

What does Assessment Look Like in Evidence-based Teaching?

ADRIENNE SALENTINY, PH.D.

RICHARD N. VAN ECK, PH.D.

This Session Will Cover...

Types of assessment (and when they work best)

How assessments are aligned with outcomes

How to score “rich” assessments

- Challenges and value

Creation of unbiased, logical rubrics and checklists

What Are the Goals of Assessment?

Assessment is used to evaluate the knowledge, skills, and/or attitudes of learners.

We assess to:

- Ensure that outcomes are being met
- Appraise students' ability to use knowledge outside of the classroom

Healthcare Education Triple Aim

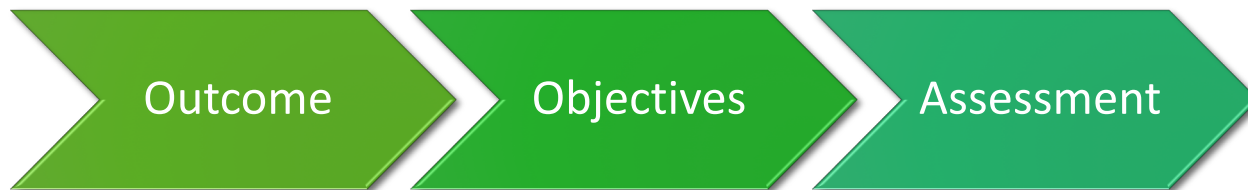
The learner's lament

- They always teach/test things they never test/teach

The educator's lament

- Part I:
 - It doesn't matter how much I talk about ____: they just don't get it!
- Part II:
 - They pass the tests but they can't do the work!

Solution



Assessment Principles

Practice ≠ assessment

Process vs. ‘outcome’ objectives

- Scientific method vs. experimental results

The role of *performance* vs. “knowing”

Authentic assessment

How Would You Measure It?

Verbal Information: Given a simulated EpiPen, the learner will be able to state which component(s) of the EpiPen must be removed before use with 100% accuracy.

Defined Concept: Given examples of simulated patients who are choking or not choking, the learner will be able to classify patients as choking or not choking by labeling or selecting them with 100% accuracy.

Problem Solving: Given a simulated patient, the learner will be able to generate a diagnosis and treatment plan for anaphylaxis with 100% accuracy.

Motor Skill: Given a real or simulated EpiPen and a simulated patient, the learner will be able to execute the use of the EpiPen by using the EpiPen with 100% accuracy.

How Would You Score It?

Type of K, S, or A	Example	Appropriate Assessment	How To Score
Verbal Information	Declarative knowledge that can be recalled later	MCQ, Matching, T/F, Fill-in-the-blank Short Answer	
Intellectual Skill	Problem-solving, Rules, and definitions	MCQ, Matching, T/F, Fill-in-the-blank Project, Essay, Case, Short Answer	
Attitude	Preferences that govern an individual's choice behavior	Indirect Project, Essay, Case, Journaling Likert-Type Scales Direct Observation/Performance	
Motor Skill	Perform a physical action	Direct Observation/Performance	

How Would You Score It?

Type of K, S, or A	Example	Appropriate Assessment	How To Score
Verbal Information	Declarative knowledge that can be recalled later	MCQ, Matching, T/F, Fill-in-the-blank Short Answer	Answer Key Rubric
Intellectual Skill	Problem-solving, Rules, and definitions	MCQ, Matching, T/F, Fill-in-the-blank Project, Essay, Case, Short Answer	Answer Key Rubric
Attitude	Preferences that govern an individual's choice behavior	Indirect Project, Essay, Case, Journaling Likert-Type Scales Direct Observation/Performance	Rubric Answer Key Checklist
Motor Skill	Perform a physical action	Direct Observation/Performance	Checklist

Rich Assessment Methods

VALUES

- Demonstrate students' ability to perform, synthesize, problem solve, think critically, and more
- Expose students to complex, sometimes interdisciplinary problems
- Increase transfer of KSAs (from class to career)

CHALLENGES

- Can't auto-grade using of electronic tools or scantron
- Difficult to grade objectively
- Difficult to grade consistently across multiple instructors/TAs
- Usually inefficient to grade, especially in a large class

Addressing the Challenges of Rich Assessment Methods

- **Challenge 1: Auto-grading.**

- This is still a challenge, but a rubric or checklist streamlines the process to ensure that the same points are addressed in each paper.
- Peer- and self- assessment may also lighten this load.

