

Assessment Literacy and Student Learning Objectives

Monday, December 8, 2014

Objectives

- To provide an overview of the importance of Assessment Literacy as it relates to Student Learning Objectives.
- Provide access to resources to further the discussions.

You Won't be an Expert

At the conclusion of this session,

- You may recognize additional professional development is needed.
- You will know about SLOs and why assessment literacy is important in the process.
- But, you won't be an assessment literacy expert.

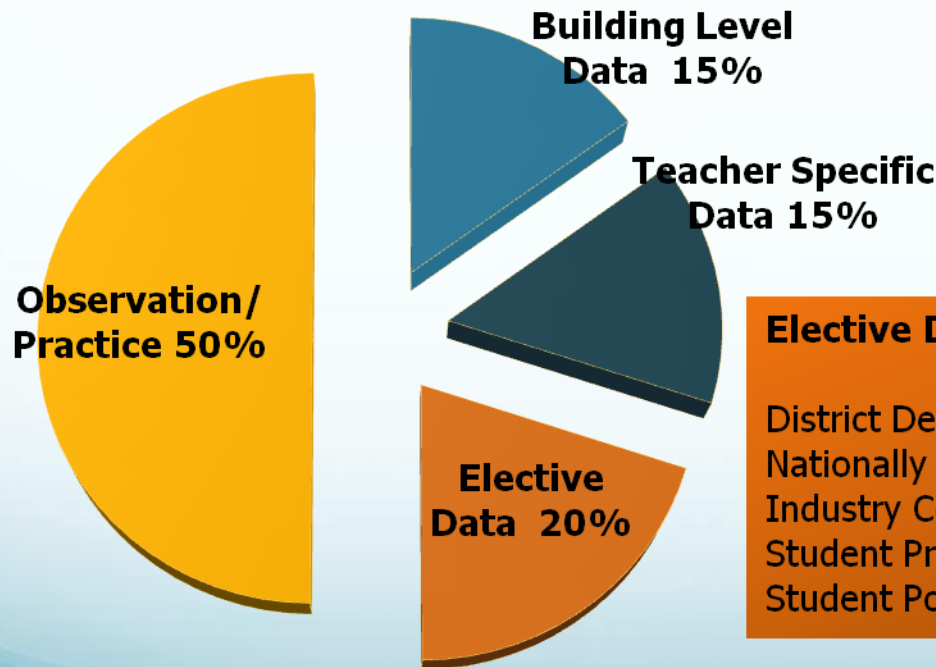
Teacher Effectiveness System in Act 82 of 2012

Teacher Observation/ Practice

Planning and Preparation
Classroom Environment
Instruction
Professional Responsibilities

Building Level Data/School Performance Profile

Indicators of Academic Achievement
Indicators of Closing the Achievement Gap, All Students
Indicators of Closing the Achievement Gap, Historically Underperforming Students
Indicators of Academic Growth/ PVAAS
Extra Credit for Advanced Achievement



Teacher Specific Data

Student Performance on Assessments
PVAAS 3-Year Rolling Average
IEP Goals Progress*
LEA Developed Rubrics*

Elective Data*

District Designed Measures and Examinations
Nationally Recognized Standardized Tests
Industry Certification Examinations
Student Projects Pursuant to Local Requirements
Student Portfolios Pursuant to Local Requirements

*Student Learning Objective Process



Do I know if these assessments measure these standards?

