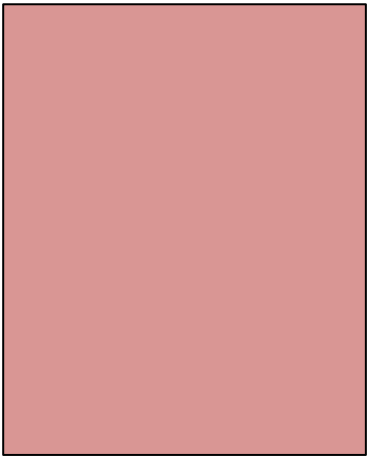


Student Learning Objectives (SLOs)



Guidance on Selecting Assessments

2013



Guidance on Selecting Assessments for SLOs

Selecting and approving assessments can be one of the most challenging and important steps of the SLO process. These measures enable teachers and teacher teams to determine growth toward and attainment of the SLO. *Before selecting assessments, local education agencies (LEAs) should note that the Ohio Department of Education (ODE) strongly recommends that districts do not allow assessments created by one teacher for use in his or her classroom for an SLO.* In the rare case where a teacher must create an assessment that is unique to his or her classroom, *ODE strongly recommends that the teacher develop the assessment in consultation with a school or district administrator with expertise in assessments, a special educator, an English language learner (ELL) specialist, and/or a content team member.*

Criteria for Selecting Assessments

When selecting assessments, teachers or teacher teams need to consider two major questions. This section provides guidance around each of these questions.

1. Is the assessment aligned to both my students' learning objectives and to the appropriate grade- or content-specific standards?

The assessment should cover the key subject and grade-level content standards and curriculum that will be taught during the interval of instruction. When examining assessments for alignment, teachers and teacher teams should look for the following:

- Items on the test should cover all key subject/grade-level content standards.
- No items on the test should cover standards that the course does not address.
- Where possible, the number of test items should mirror the distribution of teaching time devoted to concepts or the curriculum focus. For example, if a foreign language teacher devotes almost equal amounts of time to developing students' reading comprehension, listening comprehension, oral communication, and written communication skills, he or she should not use a test that devotes 90 percent of the test to reading comprehension. Instead, the distribution of the test should mirror instruction, meaning that about a quarter of the test should focus on each of the four skills listed above.

Resources for Creating and Evaluating Assessments

The Standards for Educational and Psychological Testing developed jointly by the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education

Educative Assessment: Designing Assessments to Inform and Improve Student Performance by Grant P. Wiggins

[Guidance for Developing and Selecting Assessments of Student Growth for Use in Teacher Evaluation Systems](#) by Joan L. Herman, Margaret Heritage, and Pete Goldschmidt

Teacher-Made Assessments: How to Connect Curriculum, Instruction, and Student Learning by Christopher R. Gareis and Leslie W. Grant

Creating Effective Classroom Assessments, a [webinar](#) presented by Stuart R. Kahl, Deborah Farrington, and Ellen Vorenkamp

- The items or tasks should match the full range of cognitive thinking required during the course. For example, if the main foci of the mathematics content standards are solving word problems and explaining reasoning, some questions or items on an assessment should require students to solve word problems and explain how they arrived at their answers.

The assessment should require students to engage in higher-order thinking where appropriate. These items or tasks may require students to use reasoning, provide evidence, make connections between subjects or topics, critique, or analyze.

Examples of assessment alignment with SLOs and the appropriate grade- or content-specific standards:

- A. An AP Biology teacher is evaluating available assessments to use for his SLO, which must align with the content of the AP course. He locates a district-created assessment geared toward first-year college students that covers both the major topics and the important skills associated with the course. The assessment includes 30 questions on organisms and populations, 14 questions on molecules and cells, and 16 questions on heredity and evolution.

The assessment sufficiently aligns with the content of the course. Because the AP course is designed to be the equivalent of a college introductory biology course, a test written to assess first-year college students would be appropriate. In addition, the distribution of questions nearly follows the AP Biology Development Committee's recommendations that teachers spend 50 percent of the time teaching organisms and populations, 25 percent teaching molecules and cells, and 25 percent studying heredity and evolution. Finally, the types of items mirror the AP Biology test, not just the content.

Examples of assessment alignment with student learning objectives and the appropriate grade- or content-specific standards:

- B. A second-grade mathematics teacher team is evaluating a commercially available end-of-course assessment. Looking at the items on the test, the team compares the foci of the test with the Common Core State Standards for Grade 2.

Foci of the Test	Common Core State Standards for Grade 2
Operations and Algebraic Thinking	Operations and Algebraic Thinking
Number and Operations in Base Ten	Number and Operations in Base Ten
Fractions	Measurement and Data
Geometry	Geometry
Statistics	
Probability	

The chart above indicates a discrepancy between the content of the test and the content of state standards. The test covers more topics than are taught with the Common Core and does not include measurement and data. An item analysis would be necessary to make a final determination, but given that the Common Core focuses on fewer topics in greater detail than may be on the test, a different assessment would be more appropriate.

