

CHAPTER 8

AUTHENTIC ASSESSMENT

Looking Ahead in This Chapter

A modern wrinkle in classroom assessment is the movement toward “authentic” assessment. Defining what makes assessment *authentic* is not a simple matter, though. Methods for ensuring that assessment is authentic and scored reliably are presented. This chapter also shares many real-world examples of authentic assessment across many subjects and grade levels. A research-based scoring rubric that can be used to assess the authenticity of an assessment is included at the end of the chapter.



Objectives

After studying this chapter, you should be able to

- Identify the characteristics of assessment that make it authentic
- Explain the reasons today's teachers wish to make their assessments authentic
- Design authentic assessments that can be scored fairly and objectively
- Identify authentic assessment tasks for different populations and subject areas
- Describe what a schoolwide approach to authentic assessment might look like



Mr. Hernandez Gets Real

"All right, it is time to start math, so please get your homework from yesterday with a red pen to grade," Mr. Hernandez said.

Riel started complaining and was slowly pulling out his math book from his desk. He took forever finding a red pen to grade the homework. Mr. Hernandez asked students to open to Chapter 14 over the measures of central tendency—mean, median, mode. "Riel, face the front please. Open your book please."

Now Riel began complaining about how stupid math is. Mr. Hernandez continued with the lesson, working a few examples on the board, having students work a few, and then assigning another worksheet for additional practice. Riel got the worksheet, put it in his folder without working on it, and then started making faces.

"Have you completed your assignment?"

"Almost. I can do the rest at home though."

"Let me see it, please." Looking at a blank assignment, Mr. Hernandez said, "Let me watch you do a few to make sure you understand what we are doing."

"I know *how* to do it; I don't know why we *have* to do it. When am I ever going to have to calculate the mode of anything in my life? I have never heard my mom or dad use the word *mode*, so why do I have to learn it?"

That was a good question. Mr. Hernandez had thought about this before, of course. He had heard the "why do we need to know this?" question before. It was a dilemma he had faced when it came to assessment, as well. Usually the way he tested for a topic or a skill wasn't very much like the way the real world would expect that knowledge or ability to be used. It was a problem he had always struggled with for much of what he taught. This time, though, he began to think of a more permanent solution to the problem . . .

(To Be Continued)

For decades, one of the criticisms leveled at education by some is that students aren't being equipped with the right tools to do well in the real world (e.g., on the job, in college, in a democracy, in an information-driven economy), and schools are constantly evaluating and modifying curricula to meet the needs of "the marketplace." It makes sense, then, that a parallel concern in classroom assessment would arise, and a growingly accepted principle in the modern world of classroom assessment is that it is best if assessments are "realistic," whatever that means.

Of course, there is something intrinsically artificial about testing situations, and it may create a paradox to attempt to test reality. A classroom teacher may find it difficult (and perhaps it is theoretically impossible) to observe student classroom performance in ways that reflect how students would perform outside the classroom. Nevertheless, it is reasonable on its face that assessment tasks that somehow match real-world tasks and expectations should provide a more valid picture of the student skills that matter most, those skills and knowledge sets that prepare students for success in whatever comes next for them in the world. For example, it is clear that in the real world students do not answer multiple-choice questions very often, so any multiple-choice test is bound to be more artificial than, for example, asking students to write a persuasive letter or e-mail.

Performance Assessment and Authentic Assessment

Good Question!

Q: Is authentic assessment the same as performance-based assessment?

A: A simple question with an almost simple answer. And an important enough question that it was asked before in Chapter 1. Let's start with definitions of those terms. We have been defining performance-based assessment as *assessment that requires examinees to perform or produce something for evaluation that is intended to assess skill or ability*. Let's use the simplest definition of authentic assessment (though this chapter explores the usefulness of more complex definitions): *Assessment that aligns with real-world tasks and expectations*. So the two approaches to assessment are not the same. Certainly, some performance-based assessments are authentic. The performance-based format, because it usually includes an intrinsically meaningful task and assesses generalizable skills, lends itself to authenticity. But one could easily design a performance-based assessment that is not authentic. It is the way one assesses that determines whether it is authentic, not simply whether one is assessing skill or basic knowledge.

IN SEARCH OF "AUTHENTIC" ASSESSMENT

The development of standardized testing procedures and the statistical methods of analyzing items and test scores, beginning early in the 20th century,

brought the tools of *science* to the *art* of education and classroom assessment. As the emphasis on standardized test administration and performance grew to eclipse the perceived value of other assessment approaches, however, criticisms of the artificially low level of task complexity and lack of teacher control in large-scale objectively scored testing resulted in a movement to return classroom assessment to a more realistic, student-centered approach that measured more complex and deeper student thinking. This approach has been labeled as **authentic assessment**.

A frequent piece of advice given to teachers designing their own assessments is that the *best* classroom assessments are *authentic* (Archbald & Newmann, 1988; Burke, 2009; Gronlund, 2003; Swaffield, 2011; Wiggins, 1989; Wilson & Schwier, 2012), which in most textbooks and in teacher training materials is usually defined as some version of “realistic” or “mirroring the real world.” However, while most agree that “authentic” assessment is best practice, there are a variety of definitions of authenticity presented in the research literature and the teachings of experts in classroom assessment.

The earliest reference to *authentic* tests is likely that made in a book critical of standardized testing by Archbald and Newmann in 1988. Newmann argued that assessment is authentic if it assesses tasks that “have meaning or value beyond success in school” (Newmann, Brandt, & Wiggins, 1998, p. 19). The other early advocate, and the most cited, is Grant Wiggins. “‘Authentic’ refers to the situational or contextual realism of the proposed tasks,” he tells us (Newmann et al., 1998, p. 20). However, while many others also speak of authentic in the context of application outside the classroom, some don’t, and some emphasize other aspects of assessments as “authentic.” For example, Wiggins and early advocates of authenticity emphasized the importance of taking a mastery approach in the assessment process (i.e., using a criterion-referenced philosophy with a goal of moving all students toward the same teacher-determined level of mastery), although most others who wrote later pay little attention to that component.

If the field recommends that teachers should do a particular “thing,” it is important to know what that thing is. What criteria should be used in determining whether any specific teacher-made assessment is authentic and, therefore, produces the benefits presumably associated with authenticity? A recent study reviewed hundreds of journal articles, presentations, books, and dissertations to identify concrete criteria for evaluating the authenticity of an assessment (Frey, Schmitt, & Allen, 2009). Surprisingly, beyond the simple requirement that assessments should be realistic (whatever that means), nine different components or dimensions were identified in the literature as characteristics of authentic assessment.