- **Challenge 2: Fair grading.**

- Rubrics lay out constructive, non-value-driven guidelines at each predetermined performance level.

- **Challenge 3: Multiple graders.**

- As with fair grading, rubrics lay out the guidelines clearly.
- If a project contains X, and X addresses A, B, and C as described by the rubric, it receives a predetermined score. Rubric design is key.

- **Challenge 4: Cumbersome in a large class.**

- This is still a challenge, but a rubric or checklist delivered to students before work begins will communicate expectations.
- Expectations of peer- or self-assessment also decrease instructor time commitment.

Group Discussion

WHAT KINDS OF ASSESSMENT CHALLENGES DO YOU HAVE?

At-a-Glance

Rubrics and Checklists:

- Cut down on subjectivity and inconsistency in grading
- Provide a framework the grading process of complex assessments or experiences
- Allow for peer & self-assessment
- Help communicate final expected performance to students
- Can be used for formative feedback between drafts
- Are sometimes built into LMS tools (Blackboard, etc.).

Design: Rubric or Checklist?

- Use a rubric if:
 - Students are graded on multiple criteria, or
 - Students are graded by their level of achievement
- Use a checklist if:
 - Students must either achieve, or not achieve (no gray area)
 - You are evaluating a performance/sequence of steps
 - You are observing a skillset and may not have time to consider elements of a table during the timeframe
- Sometimes, you may need a combination rubric/checklist

Rubric Design: Layout

Rubrics generally appear in a table format.

Consider:

- What performance criteria will you measure? (Look at syllabus/assignment descriptions)
 - Assignment components
 - Logic & flow
 - Inclusion of number and type of sources
 - Spelling, grammar, length, formatting
- How many levels of achievement will you have?
- How will you mark students' scores? (Check, circle, etc.)
- Will you want the ability to write comments? (We recommend it!)

Rubric Design: For Example...

When you want to assess the *quality* of a performance or product against *a set of criteria*:

- Place criteria in the first column
- Place quality levels in the first row
- Include descriptors of criteria at different quality levels in each cell

Ideal for authentic assessment

Rubric Example

	1	2	3	4	5
Coordination					
Briefing	No briefing or missing most components (introduction, use of names, discussion of monitoring or cross monitoring)		Some introduction and/or occasional use of names; discussion of monitoring/cross monitoring minimal or weak		Briefing includes introduction and names; next steps and monitoring/cross monitoring explicitly addressed
Verbalize Plan	Plan is not routinely verbalized		Plan, vision, goals verbalized in most cases		Plan, vision, goals clearly articulated at all relevant times
Timeframe	Timeframe is not routinely verbalized		Timeframe is verbalized in most cases		Timeframe clearly articulated at all relevant times
Debriefing	Debriefing missing or significantly flawed; most components missing or weak (what went well/not well, what was learned)		Debriefing incomplete; may be missing relevant info; all component present with some weakness (what went well/not well, what was learned)		Includes what went well, what could be done differently, what was learned with no weakness
Team Roles	Does not Discuss		Discusses; Roles may be vague or incomplete or some roles missing		Discusses roles comprehensively, inclusively, and without error
Patient Inclusion	Patient is not included		Patient is involved; may not be included in all relevant decisions or discussions		Patient is a full member of discussions and decisions where relevant
Resource Use	Relevant resources routinely ignored and/or irrelevant resources used		Uses some resources; may miss some or select non-relevant		Uses relevant resources appropriately in all circumstances
Workload Management	Little to no delegation; allows some to do more work than others without cause		Some delegation; may be unbalanced workload at times		Delegates effectively; ensures workload is balanced across team
Patient Safety	Frequent or significant actions that present potential risk to patient		Majority of actions ensure patient safety; minor errors may be present		Patient safety guides all actions; no anticipatable unsafe actions
Cooperation					
Request external resources	Team members do not ask for help from outside team on more than one occasion when they should have		Sometimes asks for help from outside team when warranted; may be delayed or absent in some instances		Asks for help outside team when warranted (consults, equipment, other clinicians, rooms)
Ask for help from team	Members do not ask for help when needed and/or take too long to do so		Members occasionally ask for help if needed; may be some delay in requests		Members routinely ask for help if needed in a timely manner
Verbally request team input	Members rarely ask for suggestions, opinions, comments or ideas		Members occasionally ask for suggestions, opinions, comments or ideas		Members routinely ask for suggestions, opinions, comments or ideas
Cross Monitoring	Few concerns stated, acknowledged, or responded to; little to no workload management		Some concerns stated, acknowledged, responded to; informal verbal updates; workload management minor or missing		Team state concerns and acknowledge/respond to stated concerns; make adjustments to manage workload; routine informal verbal updates
Verbal Assertion	Rarely verbalizes when members are unclear, uncomfortable, have concerns about actions; failure to persist when concerns are not acknowledged or responses are unsatisfactory; escalations missing or significantly delayed		Some verbalizations when members are unclear, uncomfortable, have concerns about actions; may delay or fail to persist in some cases if concerns are not acknowledged or responses are unsatisfactory; escalates when needed		Routine verbalizations when members are unclear, uncomfortable, have concerns about actions, see something unsafe; team persists if concerns are not acknowledged or responses are unsatisfactory; escalates when needed
Receptive to assertion and ideas	Routine failure to acknowledge ideas /concerns and/or failure to be respectful in responses		Team members respectfully acknowledge/are receptive to concerns/ideas of others in most cases		Team members respectfully acknowledge/are receptive to concerns/ideas

