

REFLECTING ON ASSESSMENT DESIGN

NARRATOR'S SCRIPT

CONTENTS

Introduction and Purpose	1
Key Concepts	1
The Assessment Checklist	1
Alignment	1
Rigor	2
Precision	3
Bias	3
Scoring	4
Check for Understanding	5
Conclusion	5
Sources	6

INTRODUCTION AND PURPOSE

In this module, we focus on a tool that you can use to reflect on an assessment you have written or selected. We simply call it the "Assessment Checklist." It focuses you on five elements of assessment design, which help you design classroom assessments that have an *appropriate* level of validity and reliability for use in your classroom.

By the end of this module, you should be able to use the assessment checklist to help determine whether an assessment you design appropriately addresses the five elements of assessment design featured in this series of modules.

KEY CONCEPTS

The Assessment Checklist

You can use the "Assessment Checklist" to reflect on the design of your assessment. It collapses into a single document the concepts you need to consider when you plan, write and select assessments.

The checklist includes five sets of questions, one for each element of assessment design. In this module, we work our way through the checklist, one element and one set of questions at a time. For each element, we summarize each set of questions and use examples from other modules to illustrate how they can help you reflect on your assessment design. This module will feel a lot like a review.

Alignment¹

As you reflect on an assessment you have written or selected, the first element you must consider is alignment.

An assessment that is *aligned with standards* measures student performance against those standards.

¹ Kansas State Department of Education, "Assessment Literacy Project" and Ohio Department of Education, "Assessment Literacy: Identifying and Developing Valid and Reliable Assessments" (2013).

To check for alignment, you can ask whether each assessment item is aligned with the standard you intend to measure and teach in the classroom.

If an assessment item is not aligned with the standard you intend to measure, it may measure something else, which can mask what your students actually know and can do. For example, in the module about alignment, we examine the sample item “What is $12 \div 5$?” While this item could be fine in another context, we find that its content is not well aligned with the relevant Tennessee mathematics standard, which states that students should be able to solve division problems using whole numbers. This problem is beyond the scope of the standard because it involves decimals. For example, a student who could easily calculate $12 \div 6$ might struggle with the fact that 5 doesn’t go evenly into 12. In other words, students may answer this problem incorrectly even when they have mastered the relevant standard.²

The checklist prompts you to ask a single, simple question about whether the items in your assessment are aligned with the standards you taught or intend to teach. If you find that they are not aligned, you need to revise them.

Rigor³

The second element you must consider is rigor.

An assessment has an *appropriate level of rigor* if two conditions are met. First, the assessment includes assessment items that match the level of rigor of the skill you intend to measure. Second, the assessment measures a range of student thinking and understanding so that it measures what all students know and can do.

To check for an appropriate level of rigor, you can ask whether the level of rigor of each assessment item matches the cognitive complexity you intend to measure and whether the assessment measures a range of student thinking and understanding so that it measures what all students know and can do.

If your assessment includes only items that require lower-order thinking, and the skill or skills you intend to measure require lower- and higher-order thinking, your assessment results could indicate that students have mastered the standard or standards when they may not have.

For example, the module about rigor includes a multiple-choice item, “Which of the following words is an antonym of *tense*”? As we indicated in that module, this item is not well aligned with the rigor of the relevant skill because it merely asks students to *identify* an antonym from a list, when the skill in the standard expects them to be able to *use* the relationship between antonyms to better understand each of the words, which is a higher-level skill.⁴

The checklist asks you to think about two questions to make sure that your assessment has an appropriate level of rigor. If the answer to either of them is no, you need to go back and revise one or more items of your assessment.

² Tennessee Department of Education, “Tennessee’s State Mathematics Standards: Grade 3” (2010).

³ Kansas State Department of Education, “Assessment Literacy Project” and Ohio Department of Education, “Assessment Literacy: Identifying and Developing Valid and Reliable Assessments” (2013).

⁴ New York State Department of Education, “New York State P-12 Common Core Learning Standards for English Language Arts & Literacy” (2010).

Precision⁵

The third element to consider is precision.

A *precise* assessment measures students' knowledge and skills, not their misinterpretations or lack of unrelated background knowledge.

To determine whether your assessment is as precise as possible, you can ask questions to determine whether your assessment items are accurate and clear. The questions you ask will vary depending on the type of assessment items in your assessment.

For example, if your assessment includes a number of multiple-choice items, a lack of precision could indicate that students have mastered a standard or standards when they may not have. The right answers may be obvious or easier to guess because your distractors are implausible or because one answer choice is much longer and more detailed than the others. Here is an exaggerated example:

Which of the following is a type of mammal?

- a. carrot;
 - b. tree;
 - c. apple; or
 - d. whale
-

Students will likely choose the correct answer because they know that a mammal is a type of animal and choices "a" through "c" are not animals, but that doesn't give you any information about what students know about mammals.