These different dimensions that define authenticity amount to a nice list of qualities to which all classroom assessment, regardless of format, can aspire. Authentic assessment is defined in a wide variety of ways, usually including one or more of these nine characteristics, which can be grouped into three broad categories:

- The context of the assessment
 - Realistic activity or context.
 - The task is performance-based.
 - The task is cognitively complex.
- The role of the student
 - A defense of the answer or product is required.
 - The assessment is formative.
 - Students collaborate with each other or with the teacher.
- The scoring
 - The scoring criteria are known or student developed.
 - Multiple indicators or portfolios are used for scoring.
 - The performance expectation is *mastery*.

Notice that two components of authenticity are two of our broad approaches to modern classroom assessment; authentic assessment should be *performance-based* and *formative*. This illustrates the overlapping understanding in the field as to what *authenticity* means as a distinct assessment philosophy. Table 8.1 presents a small sampling of the different definitions of authentic assessment one finds across textbooks and research articles. Some experts emphasize the realism aspect, some stress the importance of student participation or collaboration in the assessment process, and some place importance on the scoring criteria or the portfolio nature (multiple indicators or samples of work) of the assessment.

Defining Authentic Assessment

Good Question!

Q: I understand that different researchers define authentic assessment differently, but this is a textbook, right? Can't you just provide a simple definition, if we promise to remember that it is a complicated issue?

A: Fair enough. Assessment is authentic when the tasks, content, expectations, and evaluation methods of the assessment are similar to the meaningful tasks, content, expectations, and evaluation methods outside the classroom in the real world. The real world for students depends on their age; it could be playing and socializing with others, engaging in higher education, or performing on the job, now or in the future.

Table 8.1 Varying Definitions of Authentic Assessment

	Context			Student Role			Scoring		
	realistic activity or context	performance-based	cognitively complex	defense required	formative assessment	collaborative	known or student developed criteria	multiple indicators or portfolios	mastery expectation
Abernethie, 2006									
Archbald, 1991									
Darling-Hammond, Aness, & Falk, 1995									
Dez, Moon, & Meyer, 1992									
Dutt-Doner & Maddox, 1998									
Engel, Pulley, & Rybinski, 2003									
French, 2003									
Green, 1998									
Herrington & Oliver, 2000									
Jolly & Kettler, 2000									
Kellaghan & Madaus, 1993									
Lawton, 2000									
Maden & Taylor, 2001									
Meisels, 1996									
Meyer, 1992									
Moorcroft, Desmarais, Hogan, & Bekowitz, 2000									
Mueller, 2005									
Newmann, Brandt, & Wiggins, 1998									
Paris & Ayres, 1994									
Ratcliff, 2001									
Schnitzer, 1993									
Spinelli, 1998									
Stripling, 1993									
Suen, 1997									
Torrance, 1993									
Wiggins, 1989									

Source: Adapted from Frey, Schmitt, & Allen, 2009.

Note: Shaded areas indicate presence of the component.

WHAT WE KNOW ABOUT AUTHENTIC ASSESSMENT

Advocates for authentic assessment suggest that assessment systems that focus on higher-order thinking skills, problem solving, investigation, and analysis (some of the key real-world skills) can drive improved teaching and curriculum coverage (Torrance, 2009). These perceived strengths of authentic assessment are difficult to evaluate in the current political and policy environment that now treats assessment as a critical step in the school improvement process (and, of late, critical data in evaluating teachers).

Scores on tests are no longer treated as indicators of learning, but more and more often they have *become* the outcome of interest. Learning is not the goal; high test scores are the goal. Of course, the tests in this perspective are typically large-scale standardized tests, not teacher-made classroom assessments, but, nevertheless, the national discussions about education tend to merge summative tests with curriculum and classroom objectives as if they are all the same thing. The assessment *is* the objective. Consequently, this sort of risky thinking can be fueled by any suggestion that an assessment approach can improve teaching. The point as it applies to authentic assessment, though, is reasonable. Assessment that involves students collaboratively in identifying criteria for quality of performance, choosing tasks or questions that require thinking at the high levels of Bloom's Taxonomy (for instance), and developing realistic problems or complex creative activities to evaluate requires instruction, teaching examples, and modeling that supports development of the necessary cognitive skills and knowledge base. Once outcomes are authentic, teaching should ultimately become integrated with assessment.

Knowledge and Context

A theory about conceptual knowledge and the difficulty that students have transferring that knowledge to other situations suggests that abstract learning generalizes better when it occurs in multiple settings with students playing multiple roles. If true, then authentic assessment would seem to be the best approach to support generalization and transfer.

Brown, Collins, and Duguid's (1989) *Theory of Situated Cognition* is based on research that found, contrary to classic beliefs by educators for centuries, that conceptual knowledge cannot easily be transferred out of a specific context and applied to a new problem. Instead, the theory goes, knowledge is situated and is part of the activity, context, and even culture in which it is developed and learned. An

(Continued)

(Continued)

instruction and assessment strategy that the authors call **cognitive apprenticeship** is suggested as an alternative to traditional practice.

Cognitive apprenticeship is a collaborative approach that emphasizes group learning. Key components of the instructional and (performance) assessment tasks include these:

Collective problem solving. Groups act as more than just separate insights that are combined and cataloged. They often produce results that are more than just the sum of parts.

Displaying multiple roles. Students need to understand different parts of the process, different pieces of the product. Opportunities are created for students to play different roles in the activity, and by adding a reflective component, the task becomes more authentic.

Confronting ineffective strategies and misconceptions. Authentic assessment includes identifying what does not work or what is wrong. The discussions and observations available in group interactions allow for presenting misunderstandings.

Assessing collaborative work skills. The real world of work requires working together. Collaborative tasks allow for assessment of collaborative skills as students work and learn as a group.

It turns out that using authentic assessment well can be tough. After reviewing a series of studies of implementation of authentic assessment systems in schools a few years ago across the United Kingdom, Torrance (1995) concluded that teachers “had enormous difficulty in interpreting, conducting and assessing the tasks—precisely because they were ‘authentic’, they were too complicated to communicate easily . . . and too demanding for teachers to conduct under ordinary classroom conditions” (p. 55). He suggests that teachers’ willingness and ability to adopt new assessment approaches is influenced by their long-standing notions of what assessment is and what its purpose is. Authentic assessment might be adopted more easily if it is viewed as a framework for assessment, as opposed to a particular distinctive approach. In other words, performance-based assessment, formative assessment, and even traditional assessment might benefit, in terms of validity, from the layering on of authentic elements to the established components of these other more accepted approaches.

