

Assessing Skill Development

By Jon Mueller

Most educators are familiar with instances of authentic assessment of *content* within the disciplines (e.g., ask elementary science students to present a report describing the diet and feeding habits of an animal of their choosing) or of authentic assessment of *discipline-specific skills* (e.g., ask social studies students to interpret a political cartoon). In such authentic assessments, students apply the knowledge and skills of the discipline to situations or tasks that replicate real world challenges. The measurement of skills is particularly well suited to authentic assessment because meaningful demonstration of skill acquisition or development requires a performance of some kind.

Yet, we rarely assess, authentically or otherwise, the critical *cross-curricular or process skills* such as self-assessment, information literacy, collaboration, or metacognitive skills. Certainly, one reason is the elusive nature of some of these skills that can make it a challenge to find effective measures of assessing them. How would one measure metacognitive abilities or leadership skills or the ability to evaluate the accuracy and relevancy of information? Furthermore, these skills are infrequently included in state goals or standards, so they receive less assessment attention or urgency.

Nevertheless, these critical skills definitely can and should be assessed. To explain how, I will describe the *summative* assessment of skill development, that is, “final” judgments of whether or not a skill has been acquired and can be applied. If, for example, you set a goal of producing students who can assess their own progress and make adjustments when necessary, and you teach those specific skills because you believe they are critical to your students’ future success in so many contents, you will want to know if your students have successfully met this goal.

SUMMATIVE ASSESSMENT OF SKILLS

Because skill development so naturally fits with authentic assessment, it makes sense to adapt the four-step process of creating authentic assessments described in my online text, *Authentic Assessment Toolbox*, found at <http://jonathan.mueller.faculty.noctrl.edu/toolbox/index.htm>.

Step 1: Writing Skills as Standards

The process of creating an authentic assessment begins with identifying a good standard, including writing it in observable and measurable language.

When assessing content, some educators struggle with writing a standard that captures comprehension of certain material without relying on verbs such as “understand” or “know.” They reasonably ask, “What if what I really want is for my students to *understand* a concept or process? How can I state that in an observable form and capture it authentically?”

A simple, direct approach to address that question is to ask “How could the students demonstrate that they understand the concept or process? What would that look like?” For example, if you want students to understand the differences between books, periodicals, and databases, you could ask, “How could they show me they understand the differences between them?” The students could describe the differences between books, periodicals, and databases in a paragraph or two; they could locate each type of resource in a library when asked to; they could identify which type of resource is most likely to contain certain types of information. Each of those statements (i.e., describe the differences between books, periodicals, and databases; locate each type of resource in a library; identify which type of resource is most likely to contain certain types of information) then can serve as a standard to be taught and assessed. They describe observable and measurable behaviors that are important to information literacy.

A similar approach can be followed in writing standards for skills. Do you want students to be reflective learners, good leaders, and information literate citizens? How could you judge whether or not they are able to do so? Ask, “What do these skills look like if they are done effectively?” “What behaviors are commonly exhibited by reflective learners, good leaders, and information literate citizens?” That can help you narrow down broad and often nebulous goals such as “being information literate” into more specific and observable standards. For example, in answering these questions, I would argue that an information literate person should be able to:

- frame a question to be answered
- locate appropriate sources of information
- access information from a variety of sources
- evaluate the accuracy, relevancy, authority, and currency of information
- determine when sufficient information has been acquired to answer the question.

Step 2: Designing Tasks to Assess the Skills

In Step 1, you identified an important skill. How can students demonstrate that they have acquired this skill and apply it in relevant contexts? Give them opportunities to do so. Create simple or complex tasks by asking questions such as “When would someone ever use this skill?” or “Why would someone ever need to know how to do this?”

Some situations identified by answering such questions will be brief, simple tasks commonly performed. For example, one skill I work to develop in my students is the ability to evaluate claims presented in the media or in the course of an informal conversation. This is a skill they will have many opportunities to apply. Are my students prepared to do so? Are they capable of judging the validity of a claim based upon the evidence presented? To assess that, you can present students with brief scenarios in political, scientific, or other contexts in which you ask them to evaluate a claim in light of the accompanying evidence. For example, see Figure 6.1.

FIGURE 6.1

EVALUATING A CLAIM SCENARIO

A large study in Finland found evidence that people who ate fish less than once a week ran a 31 percent higher chance of mild to severe depression than people who ate it more often. The researchers concluded that eating fish lowers the risk of depression.

1. Is the researchers’ conclusion warranted given the results? Explain.
2. Circle the headline below that you believe BEST fits the research results.
 - a. Eating Fish Can Make You Happy
 - b. Being Depressed Lowers Fish Consumption
 - c. Fish Eaters Are Happier People
3. Explain why you selected that headline.

“The process of creating an authentic assessment begins with identifying a good standard, including writing it in observable and measurable language.”

To serve the function of summatively assessing students, these brief, constructed-response items could be incorporated into a larger test or project so a sufficient number of items could be administered to address the relevant standards. In fact, the question in Figure 6.1 is one of 25 multiple-choice and short-answer essay questions included in a test we give our majors in psychology to assess their scientific thinking.

In other contexts, the skill set needed to address a problem or question may be more extensive and require more time. For example, if someone must support his own claim with evidence, a more complex task—such as a research paper—can be assigned to students to assess whether or not they can apply the full set of information literacy skills in a meaningful context.