STUDENT LEARNING OBJECTIVE (SLO) PROCESS TEMPLATE					
SLO is a process to document a measure of educator effectiveness based on student achievement of content standards. SLOs are a part of Pennsylvania's multiple-measure, comprehensive system of Educator Effectiveness authorized by Act 82 (HB 1901).					
1. Classroom Context					
1a. Name		1b. School		1c. District	
1d. Class/ Course Title		1e. Grade Level		1f. Total # of Students	
1g. Typical Class Size		1h. Class Frequency		1i. Typical Class Duration	
2. SLO Goal					
2a. Goal Statement					
2b. PA Standards					
2c. Rationale					
3. Performance Measures (PM)					
3a. Name	PM #1: PM #2: PM #3: PM #4: PM #5:	3b. Type	<input type="checkbox"/> District-designed Measures and Examinations <input type="checkbox"/> Nationally Recognized Standardized Tests <input type="checkbox"/> Industry Certification Examinations <input type="checkbox"/> Student Projects <input type="checkbox"/> Student Portfolios <input type="checkbox"/> Other: _____		
3c. Purpose	PM #1: PM #2: PM #3: PM #4: PM #5:	3d. Metric	<input type="checkbox"/> Growth (change in student performance across two or more points in time) <input type="checkbox"/> Mastery (attainment of a defined level of achievement) <input type="checkbox"/> Growth and Mastery		
3e. Administration Frequency	PM #1: PM #2: PM #3: PM #4: PM #5:	3f. Adaptations/ Accommodations	<input type="checkbox"/> IEP <input type="checkbox"/> ELL	<input type="checkbox"/> Gifted IEP <input type="checkbox"/> Other	
3g. Resources/ Equipment	PM #1: PM #2: PM #3: PM #4: PM #5:	3h. Scoring Tools	PM #1: PM #2: PM #3: PM #4: PM #5:		
3i. Administration & Scoring Personnel	PM #1: PM #2: PM #3: PM #4: PM #5:	3j. Performance Reporting	PM #1: PM #2: PM #3: PM #4: PM #5:		



4. Performance Indicators (PI)				
4a. PI Targets: All Student Group	PI Target #1:			
	PI Target #2:			
4b. PI Targets: Focused Student Group (optional)	PI Target #3:			
	PI Target #4:			
	PI Target #5:			
4c. PI Linked (optional)		4d. PI Weighting (optional)	PI	Weight
			#1	
			#2	
			#3	
			#4	
			#5	

5. Elective Rating				
5a. Level	<i>Failing</i> 0% to ___% of students will meet the PI targets.	<i>Needs Improvement</i> ___% to ___% of students will meet the PI targets.	<i>Proficient</i> ___% to ___% of students will meet the PI targets.	<i>Distinguished</i> ___% to 100% of students will meet the PI targets.

Teacher Signature _____ Date _____ Evaluator Signature _____ Date _____

5b. Rating	<input type="checkbox"/> Distinguished (3)	<u>Notes/Explanation</u>
	<input type="checkbox"/> Proficient (2)	
	<input type="checkbox"/> Needs Improvement (1)	
	<input type="checkbox"/> Failing (0)	

Teacher Signature _____ Date _____ Evaluator Signature _____ Date _____



Are you confident that the elective rating reflects strong alignment between the chosen standards and the performance measures?
Is the rating a true measure of a teacher's effectiveness?

Why Be Assessment Literate?

- ED Test and Measurement 101 was a long time ago
- Assuring students are being accurately assessed
- Making sure we are assessing standards as they are intended, not as they are interpreted
- Helping teachers develop better assessments
- Specifically related to SLOs
 - Helping staff identify appropriate Performance Measures
 - Having conversations with your staff during the SLO approval phase

Characteristics of Quality Assessments

PM #4

Quality Assessments

- Measure what you intend them to measure
- Are aligned to the rigor of the intended standards being assessed
- Follow good test development protocols
- Weight important topics more heavily
- Are written at the appropriate grade level
- Sequenced from easiest to hardest
- Consistent with instructional strategies used

Quality Assessments (cont.)

- Are easily read because the following was considered:
 - Spacing of questions
 - Amount of text
 - Length of sentences
 - Text style
- Directions are clear and concise
- Visuals may be used

Consider reading the article provided in your handouts:

Creating and Grading Valid and Accessible Teacher-Made Tests

What Impacts Assessment?

- Reliability
 - The extent to which assessment results are consistent
 - Imagine a scale...
 - 5 lbs. of potatoes should weigh 5 lbs. every time they are placed on the scale; regardless of the time of day, no matter how many times you weigh the potatoes.
- Validity
 - The accuracy of the assessment results; do they measure what they are supposed to measure?
 - Think of that scale...
 - If the scale says you weigh 130 lbs. but you actually weigh 145 lbs., the measurement is not valid.

Reliability

- Reliability is a Correlation
 - Test - Retest
 - Original - Alternate Form
 - First Half – Second Half
- 0 would mean no reliability
- 1.0 is a perfect reliability
- .80 considered very good reliability
- Less than .50 would be an unreliable test

How Reliable are these Tests?