THE CASE FOR AUTHENTIC ASSESSMENT

Let’s assume that, whatever the precise definition of authentic assessment is (or should be), it includes the dimensions of realism, student involvement, and multiple components for scoring. Those basic aspects have strong support

across most experts. What, then, makes that approach a best practice? What is the theoretical mechanism by which the approach should work? What makes authentic assessment valid?

Judith T. M. Gulikers (2006), a Dutch researcher, suggests several reasons for embracing authentic assessment. Though writing primarily about college students and others training for specific jobs, her conclusions also make sense regarding elementary and secondary school. First, contextualizing assessment in interesting, real-life tasks is a “crucial element” of competency-based assessment that is consistent with modern approaches to education. Second, authenticity should increase validity. This could happen in several ways. Construct validity (i.e., the assessment measures the trait of interest) should be increased when tasks “represent the real-life problems of the knowledge domain assessed and that the thinking process that experts use to solve the problem in real life are also required by the assessment task” (p. 21). Further, the common assessment concern of underrepresenting the construct can be countered by providing a richly detailed context, which is frequently seen in authentic assessment and in real life. Another potential validity benefit suggested by Gulikers has to do with the process by which assessment tends to eventually drive instruction. By focusing the assessment on real-world skills and expectations, the curriculum remains likewise focused, as does instruction.

Gulikers suggests two mechanisms by which authentic assessment might influence student learning. This would place authentic assessment in the same rarified air as formative assessment—a classroom assessment approach that can actually increase learning. First, it may stimulate a “deep study” approach leading to greater understanding and skill development. Second, authentic assessment may increase student motivation to learn when tasks are seen as relevant and useful.

Theories of Learning

There is a classic theory of how student learning happens based on the behavior of students, and there is a modern theory of learning, which focuses on how the mind works. As one might expect, the modern theory is more supportive of modern classroom assessment approaches like authentic assessment.

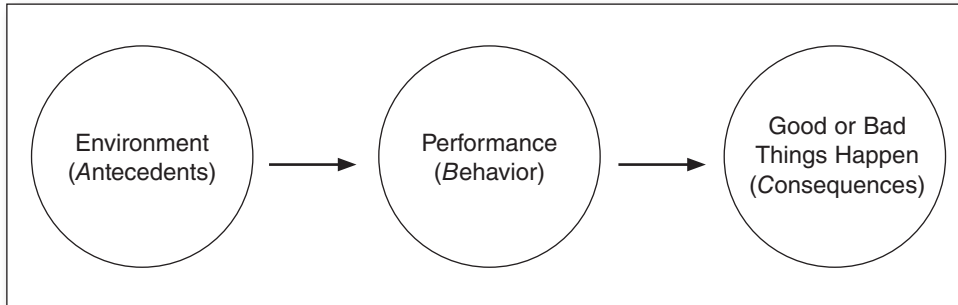
ABC Model

The traditional theory explains learning by describing the classroom as a series of (A)ntecedents, the environmental context that leads to some behavior; the

(Continued)

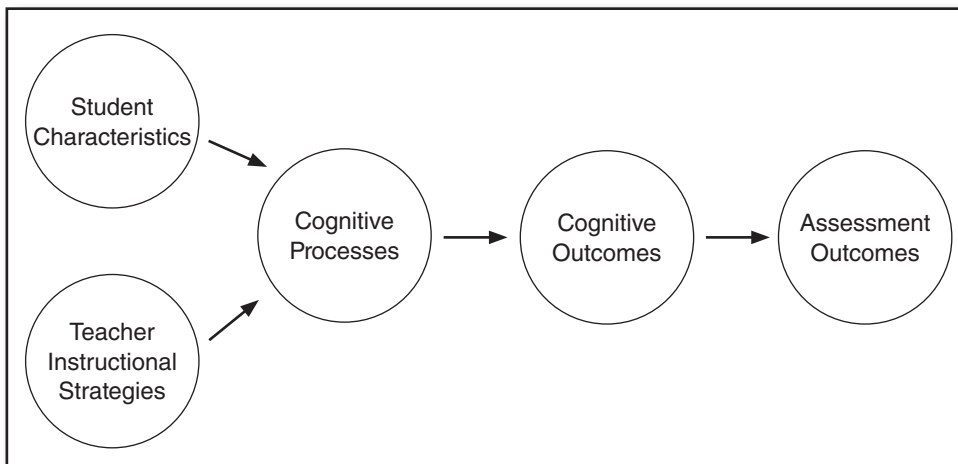
(Continued)

(B)ehavior of interest; and the (C)onsequences, which are whatever happens to the student after the behavior. For example, the teacher lectures (Antecedent), the student takes notes (Behavior), and then the student receives a high score on the test (Consequence).



Cognitive Model

A recent theory of learning (starting to appear about 40 years ago, which is just yesterday in terms of human civilization) understands student learning as a cognitive process. The teacher and student bring certain elements to the process—the student has particular characteristics (intelligence, skills, experience,



motivation) and the teacher provides some instruction. This results in some thinking inside the head of the student, learning occurs, and assessment makes that learning observable.

Tombari and Borich (1999), in a nice guide on applying authentic assessment in the classroom, point out that the behavioral ABC model envisions that the goal of a classroom environment is to produce correct answers on tests. The cognitive model assumes the goal of good teaching is to promote good thinking, the smart, cognitive processes that result in those correct answers (Borich & Tombari, 2004). This is clearly consistent with one goal of authentic assessment, to produce what Wiggins (1989) and others call healthy "habits of mind."

Reality as a Construct

Teachers might disagree with students and students might disagree with each other about whether a particular assignment or assessment task is realistic and meaningful. It depends on one's individualized experiences and perceptions (Gulikers, 2006). Teachers must choose tasks they consider realistic. This is easier said than done. What is reality, after all?

This big question about the nature of reality is, of course, one of the major philosophical questions of modern civilization. A famous story, Plato's *Allegory of the Cave*, provides a metaphor for both understanding reality and, usefully for us, interpreting test scores as mere shadowy indications of student ability and knowledge. Plato writes of an imagined conversation between a teacher, Socrates, and a student, Glaucon.