Rubric Example

	N/A	1	2	3	4
<i>Problem Identification</i> (problem is appropriate for HPT process)					
<i>Articulation of HPT Model and Approach</i> (model is described; model is based on existing models in literature; origins of model are clearly articulated; combination of approaches is explained; reference to model used to organize documentation of process in headings, narrative, etc.)					
<i>Systems Approach</i> (analysis, intervention, and evaluation consider full range of system)					
<i>Logic</i> (each stage or step in process is logically connected; data/outputs are used as inputs where appropriate; analysis drives design)					
<i>Sources of Evidence</i> (data for analysis, intervention, and evaluation are drawn from full range of appropriate sources, including extant data, qualitative data, raw data outputs from processes such as interventions or on-the-job analyses)					
<i>Analysis</i> (analysis is appropriate for problem identification, solutions, and intervention; demonstrates iterative nature of ongoing analysis at different stages of HPT process as appropriate)					
<i>Intervention</i> (range of possible interventions identified; evidence of HPT constructs, processes, and data used for selection of intervention(s); evidence of data collection/analysis used during intervention)					
<i>Evaluation</i> (results of implementation clearly articulated and informed by HPT processes; data and analysis from implementation and other relevant stages used to support conclusions; evidence of <i>judgment</i> is present in addition to analysis of findings)*					
<i>ROI/Cost-Benefit</i> (costs and benefits are discussed; ROI calculations are conducted; approach is appropriate for measurement of this construct; basis for costs and calculations are complete and reasonable)					
<i>Rationale for Current State of HPT Process/Project</i> (description of project is clear; state of implementation is articulated and supported by evidence; stage of project, if incomplete, is appropriate to scope, complexity, or other aspects of the project)					
<i>Writing</i> (appropriate for graduate/professional levels; voice is professional; ideas clearly articulated, developed, and supported appropriately; flow of logic and ideas is clear)					
<i>Editing</i> (typos, grammar, and spelling errors)					
<i>Layout</i> (headings consistent; APA where appropriate; tables and figures clear and appropriate; margins & white space; space between paragraphs)					
OVERALL DOCUMENTATION					
Comments:					

4 = Exceptional	Project documentation clearly demonstrates evidence of exceptional comprehension and application of understanding of the HPT construct(s). Employs appropriate relevant information from the HPT literature in support of analysis and conclusions. Product and design are reflective of extensive personal skills that are relevant, accurate, and consistent with the domain.
3 = Good	Project documentation demonstrates evidence of minimal competence of comprehension and application of the HPT construct(s) with no significant gaps in scope or depth of content and concepts. Employs adequate supporting information from HPT literature. Product and design are reflective of personal skills that are adequate and consistent with the domain.
2 = Fair	Project documentation demonstrates only a very <i>general</i> understanding of HPT construct(s), with notable gaps in content and concepts, depth within content and concepts, or both. Employs only limited supporting information from HPT literature. Product and design are reflective of limited personal skills consistent with the domain.
1 = Poor	Demonstrates little or no evidence of understanding design constructs or domain. Significant gaps in scope, content, or both. Fails to support analysis or conclusions with information from HPT literature. Product and design are not reflective of personal skills needed to be consistent with the domain.
NA	This construct/criteria is not relevant for this project.

