCE; Standard Patients (SPs) Medical Patient Simulators Mini-CEX (mini clinical evaluation exercise) MMI (multiple mini interview)	Cognitive: Yes. Depending on the method and focus of assessment, can get at critical thinking (clinical reasoning). Affective: Yes. Including communication and teamwork skills, depending on assessment materials. Psychomotor: Yes. Skills assessed depend on assessment materials.	Relevance to practice Can adapt questions according to examinee performance Moderate reliability if detailed scoring rubrics or multiple sessions/graders are used	Subjective grading; limited reliability Limited content coverage in a given time period; case specificity Rater training usually required Irrelevant skills (e.g. personality) Assesses best, not typical, performance; can "fake good"
Workplace	Chart Review, Practice Audits 360° (Multisource) Feedback Patient Surveys Direct Observation of Clinical Skills • DOPS – direct observation of procedural skills • CWS – clinical work sampling • Other	Cognitive: Limited. May not have context needed to ask probing questions. Affective: Yes. Skills assessed depend on assessment materials. Psychomotor: Yes. Skills assessed depend on assessment materials.	Practice-related Relatively efficient and low cost Rating forms easy to develop Best for assessing routine or typical performance. Comprehensive if multiple observations obtained	Subjective ratings; low reliability Limited assessment of cognitive skills. Irrelevant skills (e.g. personality, other) may influence ratings Multiple raters; training required

*The arrangement of methods is somewhat arbitrary, and some methods could fall under both Simulations and Workplace. The general idea is that from top to bottom the assessment methods tend to increase in realism and naturalistic context, while decreasing in standardization and reliability.

_____

References and Resources

ACGME Outcomes Project & ABMS. (2000). *Toolbox of Assessment Methods*. Chicago, IL: Accreditation Council for Graduate Medical Education and American Board of Medical Specialties.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., ... & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, abridged edition.* White Plains, NY: Longman.

Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge, England: Cambridge University Press.

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals: Cognitive Domain*. Longman.

Boulet, J., & Raymond, M. R. (2013). *Blueprinting: Planning Your Tests*. FAIMER-Keele Master's in Health Professions Education: Accreditation and Assessment. Lesson 1, Unit 2. London, England: FAIMER Centre for Distance Learning (CenMEDIC).

Downing, S. M., & Haladyna, T. M. (2006). *Handbook of Test Development.* New York, NY: Routledge.

Downing, S. M., & Yudkowsky, R. (2009). *Assessment in Health Professions Education*. New York, NY: Routledge.

McLeod, S. A. (2009). Jean Piaget's Theory of Cognitive Development. *Simply Psychology.* Retrieved January 2019 from *http://www.simplypsychology.org/piaget.html*.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine*, *65*(9), S63–7.

Raymond, M. R., & Neustel, S. N. (2006). Determining the content of credentialing examinations. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of Test Development* (pp. 181–223). Mahwah, NJ: Lawrence Erlbaum Associates.

Rider, E. A., Nawotniak, R. H., & Smith, G. (2007). *A Practical Guide to Assessing the ACGME Core Competencies*. Marblehead, MA: HCPro, Inc.