Socrates: And now, let me show in a figure how far our nature is enlightened or unenlightened: – Behold! human beings living in an underground cave, which has a mouth open towards the light and reaching all along the cave; here they have been from their childhood, and have their legs and necks chained so that they cannot move, and can only see before them, being prevented by the chains from turning round their heads. Above and behind them a fire is blazing at a distance, and between the fire and the prisoners there is a raised way; and you will see, if you look,

(Continued)

(Continued)

- a low wall built along the way, like the screen which marionette players have in front of them, over which they show the puppets
- Glaucou: I see.
- Socrates: And do you see, men passing along the wall carrying all sorts of vessels, and statues and figures of animals made of wood and stone and various materials, which appear over the wall? Some of them are talking, others silent.
- Glaucou: You have shown me a strange image, and they are strange prisoners.
- Socrates: Like ourselves, and they see only their own shadows, or the shadows of one another, which the fire throws on the opposite wall of the cave?
- Glaucou: True, how could they see anything but the shadows if they were never allowed to move their heads?
- Socrates: And of the objects which are being carried in like manner they would only see the shadows?
- Glaucou: Yes.
- Socrates: And if they were able to converse with one another, would they not suppose that they were naming what was actually before them?
- Glaucou: Very true.
- Socrates: To them, the truth would be literally nothing but the shadows of the images.

The Republic, Plato (Lindsay, 1991)

The choices teachers make when they use authentic assessment are twofold: Determine what is authentic and then determine a way of producing scores that represent reality. Whatever that is.

SCORING AUTHENTIC ASSESSMENTS

If one applies the nine key dimensions of authenticity, some characteristics hinder reliability, while others should increase reliability. The dimensions of authenticity overlap substantially with performance-based assessment; consequently, authentic assessment shares some of the same reliability issues.

Chapter 7 discusses the scoring concerns with performance assessments in general.

Subjective assessment leads to subjective scoring, so inter-rater reliability might be a problem for the authentic assessment format. On the other hand, if the scoring criteria themselves are created collaboratively with students, this suggests that both the teacher and the students may share a fairly precise understanding of how the criteria should be applied. A solid scoring rubric developed with input from all members of the learning community may provide enough concrete guidance that the subjective nature of rubric scoring is lessened. Gipps (1995) examined a series of authentic assessments put into place across Great Britain and did indeed find that the subjective nature of scoring authentic assessment tasks resulted in poor inter-rater reliability both between teachers and within individual teachers (scoring the same products differently on different occasions). The British assessments, however, were national, standardized (or intended to be) tests and developed administratively. These were not individual classroom assessments developed collaboratively with student involvement.

One of the characteristics of authentic assessment that are commonly emphasized among scholarly publications is that scoring for authentic assessment should be based on multiple indicators, at a minimum, and based on a portfolio of work products at best. As Chapter 2 explained, internal reliability, the consistency of scores across the pieces, tasks, and items within an assessment, often increases as the number of those items increases. Simply put, scores based on a large number of subscores tend to be more reliable. This aspect of authentic assessment, then, could help with reliability, perhaps protecting against the chance element of subjective scoring.

WHAT AUTHENTIC ASSESSMENT LOOKS LIKE IN THE CLASSROOM

“I can’t claim to be an authority on anything, but I can honestly say that certain matters absolutely fascinate me, and that I write about them all the time. The two basic topics which fascinate me are ‘What is reality?’ and ‘What constitutes the authentic human being?’ . . . I consider that the matter of defining what is real—that is a serious topic, even a vital topic. And in there somewhere is the other topic, the definition of the authentic human. Because the bombardment of pseudo-realities begins to produce inauthentic humans very quickly, spurious humans—as fake as the data pressing at them from all sides.”

Philip K. Dick, Science Fiction Author (1928–1982)

A recurring theme in this book is that there is often a gap between the theoretical classroom assessment approach that one wishes to apply and what that approach looks like when it is actually put into practice in some form. There are realities and logistics (e.g., time in the day, energy and motivation, policies, pressures, training) that often prevent valid application of assessment principles relevant to the chosen approach. This is very much the case with authentic assessment because, as the title of one book on authentic assessment proclaims, “authenticity is in the eye of the beholder” (Gulikers, 2006). As the quotation that opens this section suggests, agreeing on reality is perhaps the toughest step in designing authentic assessments, but it is only the first step. The rest of this chapter describes the ways that real-life classroom teachers and school districts can and have translated the principles of authentic assessment into day-to-day assessment practices. For authentic assessment, perhaps more than any other approach, though, it is the thoughtful teachers who can best evaluate what skills and knowledge base are most relevant for the real-world problems, challenges, and tasks facing their students.

Broad Strategies for Developing Authentic Assessments

Baron and Boschee (1995) and others (Burke, 2009; Meyers & Nulty, 2009) provide specific suggestions for developing assessments of a wide variety of authentic tasks. They begin by offering three points to keep in mind when planning for assessments with authenticity:

1. Not all assessments have to be authentic. Traditional paper-and-pencil tests are still very useful tools for assessing important basic knowledge and many skills. If the majority of your own classroom assessments are made up of matching and multiple-choice questions, that might be perfectly appropriate.
2. Regardless of the specific tasks or content involved, authentic assessment should be considered. Consider the extent to which each instructional objective reflects valued skills outside the classroom. This advice is consistent with the logic of backward assessment described in Chapter 3.
3. Not all authentic tasks need be assessed; they may be formative or act as instruction. Authenticity is likely to improve the quality of any type of assessment.

The most difficult part of authentic assessment is the creativity and thoughtfulness necessary to think of and identify authentic tasks that have value. Start

by identifying several critical or clearly relevant issues (within the domain of interest) and then listing one or two learner outcomes for each of those issues. Next, identify the criteria for success in each of those outcomes. Pick several criteria, but don't worry if you haven't covered them all. The tasks are likely authentic if they involve complex thought (e.g., problem solving, analysis, investigation), are interesting, and seem relevant and engaging to students.

Baron and Boschee also suggest a simple multistep process for producing scoring rules for authentic assessment:

1. Design a scoring rubric for each criterion. It should evaluate the degree to which students have incorporated all the important components of the thinking process involved.
2. Verify that the task will provide all the information necessary to produce a valid score.
3. Consider modifying the task to increase interest or the amount of information produced.
4. For your records, include on the rubric the task, learner outcome (objective), and complex thinking skill required.
5. Consider assigning different weights to each criterion if they differ in importance.
6. Share the criteria with students and your reasoning in selecting the criteria. (Consider student involvement in criteria selection.)
7. Of course, share results with students, emphasizing what they have mastered or learned.