Formatting, typos, incorrect information and other problems can all lead to an imprecise assessment. The checklist helps you check for these and other potential problems. If you find any of these imprecisions, you need to go back and revise your assessment.

Bias⁶

The fourth element to consider is bias.

An *unbiased* assessment measures students' knowledge and skills, not differences among groups of students because of their personal characteristics, such as race, gender, socioeconomic status or religion.

⁵ Kansas State Department of Education, "Assessment Literacy Project" and Ohio Department of Education, "Assessment Literacy: Identifying and Developing Valid and Reliable Assessments" (2013).

⁶ Ibid.

To check for bias, you can ask whether you ensured that each assessment item does not provide an advantage or disadvantage to any group of students because of their personal characteristics, such as race, gender, socioeconomic status or religion.

Remember that even though bias results from *personal* characteristics, we detect bias at the *group* level, not at the *individual* level. This distinction is important because an assessment can provide an advantage or disadvantage to individual students without being biased.

For example, in the module about bias, we describe how an item about a sailing trip that you write to measure reading comprehension may also measure students' familiarity with sailing—something that you do not intend to measure. Typically, sailing is an expensive activity, which means that you may also measure students' socioeconomic status. The item may be biased because students who know more about sailing likely come from more affluent backgrounds than those who do not. It would not be biased because some students simply may know more or less about sailing.

The checklist reminds you to look for bias before you administer an assessment. If you find evidence of bias, you will need to revise your assessment.

Scoring⁷

The fifth element to consider is scoring.

An assessment that has an *appropriate scoring strategy* measures students' knowledge and skills, not how or when the assessment is scored or who scores it.

To determine whether your assessment has an appropriate scoring strategy, you can ask whether you developed an appropriate scoring tool or tools, and whether your tool or tools are well designed.

For example, imagine that you administered a grade-wide essay and that several teachers on your grade-level team scored the students' assessments. Did you and the other teachers use a common rubric? If so, did you engage in a norming activity in which all raters practiced applying the rubric to the same small number of papers, and then discussed inconsistent ratings to calibrate expectations across all those who would be grading papers? Did your team select student exemplars that are representative of each possible performance level for all raters to use as guidance? If you do not employ techniques like these, each team member will likely apply different standards or look at or emphasize different components of writing. Having a plan to score consistently across classrooms if the assessment is grade wide is a strong assessment practice.

The checklist reminds you to make sure that you use strong scoring strategies after you administer your assessment. If you find that you have not identified a strong scoring strategy, develop one before you administer the assessment.

⁷ Ibid.

CHECK FOR UNDERSTANDING

We have addressed the key concepts in this module, so let's review our goal.

At the outset of the module, we set a goal that you would be able to use the assessment checklist to help you determine whether an assessment you design appropriately addresses the five elements of assessment design featured in this series of modules.

To determine whether we have achieved our goal, let's check your understanding with two assessment items.

Here's the first item:

What question can you ask to check an assessment for alignment? If you like, use the assessment checklist to help you answer the question.

Pause this video if you want a few moments to think about your answer or discuss it with colleagues.

A sample answer to the item would be: Is each assessment item aligned with the standard that I intend to teach and measure?

Here's the second item:

What question can you ask to check for bias? If you like, use the assessment checklist to help you answer the question.

Pause this video if you want a few moments to think about your answer or discuss it with colleagues.

A sample answer to the second item would be: Did I ensure that each assessment item does not provide an advantage or disadvantage to any group of students because of their personal characteristics, such as race, gender, socioeconomic status or religion?

CONCLUSION

Thank you for completing this review and for considering the checklist as a tool that you can use to reflect on the assessments you write or select before you administer them.

SOURCES

Kansas State Department of Education. "Assessment Literacy Project." <http://www.k-state.edu/ksde/alp>.

New York State Department of Education, "New York State P-12 Common Core Learning Standards for English Language Arts & Literacy." 2010.
http://www.p12.nysed.gov/ciai/common_core_standards/pdfdocs/p12_common_core_learning_standards_ela.pdf.

Ohio Department of Education. "Assessment Literacy: Identifying and Developing Valid and Reliable Assessments." 2013. <http://education.ohio.gov/getattachment/Topics/Teaching/Educator-Evaluation-System/How-to-Design-and-Select-Quality-Assessments/AL-Training-PPT-FINAL-for-Distribution.pdf.aspx>.

Relay Graduate School of Education. *Designing and Evaluating Assessments*. 2014.

Relay Graduate School of Education. *Rules for Constructed Response Item Design*. 2013.

Relay Graduate School of Education. *Rules for Multiple Choice Item Design*. 2013.

Rhode Island Department of Education. "Deepening Assessment Literacy." Accessed March 23, 2015.
http://media.ride.ri.gov/PD/Eval/Deepening_Assessment_Literacy_output/story.html.

Tennessee Department of Education. "Tennessee's State Mathematics Standards: Grade 3." 2010.
http://www.tn.gov/education/standards/math/std_math_gr_3.pdf.