MULTIPLE AND VARIED

To increase the reliability and validity of inferences drawn from student performances on these tasks, the assessments should be multiple and varied (Stiggins, 1987). In fact, using brief constructed-response items as described above permits the repeated testing of student skills across a variety of contexts. A single large task enables students to provide evidence they can integrate several skills into a more complex product, but it provides only a single sample of such skill application and integration, limiting the confidence one can have in the reliability and validity of inferences drawn. Along with a research paper, you could ask students to demonstrate information literacy skills on some smaller research tasks of different types to provide a greater number and variety of pieces of evidence of skill proficiency.

Step 3: Identifying the Criteria for the Skill

Furthermore, just because a student can complete a large task such as a research paper does not in itself indicate proficiency at the complete set of information literacy skills. To determine that, it is necessary to identify the specific characteristics of good performance on that task, that is, the criteria. What are the observable behaviors, the *behavioral indicators*, of proficiency on a particular skill? For example, what does it look like to be proficient at framing a good research question? Criteria of a good research question might include *clarity, a specific focus, appropriate breadth, and relevance to the topic.*

Or, what does good collaborative behavior look like? The characteristics of good collaborative behavior, the criteria on which you would judge student performance, might include behaviors such as *involves others in the task, participates without prompting, seriously considers the ideas of others, and offers helpful feedback.*

Step 4: Creating Rubrics for Rating Skill Performance

Once the criteria for a task have been identified, a rubric, or rating scale, can be used to judge how well someone has met the criteria for performance on that skill task. Authentic assessment of skills does not require a rubric, but the use of rubrics can increase the consistency of application of the criteria (Marzano, 2006). Additionally, by articulating the criteria and the characteristics of good performance at each level (descriptors), those learning and performing the skill and those teaching and assessing it will share a clearly defined picture of what proficiency should look like.

For example, to assess the item in Figure 6.1 on evaluating claims about eating fish and depression, you can use the following rubrics:

FIGURE 6.2 RUBRIC FOR EVALUATION OF RESEARCH CLAIMS (QUESTION 1)

Rubric for Research Claim Question		
0	2	3
<ul style="list-style-type: none"> Answered YES OR Answered NO but provided inadequate explanation 	<ul style="list-style-type: none"> Answered NO With adequate support for position 	<ul style="list-style-type: none"> Answered NO With excellent support for position

FIGURE 6.3 RUBRIC FOR EVALUATION OF RESEARCH CLAIMS (QUESTIONS 2 AND 3)

Eating Fish and Depression (Questions 2 and 3)			
0	1	2	3
<ul style="list-style-type: none"> Circled incorrect headline And provided inadequate explanation 	<ul style="list-style-type: none"> Circled incorrect headline But provided adequate explanation 	<ul style="list-style-type: none"> Circled correct headline And provided adequate explanation 	<ul style="list-style-type: none"> Circled correct headline And provided excellent explanation

“Authentic assessment of skills does not require a rubric, but the use of rubrics can increase the consistency of application of the criteria.”

To examine how you can implement the four steps of creating an authentic assessment of skill development, let’s look at another example. For Step 1, we will use the information literacy standard previously listed: *Students will be able to frame a question to be answered.* To partially assess this standard, we could use the following authentic task (Step 2):

LEON’S RESEARCH QUESTION

Leon asks you for help. He is assigned to complete a three- to five-page report on something related to the three branches of the U.S. government. Leon tells you that his dad says that judges have too much power and they are running this country. Leon thought the President was supposed to run the country. Shouldn’t he be making the decisions? And what does Congress do? Why don’t they tell the judges to stop messing things up? So, Leon decided the question he wanted to research was:

Question: What could the President do to fix things?

Has Leon identified a good question to guide his research? Why or why not?

Write another question that you think would serve as a good question to guide Leon’s research.

To assess student performance on Leon’s Research Question Task, we will create a simple rubric (Step 4) built on a set of criteria (Step 3) for judging student responses to the task prompt. (See figure 6.4)

With brief constructed-response items like the Research Question Task, we can determine if students are capable of distinguishing a good question from a poor one. We can also assess students’ ability to frame their own questions for a research topic. Thus, we will have collected good, sufficient evidence to summatively assess the application of this critical skill. A similar approach can be applied to any skill that is worth assessing and can provide observable evidence. 🌈

REFERENCES

Marzano, R. J. (2006). *Classroom assessment and grading that work*. Alexandria, VA: ASCD.

Stiggins, R. J. (1987). “Design and Development of Performance Assessment.” *Educational Measurement: Issues and Practices*, Fall: 33–42.

Jon Mueller is a professor of psychology at North Central College in Naperville, Illinois. He is also the author of *Assessing Critical Skills* (Linworth Publishing, 2009).



FIGURE 6.4 RUBRIC FOR DEVELOPING A RESEARCH QUESTION

CRITERIA	BEGINNING	DEVELOPING	PROFICIENT
Leon’s question:			
Clear focus	Difficult to identify focus of investigation		Easy to identify focus of investigation
Relevant to topic	Not related to intended focus	Somewhat related to intended focus	Question captures intended focus
Appropriate breadth	Far too narrow or broad for assignment	Somewhat too narrow or broad for assignment	Breadth of question consistent with assignment
New question:			
Clear focus	Difficult to identify focus of investigation		Easy to identify focus of investigation
Relevant to topic	Not related to intended focus	Somewhat related to intended focus	Question captures intended focus
Appropriate breadth	Far too narrow or broad for assignment	Somewhat too narrow or broad for assignment	Breadth of question consistent with assignment