Fischer and King (1995) published a concise guide to implementation of authentic assessment. They suggest that a portfolio assessment approach is best and provide a list of eight authentic characteristics that are found in real-world classroom assessments. This style of assessment will contain several of these components:

1. Represents realistic tasks in a variety of contexts done for a variety of purposes
2. Ongoing, formative assessment
3. Samples a wide range of cognitive strategies
4. Designed for different developmental levels
5. Individualized
6. Provides for collaborative reflection between students and teachers

7. Assessment guides instruction
8. Emphasizes what students know and can do

Dozens of concrete examples of what authentic assessment looks like in elementary classrooms are provided by Montgomery (2001). She points to many authentic activities already seen in most contemporary schools, such as map-making, writing plays, producing videos, writing computer programs, and making up stories. She argues that authentic assessment tasks simulate important real-world challenges and require high levels of complex thinking. These sorts of assessable tasks are becoming more and more common in modern classrooms. She suggests many authentic assignments for real-world teachers to try (pp. 36–37), including the following:

1. Keep a five-day record of precipitation and temperature for any five cities in the United States. Graph the results.

Table 8.2 Assessing Skills With Authentic Tasks

Skill	Definition	Assessable Tasks
Procedural knowledge	Knowledge of how to perform, how to do something	<ul style="list-style-type: none"> • Thinking aloud • Using a computer • Safety procedures • Driving a car • Conducting an experiment • Showing work while solving a math problem
Problem solving	Use of critical-thinking and decision-making skills to find a solution	<ul style="list-style-type: none"> • Testing a hypothesis • Writing a research paper • Making value judgments • Solving mathematical “story problems” • Judging the credibility of evidence • Deductive reasoning (e.g., geometry problems) • Concept mapping to identify variables of a problem
Collaboration	Working with others toward a shared goal	<ul style="list-style-type: none"> • Listening (e.g., eye contact, asking questions, reflective responses) • Cooperation (e.g., turn-taking, sharing, being polite) • Produce a product as a group • Present as a group
Motivation	Level of desire or willingness to do something	<ul style="list-style-type: none"> • Setting goals • Creating a plan to reach a goal • Self-assessing success • Demonstrating persistence

2. Read a rabbit care guidebook that contains information about real rabbits. Pretend that your class will be getting a rabbit to take care of as a class pet. Prepare a shopping list of things you will need to take care of the rabbit. Design and “build” paper airplanes, experimenting with different designs and use of weights (paper clips). Fly the planes with the goal of distance or time aloft.
3. Using a restaurant menu, pick out six meals for you and your friends. Calculate the cost and keep within a budget.

Authentic Assessment of Specific Skills

Common skills for which authentic assessment is especially suited include procedural knowledge, problem solving, collaboration, and motivation (Borich & Tombari, 2004). Table 8.2 provides examples of tasks that allow for observation of these skills. What makes these examples authentic is that they are part of a teaching and assessment plan with the goal of increasing generalizable skills, as opposed to a goal of high test performance. They focus on complex abilities, not low-level knowledge.

Computers and Authentic Assessment

Technology

It's likely no coincidence that the 20-some-year movement toward assessment designed to promote and measure complex thinking skills has grown parallel to the rise of computer-based educational technology. Computers and the web provide great opportunities for authentic assessment (Chang & Tseng, 2008) through, among other avenues, web-based collaboration (Chiu, Yang, Liang, & Chen, 2010; Donnan, McCormack, Battye, & Hart, 2008; Hron & Friedrich, 2003), computer-based (or supported) assessment (Laurier, 2000), and the use of multimedia (Neo, Neo, & Tan, 2011).

Web-Based Collaboration

Collaborative learning is “a joint construction of meaning through interaction with others” (Hron & Friedrich, 2003, p. 71). Authentic skills that can be assessed through collaborative educational activities include deduction, induction, synthesis, investigation, and a variety of social skills. Structured online interactions, whether local or outside the classroom, can allow for assessment through the role of moderator—monitoring the social exchange, providing predetermined discussion topics, providing formative feedback, and strategically controlling participation.

(Continued)

(Continued)

Computer-Based Assessment

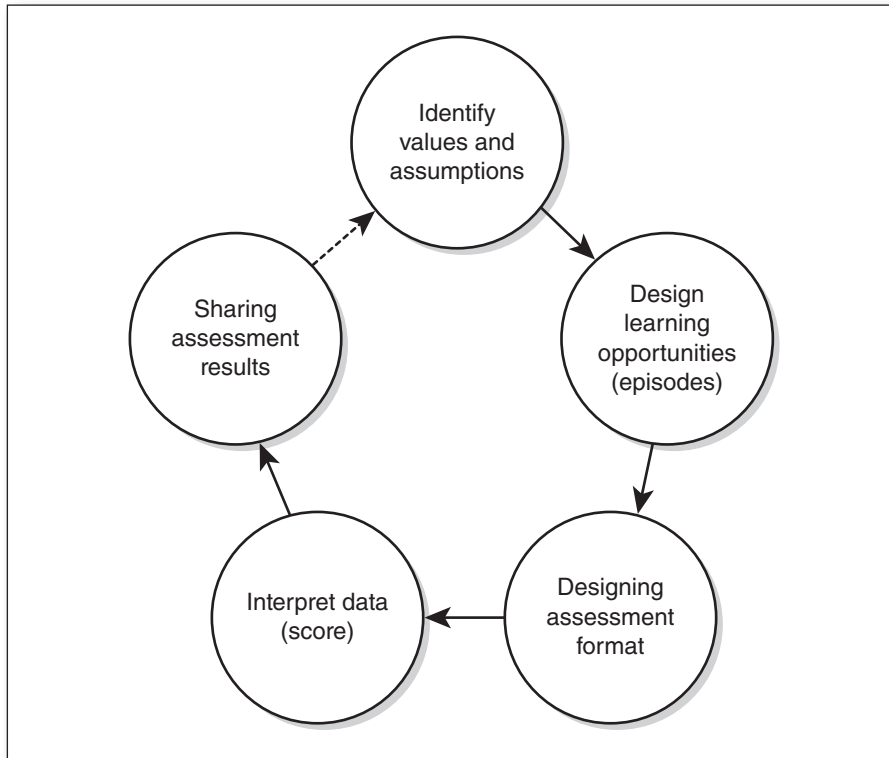
Standardized testing is now substantially administered and scored online or on computer. The authentic assessment movement is alive and well in large-scale standardized assessment, as well, but it is still an open question how well the principles of authenticity can translate to standardized tests. Some powerful elements of authentic assessment (e.g., local, student involvement in development, public defense, collaboration) might be impossible (or, at best, very difficult) for large scale assessment. The computer, though, can be a useful tool in authentic assessment. Computers can be used for assessment of those tasks that in the real world occur on computers (e.g., literacy tasks, such as chatting or discussions and writing), for efficiency (if computers allow for quicker assessment, then a greater number of assessments can be administered and included in portfolio-type systems) and, of course, to test computer skills themselves, which are often objectives in the modern classroom (Laurier, 2000).