- Science Test
 - .93
- Physical Fitness Test
 - .66
- Writing Assessment (Scored with a Rubric)
 - .75
- Social Studies Final
 - .48

Important Notes about Reliability

- It is unrealistic to think one would conduct these correlation calculations for all assessments
- But it is important to ask yourself, can I rely on these results?
- Reliable does not mean Valid
 - A scale consistently measuring one at 130 lbs. when they weigh 145 lbs. is very reliable – a perfect correlation, but the results aren't valid.
 - If students are scoring consistently well on a final exam, but the final exam is not really measuring the intended goals of the course, the test is reliable, but not valid.

Validity (1)

- Does the assessment measure what it is supposed to measure?
- Does the content of the test measure the intended instructional objectives or standards?
- Did you follow good test development protocols to ensure you are testing what you think your testing?
- Are your questions clearly understood by students?
- Do students have enough room to answer the questions?

Validity (2)

- In some cases, valid assessments can be predictive.
- The extent to which scores are in agreement with or can predict other criterion.
 - SAT scores predicting future college success
 - Third grade reading scores predicting future school success
 - PVAAS scores predicting future PSSA scores

Validity (cont.)

- Great news about Validity...
 - We know...
 - A reliable assessment isn't necessarily valid
 - BUT
 - Valid assessments are almost always reliable!
- If we focus on making sure our assessments are valid, we can be confident in their reliability!

Teachers Should Ask “Are My Assessments Valid?”

- Do my assessments test what I think they do?
- Do I have enough assessment items for each standard?
- Are my assessment items rigorous enough to assess the intent of the standard?
- If the assessment is teacher created, were strong test development protocols used?
 - The document *Creating and Grading Valid and Accessible Teacher-Made Tests*) is a good reference for this.

Validity

- For the purposes of SLOs, we'll focus on validity.
- A final exam that only included content covered during the last 6 weeks would not be a valid measure of the entire course.
- If the standard being assessed expects students to apply their knowledge, and the test measures only content, not application, then the assessment would not be a valid measure of the standard.

Pop Quiz



PM #5

- A parent called you to ask about the reliability coefficient on a recent standardized test. The coefficient reported was .89 and the parent thinks that must be a very low number. How would you explain that .89 is an acceptable coefficient?
- Your school district is looking for an assessment instrument to measure reading skills. They have narrowed the selection to two possibilities—Test A provides data indicating that it has high validity, but there is no information about its reliability. Test B provides data indicating that it has high reliability, but there is no information about its validity. Which Test would you recommend? Why?

Depth of Knowledge



Webb's Depth of Knowledge Levels

Recall and Reproduction: Level 1

Skills & Concepts: Level 2

Strategic Thinking: Level 3

Extended Thinking: Level 4

Recall and Reproduction: Level 1

- Curricular elements that fall into this category involve basic tasks that require students to recall or reproduce knowledge and/or skills.
- DOK 1 requires recall of information, such as a fact, definition, term, or performance of a simple process or procedure.
- It may also involve use of simple procedures or formulas. It can be difficult without requiring reasoning.
- A student answering a Level I item either knows the answer or does not; that is, the answer does not need to be figured out” or “solved.

DOK Level 1 Examples

- List animals that survive by eating other animals.
- Locate or recall facts explicitly found in text.
- Describe physical features of places.
- Determine the perimeter or area of rectangles given a drawing or labels.
- Identify elements of music using musical terminology.
- Identify basic rules for participating in simple games and activities.

Skills/Concepts: Level 2

DOK 2 includes the engagement of some mental processing beyond recalling or reproducing a response. Items require students to make some decisions as to how to approach the question or problem.

These actions imply more than one mental or cognitive process/step.

DOK Level 2 Examples

- Compare desert and tropical environments.
- Identify and summarize the major events, problem, solution, conflicts in literary text.
- Explain the cause-effect of historical events.
- Predict a logical outcome based on information in a reading selection.
- Explain how good work habits are important at home, school, and on the job.
- Classify plane and three dimensional figures.
- Describe various styles of music.

Strategic Thinking: Level 3

DOK 3 requires deep understanding as exhibited through planning, using evidence, and **more demanding cognitive reasoning**. The cognitive demands at Level 3 are complex and abstract.

An assessment item that has more than one possible answer and **requires students to justify the response they give** would most likely be a Level 3.