- C. The fifth grade science curriculum contains three curricular units: cycles and patterns in the solar system; light, sound, and motion; and interactions within ecosystems. However, through these curricular units, students are expected to develop scientific inquiry skills in accordance with state standards. When the teacher team evaluates various available assessments to use with their SLOs, it finds that most of the assessments are multiple-choice questions that require basic recall, like “Which of the following is not a characteristic of Venus?”

The teacher team instead creates its own assessment that integrates the content of the course with scientific processes and inquiry. In addition to having a few multiple-choice questions, the assessment requires students to provide written explanations for scientific phenomena, analyze and interpret data relevant to the course content, and describe how they would construct a basic scientific investigation. The resultant assessment is rigorous and is better aligned to the expectations of the state content standards.

Examples of assessment alignment with student learning objectives and the appropriate grade- or content-specific standards:

- D. A seventh-grade social studies curriculum covers relevant world developments from 750 B.C. to 1600 A.D. A teacher examines an available district-created assessment for potential use with SLOs. The assessment features 40 questions, 20 of which are focused on Ancient Greece and 20 of which are focused on Ancient Rome.

The assessment mentioned above might be a district-created unit test that focuses on development in Ancient Greece and Ancient Rome. It does not adequately cover the breadth of the course, which covers world history up through global exploration. In order for the assessment to be aligned to the course, the assessment would have to measure student growth in understanding of key developments in a variety of cultures—not just Ancient Rome and Greece—over a larger period of time.

2. Does the assessment allow high- and low-achieving students to adequately demonstrate their knowledge? In other words, does the assessment have enough stretch?

All students should be able to demonstrate developmentally appropriate progress on the assessment(s) used with an SLO. In order for the assessment to work for most or all students, the assessment must have sufficient “stretch,” meaning that it contains questions that are of varying difficulty and covers some basic, low-level and advanced knowledge or skills. Teachers may not be able to make an informed judgment about the needed stretch of the assessment until they have analyzed the baseline or preassessment performance of students. When evaluating the assessment for sufficient stretch, teachers and teacher teams should keep their lowest performing and highest performing students in mind. Based upon students’ recent performance, will they be able to demonstrate growth on this assessment?

- All students should be able to demonstrate growth on the assessment.
- The test includes items that cover basic knowledge and skills and appropriate, content-relevant items that will challenge the highest performing students.

Examples of assessment that allow high- and low-achieving students to adequately demonstrate their knowledge:

- A. A teacher examines a district-created assessment of fourth grade reading. The assessment covers all reading standards for informational text and literature for fourth grade and often includes questions that are slightly less or more challenging than grade-level expectations. In addition, questions throughout the assessment cover the third grade and fifth grade expectations of the same standard. For example, three assessment tasks are aligned with fourth grade standards and require students to compare and contrast a firsthand and secondhand account of the same event and describe the differences in the two accounts in terms of focus and information. In addition, one question asks students to distinguish their own point of view from that of an author (a third grade expectation), and one task asks students to analyze multiple accounts of the same event or topic and note differences in points of view (a fifth grade expectation).

Given that the teacher has one student who began the year reading below grade level and three students who were reading above grade level, this assessment has sufficient stretch.

- B. A high school band teacher distributed a district-created high school I music preassessment at the beginning of the year to his high school I class. Looking at the results, the teacher was surprised to see that a third of his students scored 85 percent or higher on the preassessment. The teacher is scheduled to distribute a postassessment to students at the end of the year that contains questions of the same difficulty level.

Because so many students demonstrated mastery of course content at the beginning of the year, the high school I music posttest currently planned does not have enough stretch. In order to ensure that all students will be able to demonstrate developmentally appropriate growth, the teacher might need to supplement the high school I postassessment with more challenging questions or tasks.

3. Is the assessment valid and reliable?

The assessment should be both valid and reliable. In other words, the assessment should measure accurately what it says it measures and should produce consistent results (that is, it should be administered in such a way that students with the same skills should obtain similar scores). When evaluating assessments for validity and reliability, teachers and teacher teams should consider the following:

- Unless the assessment aims to test reading skills, a test should not include overly complex vocabulary. For example, a mathematics test that includes word problems with complex names and language may be assessing reading skills rather than mathematical reasoning.
- Items or tasks should be written clearly and concisely. Performance-based assessments should contain clear directions that are easily understood.

- Clear scoring rubrics or guidance should be included for performance-based items.
- The teacher or teacher team should determine how the assessment will be administered consistently across classes. Testing conditions, instructions, and test items (if using different forms of a test across classes) should be similar across classes.

Examples of ensuring assessments are valid and reliable:

- A. The teacher evaluates a ninth grade preassessment and postassessment in social studies. The tests are aligned with the content standards, contain sufficient stretch, and are sufficiently rigorous. However, the teacher notices that most of the questions are written at a 12th-grade reading level.