Multimedia Authentic Assessment

Though computer-based testing and authenticity have yet to fully embrace each other, computer technology is great at providing interactive multimedia assessment environments. Videos, sound, animations, games, graphs, journals, and simulated documents can all be integrated into a coordinated environment of learning, exploration, and problem-solving tasks. This supports authentic assessment, especially in terms of providing a multidimensional, complex context, student participation and motivation, multiple indicators of performance, and fairly unstructured challenges (Herrington & Herrington, 1998).

Authentic Assessment of Literacy

The teaching and assessment of reading, writing, listening, and speaking is foremost among the goals of modern education. Though there are standardized tests to assess these skills, they do not assess the high levels of literacy needed to fully engage in today's world (Koda, 2005; Ratcliff, 2002). Consequently, authentic assessment of reading and other literacy skills is a hot topic that has received a great deal of notice in teacher education. The analysis of important aspects of authenticity discussed at the start of this chapter found that among language arts researchers (e.g., Hirvela & Pierson, 2000; Laurier, 2000), the important components of assessment were a match between language tasks, such as reading, and the way these skills are applied outside the classroom and the use of multiple indicators of skill, such as portfolios (Chang & Tseng, 2008).

Figure 8.3 Authentic Literacy Assessment Process

Hancock, Turbill, and Cambourne (1994), a team of Australian literacy researchers, developed authentic classroom assessment activities while observing real-world teachers. They described their work some years ago, but their insights and suggestions are modern. The authors suggest a process by which authentic assessment can be developed, evaluated, and shared. As shown in Figure 8.3, the assessment process begins with identifying one's beliefs and assumptions about the skills to assess (e.g., *what is literacy? how is it learned?*) and ends with the sharing of assessment results with students and parents.

Among the educators observed was an experienced 6th grade teacher who provided an example of how he filled the 2-hour “literacy block” in his classroom with authentic assessment. Literacy blocks are very common today in U.S. schools, and his approach is a good model for the multifaceted structure

Table 8.3 Authentic Assessment as Part of a 2-Hour Literacy Block in a 6th Grade Classroom

Segment	Time	Purpose	Assessment
Teacher reading	15 mins	Shows reading is fun Demonstrates fluent reading Encourages students to ask questions	Observation Learner logs
Sustained silent reading	20 mins	Shows reading is rewarding Students responsible for choices Uninterrupted period to connect to text	Observation Reading logs
Modeling	10 mins	Demonstrates assigned tasks Models strategies	Conferences Interviews Surveys
“Workshop” tasks	60 mins	Working in groups Opportunity to listen Opportunity to clarify ideas Using language to communicate ideas	Observations Retellings
Discussion and sharing	15 mins	Identifies what learning has occurred Identifies common difficulties in learning Guides students to respect others	Learning logs

Source: Adapted from Hancock, Turbill, & Cambourne, 1994.

of authentic assessment and what an authentic literacy assessment environment might look like in the modern classroom. The different “episodes” (segments) in a typical literacy block in his classroom are presented in Table 8.3.

Authentic Assessment in Mathematics

It is more difficult to identify authentic tasks in math than it is in many other school subject areas because there seems to be a disconnect between the common mathematics activities in classrooms and the tasks that real-world mathematicians (or other professionals who use math in their work) engage in (Goos, Stillman, & Vale, 2012; Lajoie, 1995). Inside the classroom, solving a “math problem” is often a solitary activity with an emphasis on mental calculation; outside the classroom, tools and technology are often available and the process may take place in a collaborative group setting. (See Chapter 7 for a scoring rubric based on behaviors of real-world mathematicians.) Consequently, some math education researchers suggest that *authenticity* be defined a bit differently for this subject matter. Instead of focusing on how

math is practiced in the real world, those mathematical skills that are *valued* in the outside world might be a more reasonable focus.

Worthwhile mathematical tasks, those that are valued outside the classroom, include the following:

- **Problem-solving.** Challenging cognitive processes include gathering and discovering knowledge and analyzing numbers. Authenticity can be increased by using real data and assigning interesting problems.
- **Communication.** Reading about, writing about, and discussing mathematical ideas. Activities that encourage reflection, synthesis, collaboration, and a “public” defense support authenticity.
- **Reasoning.** Understanding the structure and assumptions of a math problem and applying deductive and inductive thought to reach a solution. Authenticity is strengthened when problems are complex and ill-defined and are experienced across a variety of contexts.
- **Making Connections.** A curriculum that incorporates math tasks and assessment across topics encourages students to generalize and transfer their skills instead of treating each task or application in isolation. An authentic assessment environment incorporates assessment of all these math skills together throughout the entire mathematics curriculum (Lajoie, 1995).

Because mathematical skill assessments that are meant to be authentic do not always match the way mathematical tasks are completed outside the classroom, some argue that a balance can be found between opportunities to simulate workplace practices (the realistic aspect) and opportunities for students “to act like children learning mathematics” (Moschkovich, 1998, p. 16).

Authentic Assessment in Visual Arts

“Artwork created by students are objects of meaning that reflect artistic valuing and aesthetic intents that provide sensory perception and appreciation because they involve elements of human motivation and interactions between the student and his or her environment. . . . (They) contribute to the enrichment of conscious life experiences through providing meaning on a symbolic level and affectively through feelings that contribute to the enrichment of sensory competence and cognitive enrichment.”

Dorn, Madeja, and Sabol (2004, p. 1)

In the current budget-focused climate, art educators often face pressure to justify their content area as representing a set of authentic and necessary real-world skills. As the quotation that opened this section demonstrates, however, visual arts assessment has the potential to be authentic and realistic because it reflects authentic human experience.

Dorn et al. (2004) were involved in large-scale attempts to develop art instruction and authentic assessment in schools in Florida, Indiana, and Illinois. When developing authentic tasks for use in the projects' assessment, they followed a set of guidelines (p. 100) for maintaining the validity of the process:

1. Identify both the procedural and focal knowledge of the students needed for them to know how and be able to do various learning activities in the arts.
2. Identify the core performance roles or situations that all pre-K–12 students should encounter and be expected to master.
3. Choose the most salient discriminators that can be used in evaluating performance.
4. Design tasks with sufficient depth and breadth to assess competence.
5. Train evaluators to (among other things) reach high levels of inter-rater reliability.
6. Apply a clear understanding of the intended audience.