DOK Level 3 Examples

- Compare consumer actions and analyze how these actions impact the environment.
- Analyze or evaluate the effectiveness of literary elements (e.g. characterization, setting, point of view, conflict and resolution, plot structures).
- Solve a multiple-step problem and provide support with a mathematical explanation that justifies the answer.
- Sort objects into a variety of categories and explain the rule you used.

Extended Thinking: Level 4

- DOK 4 requires high cognitive demand and is **very complex**. Students are expected to make connections—relate ideas *within* the content or *among* content areas—and have to select or devise one approach among many alternatives on how the situation can be solved.
- Students are engaged in conducting multi-faceted investigations to solve real-world problems with unpredictable solutions.
- Due to the complexity of cognitive demand, DOK 4 often requires an extended period of time.

DOK Level 4 Examples

- Research a wetlands animal using internet resources and create a realistic fiction text with the wetlands animal as the main character.
- plan and curate the exhibits for a museum of the ancient river civilizations of Mesopotamia, Egypt, Indus Valley and China.
- Conduct a river study by collecting data over time, taking into consideration a number of variables, and analyze the results.

DOK is about Complexity

- Level 1 requires students to use simple skills or abilities.
- Level 2 includes the engagement of some mental processing beyond recalling.
- Level 3 requires some higher level mental processing like reasoning, planning, and using evidence.
- Level 4 requires complex reasoning, planning, developing, and thinking over an extended period of time



DOK is about Complexity (cont.)

- The intended student learning outcome determines the DOK level. What mental processing must occur?
- While verbs may appear to point to a DOK level, it is what comes after the verb that is the best indicator of the rigor/DOK level.
 - **Describe** the process of photosynthesis.
 - **Describe** how the two political parties are alike and different.
 - **Describe** the most significant effect of WWII on the nations of Europe.

Resource to Consider

- Hess' Cognitive Rigor Matrix
- Take a few minutes to look at the Matrix
- This may be a tool to use with staff when reflecting on rigor in our assessments



Blueprinting Assessments

Blueprinting



PM #7-10

- On your own,
 - Refer to the blueprinting documents
 - When would you use these?
 - With whom would you use them?
 - Why would you use them?
 - What would be their purpose?
- Turn and share your thoughts with an elbow partner – record your answers on PG 11.

Using the Blueprints Help Us Answer Questions

- Do I have enough assessment items to fully assess a standard?
- Are the standards being assessed with the same value/weight?
 - Do they need to be?
- What needs to be altered, if anything?

Scoring Tools

Answer Keys and/or Rubrics

Developing Rubrics

- Two Types
 - Holistic Rubrics
 - Analytic Rubrics



Rubrics – Keep In Mind

- Holistic Scoring
 - Provides a single score based on an overall determination of student performance
 - Assesses a student's response as a whole for the overall quality
 - Very difficult to calibrate
 - Often multiple evaluators will choose different scores

Rubrics – Keep In Mind

- Analytic Scoring
 - Identifies and assesses specific aspects of a response
 - Multiple Dimensions or Categories are assessed
 - Provides a logical combination of sub-scores to the overall assigned score.
 - Much easier to score, less subjective

Rubrics – Keep In Mind

- Considerations:
 - If it's important enough to score, it needs to be in the rubric.
 - Rubric needs to align to the standards being assessed.
 - Avoid words like: many, some, few –instead give a numeric description.
 - Avoid subjective words: creativity or effort, excellent or inadequate

Principles of Well-Designed Assessments

- Be built to achieve the designed purpose
- Produce results that are used for the intended purpose
- Align to targeted content standards
- Contain a balance between depth and breadth of targeted content
- Be rigorous, and fair
- Be sensitive to testing time and objectivity

Basic Assessment Literacy Training

- Available at your Local Intermediate Unit
- Would include deeper dives in:
 - Validity
 - Test development protocols
 - Webb's Depth of Knowledge
 - Test Blueprinting
 - Rubrics

For Technical Assistance

- Contact your Local Intermediate Unit
- Facilitators of these sessions:
 - Cori Cotner: ccotner@iu17.org
 - Karen Ruddle: kruddle@caiu.org
 - Cristine Wagner-Deitch: clw@bviu.org
- SLO State Lead, Dave Deitz: oddeitz@comcast.net