Rubric Design: Scoring

Score based on what is important to the course, field, and intended outcome.

- Will it be scored with...
 - Points?
 - Check/plus/minus?
 - A different rating scale?
- What is the assignment worth (per individual component and overall)?
 - Is each section worth the same points, or should they be weighted?
 - Can (and should) a student achieve the desired level of mastery if they receive a low score on one or more sections?

Rubric Design: Wording

DO

Use unbiased, constructive language to communicate expectations, points for improvement, and predetermined benchmarks of achievement.

Eight or more primary sources were consulted to produce a paragraph that is structured by topic, not serially.

DON'T

Use subjective language or undefined terms to pass judgment on whether a student has achieved the “right answer.”

Good use of literature from the field, organized in a logical way; provides a significant contribution to the study.

More About Checklists

Breaks performance down into component parts

Often sequential, but not always

Behavior checklist

- Measures IF the behavior occurred

Behavior frequency checklist

- Measures how OFTEN the behavior occurred

Behavioral Checklist Examples

Customer Service Checklist	Yes	No	NA
Greets Customer			
Takes Order			
Offers Additional Products			
Takes Payment			
Provides Change			
Submits Order Ticket			
Delivers Order			

TABLE 4

Summative Assessment, Combination Checklist and Ratings Scale for Assessing Healthy Upper String Technique

Behavior	Not Observed	Observed
Left-hand contact point		
Effective posture		
Relaxed thumb		
All fingers down—"boxed" first finger		
Knuckle angle toward scroll		
Clean string crossing (plays one string only)		
Plays low 2 notes (C or F)		
Interval between low 2 and 3 is a major second		
Bow does not travel in the bout		
Tone is balanced on all four strings		
Relaxed bow hold		
Left wrist gently curved		
Straight bows		
Big bows		
Adjusting intonation		
Consistent tempo maintained		
Began on correct bow direction		
Slurs executed correctly		
Equal amount of bow used in slurs		
Interval accuracy		
Shifted on correct finger		
Entire hand moves as one		
Pitch accuracy on shift		
Third-position finger pattern accuracy		
Shift in time		
Total points		/25

Behavioral Checklist Examples

Interval Recording Form

Student's Name _____

Target Behavior: Rodney taps his feet, flaps his hands, or engages in other self stimulatory behavior.

Observation Length: 1 hour

Interval Length: 5 minutes

[illegible]

Summary and Closing Thoughts

“Too often, rubrics suggest only that poor work has “less” of the same types of qualities as better work.” (O’Neil, 1994, p. 5)

Rich, complex assessments such as cases, projects, research papers, and demonstrations effectively measure mastery of learning outcomes, but there are barriers to grading them efficiently and objectively.

- Rubrics and checklists provide a framework for the fair and consistent grading of these assessments.
- They can be as complex or simple as needed, and can operate within a variety of scoring strategies.
- They should use consistent, non-judging, language that indicates tangible benchmarks (not opinions or “good-/poor-”ness).

Distribute the rubric or checklist to students and encourage (or require) their use of this tool for formative feedback during the process and/or peer-evaluation prior to submission for grading.

References

Angelo, T. A. & Cross, K.P. (1993). *Classroom assessment techniques: a handbook for college teachers* (2nd ed.). San Francisco: Jossey-Bass.

Gagne, R. M., Wager, W. W., Golas, K. C., Keller, J. M. (2005). *Principles of instructional design* (5th ed.). Belmont, CA: Wadsworth/Thomson Learning.

Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. Needham Heights, MA: Allyn & Bacon.

McMillan, J. H. (1997). *Classroom assessment: Principles and practices for effective instruction* (7th ed.). Needham Heights, MA: Allyn & Bacon.

Pellegrino, K., Conway, C. M., & Russell, J. A. (2015). Assessment in performance-based secondary music classes. *Music Educators Journal*, pp. 48–55.