In their description of scoring rubrics for the authentic assessment of art assignments, they emphasize several points, all consistent with the criteria for quality rubrics presented in Chapters 6 and 7. The rubric should be written in ordinary language and be based on directly observable characteristics of the product or performance, and the criteria should be based on the critical demands of performance.

Posters

An authentic assessment format that can be used for any content area, not just art, is the poster. Teachers sometimes choose the "poster" as an assessment format for a variety of contexts. MacAndrew and Edwards (2002) describe the characteristics of a poster as an assessment, compare it to the more traditional *essay*, and provide suggestions for a valid scoring

rubric. Posters for this purpose are similar to those one sees at professional conferences (present a question, describe the methods of research, and share and interpret results). In format, poster sessions have much in common with the traditional science fair (an example of authentic assessment that has been around long before modern approaches to assessment). The posters are displayed at a “conference” gathering of students and teachers and, often, parents and staff. Students stand in front of the poster and answer questions about it and the work they have done.

To score the posters, a rubric should include assessment criteria developed collaboratively by students and teachers, but reasonable criteria include

- Title
- Clear goals
- Well organized
- Easy to read
- Effective use of graphics
- Research quality (appropriate sources cited, evidence of reading in the field, methodology is sound, interpretation of results is correct)

Students and other stakeholders (parents, administrators) tend to be comfortable with essays because they are used to them. It is important to train students as to the nature of a research poster and the characteristics of quality before expecting high levels of performance.

AUTHENTIC ASSESSMENT FOR ENGLISH LANGUAGE LEARNERS

There has been quite a bit of study and scholarly reflection on authentic assessment of English language learners (ELL) and students who use English as a second language (ESL). Wise suggestions for developing authentic assessments for this population can be found in books by Ekbatani and Pierson (2000), who are interested in the self-assessment and formative aspects of authenticity, and O’Malley and Pierce (1996), whose work focuses on practical advice for authentic classroom assessment. The suggestions that follow are found in these books and the work of others (Ekbatani, 2008; Shibliyev & Gilanlioglu, 2009; Toscano, 2009) and are meant for evaluating English language learning.

One approach is to turn the classroom into a writing environment, not a grading environment (Gearhart, 2009; Hirvela & Pierson, 2000), and to establish portfolio assessment that is maintained and self-assessed by students. To allow for a developmental picture to form, the portfolios should be sequential, produced over a defined period of time, with assignments focused on a particular context or purpose. A key in owning one's own learning is to make the decisions about what to include in the portfolio; the greater the role played by the student in designing the portfolio, the more realistic the assessment becomes. North (2000) provides a template for a generic scoring scale developed by the educational advisory group *Council of Europe*, which could be used for authentic assessment rubrics for ELL or ESL students. The descriptions of the score points make use of many of the concepts that we have come to associate with authentic assessment, such as *complex*, *encountered in work*, *relevance*, and *demanding*. A condensed version of that scale, with point values added, is below.

0. Breakthrough

Can understand and use familiar everyday expressions and very basic phrases. Can interact in a simple way.

1. Waystage

Can understand sentences and frequently used expressions related to the areas of most immediate relevance.

2. Threshold

Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, and leisure.

3. Vantage

Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions within domains of interest.

4. Effective Operational Proficiency

Can understand a wide range of demanding, longer texts, and recognize implicit meaning.

5. Mastery

Can understand with ease virtually everything heard or read.

O'Malley and Pierce (1996) present detailed descriptions of many authentic assessable classroom tasks. They observed teachers' real-life efforts at applying

Table 8.4 Examples of ESL Authentic Assessments

Level	Subject	Task	Description
Elementary	Language Arts	Responding to Reading	Student reads to whole class and students give feedback.
Elementary	Language Arts	Anecdotal Records	Teacher takes notes while observing individual reading.
Elementary	Language Arts	Book Talks	Students present personal responses to readings and answer other students' questions.
Middle	Mathematics	Geoboard	Geometric concepts are assessed.
Middle	Science	Magnet Experiment	Students are observed while comparing magnets to each other.
Secondary	Language Arts	Talk Show	Students simulate a TV talk show and interview each other as if they are famous.
Secondary	ESL	Interpreting Portfolios	Two teachers examine portfolios and discuss their evaluations.

Source: Based on O'Malley & Pierce, 1996.

authentic assessment with ELL and ESL students and interviewed many others about their techniques. Among the authentic elements of the tasks they describe are organized use of multiple indicators (such as portfolios) and student relevance (students' individual reactions to text or communication activities are central). Table 8.4 (based on a figure in their book) gives examples. Some examples are for teaching language; others are for teaching and assessing other subjects for ESL students.

Simulation Computer Games

Technology

There are several useful websites that provide free and simple games that simulate real-life problem-solving situations for children. Similar games are very popular on Facebook and other social networking sites. Though not entirely realistic, performance on these games could be used as part of authentic assessment strategies or provide ideas for in-class simulations. Of course, screen these games thoroughly to make sure they meet

(Continued)

(Continued)

your standards regarding age appropriateness, Internet use, and commercialization. (All these sites have advertising and hope to sell things.) Some games can be downloaded so they could be used on classroom computers without being connected to the web. <http://www.youdagames.com/>

Youda Games. Several fun and somewhat complex simulation games are available here. Some good possibilities are *Virtual Farm*, *Goodgame Farmer*, *Virtual City*, *Hotel Mogul*, and *Roads of Rome*. <http://www.123-games.net/>

123 Games. Some richly detailed business simulation games at this site include *Cookie Tycoon*, *Shop Empire*, and *Corporation, Inc.* and a variety of “lemonade stand” type games such as *Hot Dog Stand* and *Pizza King*. Many of these games emphasize money decisions to a much lesser degree than the “move quickly and make that taco” aspect, but there are still good possibilities here. dan-ball.jp/en/

Dan-ball. Science and physics-based simulations are available here. Fun and educational games include *Planet Simulation*, where you must design solar systems; *Earth Editor*, where you develop the planet; and *Elemental Box*, where various objects react to each other using real physics. These are “sandbox” games where students can explore and try things just to see what happens.

AUTHENTIC ASSESSMENT AND YOUNG CHILDREN

Thinking of authentic assessment as a strategy to increase job-related skills or develop the abilities that employers are looking for doesn't have much meaning when one is assessing a 3-year-old. It is useful to view authenticity somewhat differently when teaching preschool children or students in the early elementary grades. Here the emphasis should be on the developmental skills supportive of future learning. One can think of a young child's “job” as growing, learning, and being a student.