This test raises validity issues. If students do better on the postassessment, would it be because their knowledge of social studies and reasoning skills has improved, or because their reading comprehension has improved? To create a more valid assessment, the teacher might convene a teacher team to create a new test that uses appropriate vocabulary and will be readable to all students.

- B. A team of band teachers in the district create a performance assessment for students. In addition to developing the tasks together, the teachers specify a set of directions and testing conditions that each teacher will follow. For example, each student will be asked to perform a short piece of music during their small-group lessons. All teachers will assess the students using the same band rubric. Prior to grading, teachers will practice using the rubric and make sure that they are all grading performances consistently.

By creating standard assessment procedures, the teacher team is increasing the reliability of the assessment. These procedures will help ensure that one student's results are not more valid than another student's. In other words, if Susie takes the test during a teacher's period 1 class and then again during the teacher's period 6 class, her results should be similar. Again, testing conditions, instructions, and test items (if using different forms of a test across classes) should be similar across classes.

Types of Assessments

The list below ranks assessment types based upon the likelihood that the assessments will be aligned to standards, rigorous, valid, and reliable.

1. **State assessment items in proportion to the content specified in the SLO**—These assessment item banks are provided by the state and include items from past Ohio Achievement Assessments and Ohio Graduation Tests. Note: Teachers instructing a course with a Value-Added measure should focus their SLOs on other courses that do not have a growth measure in place so long as the district plan permits LEA measures in addition to Value-Added data.
2. **Commercially available assessments**—Some commercially available assessments have been carefully created and reviewed by assessment and education experts. However, these assessments do not always align with state content standards. Teachers may use these measures for their SLOs, but ODE recommends that teachers review these assessments for alignment first.
3. **District or team-created assessments**—District-created or team-created assessments are appropriate for use with SLOs, provided they meet the criteria for selecting assessments. Wherever possible, the same assessments should be administered across classrooms and across the district to increase comparability across classrooms. ODE also recommends that teams harness the expertise of district or school administrators with expertise in assessments, content area specialists, special educators and ELL specialists when developing assessments.
4. **Teacher-created assessments**—ODE strongly recommends that districts not allow assessments created by one teacher for use in his or her classroom for an SLO. In the rare case where a teacher must create an assessment that is unique to his or her classroom, ODE strongly recommends that the teacher develop the assessment in consultation with a school or district administrator with expertise in assessments, a special educator, an ELL specialist and/or a content team member. Individual teachers should not create assessments *for this purpose*.

Frequently Asked Questions Related to Assessments

1. *Is a teacher-created assessment appropriate for an SLO?*

If no other assessments are available, ODE recommends that assessments be created in teacher or district teams following this guidance and the valuable guidance in other assessment literacy resources. Whenever possible, individual teachers should not create assessments.

2. *Why is baseline data important?*

Baseline data is an integral part of the SLO process. In order to demonstrate growth, teachers have to have information about their students' starting points. Teachers should consider using the following sources of information as baseline data:

- Results from a preassessment that is similar to the end-of-year assessment.
- Results from last year's end-of-course exams in the same subject.
- Data from a portfolio of student work from the prior year in the same subject.
- Results of the first unit test and other student work samples from the beginning of the year.
- Data on performance in a similar subject, if the subject is completely new to students (i.e., the class is a first-year foreign language class or an introduction to physics).

3. *I have students with Individualized Education Plans (IEPs) and 504 plans. Am I allowed to provide them with the testing modifications articulated in their plans?*

Yes, by law students are entitled to receive the testing accommodations in their IEPs and 504 plans. All teachers should provide students with testing modifications when appropriate.

Checklist for Selecting Appropriate Assessments

This checklist should be completed prior to SLO approval to ensure that the assessment chosen meets the basic requirements.

Alignment to Standards:			
<i>Is the Learning Objective clearly reflected in the assessment measure?</i>			
Yes	Somewhat	No	
			All items in the assessment align to the standard(s) addressed in the SLO.
			The assessment measure addresses the full range of topics and skills included in the SLO.
			The focus of the assessment mirrors the focus of the curriculum and standards.
			The items or task match the full range of cognitive thinking required during the course.
			The assessment requires students to engage in higher order thinking where appropriate.
Comments:			
Stretch:			
<i>Will all students be able to demonstrate growth on this assessment?</i>			
Yes	Somewhat	No	
			The test includes items that cover prerequisite knowledge and skills from prior years and appropriate, content-relevant items that will challenge the highest performing students.
			Test items cover knowledge and skills that will be of value beyond the school year.
Comments:			
Validity and Reliability:			
<i>Is the assessment measure a valid and reliable tool for the intended purpose?</i>			
Yes	Somewhat	No	
			The assessment does not include overly complex vocabulary.
			Items or tasks are written clearly and concisely.
			Clear scoring rubrics or guidance exists for open-ended questions or performance-based assessments.
			The teacher has a plan for administering assessments consistently across classes.
Comments:			