Puckett and Black (2008), in their book on assessing young children, define authentic assessment for this population as obtaining information that “truly reflects how a child pursues knowledge and skills and the outcomes of the child's efforts. [Assessment should be] teacher-mediated, child-centered . . . and based on multiple theories and knowledge about child growth and development” (p. 75). While one concern that teachers may have is that young children cannot meaningfully take part in the collaborative aspect of authentic assessment, they

argue that empathic interactions are possible with the teacher encouraging children's natural eagerness to learn and prove how "smart" they are. Ways to increase the authenticity of assessments with young children include the following:

- The teacher assumes the role of mentor, with the child as novice.
- Student and teacher share experiences from outside the classroom.
- The teacher incorporates children's interests into classroom activities and assessments.
- Learning is understood as being partly social and emotional.

AUTHENTICITY AT THE SCHOOL LEVEL

"Do we want to evaluate student problem-posing and problem-solving in mathematics? Experimental research in science? Speaking, listening, and facilitating a discussion? Doing document-based historical inquiry? Thoroughly revising a piece of imaginative writing until it 'works' for the reader? Then let our assessment be built out of such exemplary intellectual challenges."

Grant Wiggins (1990, p. 1)

Darling-Hammond, Aness, and Falk (1995; Darling-Hammond, 2012) studied real-life examples of schools that had implemented systems of authentic assessment. For their analysis, they used a framework of Wiggins's (1989) and applied four observable characteristics that distinguish authentic assessment in practice from other approaches:

1. Assessment tasks are representative of the "field." Students actually write and conduct experiments rather than taking spelling tests and recalling science facts.
2. Carefully designed standards of performance evaluate the essential qualities of performance. These aren't secret, they are shared with students and guide instruction.
3. Students play a role in evaluating their own work. Real-world contexts require that people self-assess and self-motivate to be successful, and authentic assessment aims to develop those skills.
4. Students frequently present their work "publicly." This requires that they reflect on their work and what they know and share it in an understandable way.

Three case studies of New York schools presented by the authors, in particular, provide useful models for what various aspects of authentic assessment—portfolios, projects, and collaboration with students—look like at the school level.

Central Park East Secondary School structured student performance expectations around a portfolio approach. The portfolios contain work samples—writings, math papers, and, principally, projects—which are meant to demonstrate independent reasoning and action and “habits of mind” (a concept that is emphasized across many perspectives of authentic assessment) which encourage the weighing of evidence, awareness of different viewpoints, seeing connections, speculating on possibilities, and assessing value. Portions of the portfolios are presented and “defended” by students to a committee of faculty.

Hodgson Vocational Technical High School responded to their district’s drive for higher expectations for vocational students in the areas of mathematics, literacy, and science. As a first step toward the goal of a diploma based on performance, a three-part Senior Project was instituted. The three parts are

1. A research paper. This paper is “shop”-based; students at this school are training to be carpenters and such.
2. A shop project. This is to be a large, complex project that students design and build themselves. Their research paper is meant to support the production of the project.
3. A public, formal oral presentation. Teachers, parents, and other students attend.

At the time that Darling-Hammond and colleagues visited Hodgson, the evaluations of the project’s pieces were conducted separately by the different departments (e.g., the English faculty evaluated the research paper and the presentation). An interdisciplinary approach, however, would have been more consistent with authentic assessment principles.

Another case study described concerns International High School, which emphasized collaborative learning. International High School in New York City is exclusively for recent immigrants with a relatively low English proficiency. A community learning and assessment environment was established supported by a core belief that students learn from each other’s different experiences and knowledge. Authentic assessment is operationalized as a three-pronged approach—self-assessment, peer assessment, and teacher assessment. These three sources of assessment apply to an ongoing series of formative, performance-based assessments designed to provide meaningful feedback, summative assessments that are multidimensional, and collaborative broad evaluations from multiple perspectives. Group work and working in pairs is common, and much of the assessment occurs in these contexts.

Figure 8.4 Authenticity of Assessment Scoring Rubric

	High				Low
Context of the Assessment					
Realistic activity or context The task and methods of evaluation are similar to what would be required or expected in the real world, outside of an artificial classroom environment.	20	15	10	5	0
Cognitively complex Successful performance of the task requires high levels of understanding or critical thinking.	10	8	5	2	0
Performance-based Skill or ability (as opposed to knowledge) is assessed through a performance or creation of a product.	5	4	3	2	0
Role of the Student					
Formative assessment The assessment is designed to provide feedback to students to control their own learning. Scores do not affect grades.	10	8	5	2	0
Collaborative Students work with each other or with the teacher during the task, to evaluate their performance and, perhaps, to create the assessment.	10	8	5	2	0
Defense is required Students defend their “answers” or performance. This might be a formal, oral defense in front of students and adults or a written defense as part of the assessment.	5	4	3	2	0
Scoring Procedures					
Multiple indicators or portfolio The “score” on the assessment is a composite of multiple scores reflecting the quality of multiple components or a portfolio of products and student work.	20	15	10	5	0
Criteria known by students Scoring rules are well understood by students or they participated in their creation. Teachers may have used these criteria as part of their instruction.	15	12	8	4	0
Mastery expectation The task and scoring are designed to provide feedback on whether the student has mastered a skill or ability (as opposed to comparing the student with other students).	5	4	3	2	0
Total					

From Frey, B. B., Schmitt, V. L., & Allen, J. P. (2009, November). *Assessing authentic assessment*. Presented at the Annual Meeting of the American Evaluation Association, Orlando, FL.

ASSESSING THE ASSESSMENT

The different definitional dimensions of authenticity discussed in this chapter can be organized into a scoring rubric to evaluate the degree of authenticity within a teacher-made assessment (Frey et al., 2009). To find out “how authentic is it?” with your own assessments or others’, apply this set of scoring guidelines based on those nine components. The different point values are rough approximations of the relative frequency that these components were offered as critically necessary for assessments to be authentic. For the publications dealing with authenticity for school-aged students, the relative frequencies of the dimensions required as part of the definition of authentic assessment were *assessment task mirrors reality outside of the classroom*, 60%; *multiple indicators or portfolios for scoring*, 54%; *known or student developed criteria*, 47%; *formative assessment*, 31%; *cognitively complex*, 30%; *performance-based*, 23%; *collaborative*, 20%; *a required defense*, 15%; and *mastery expectation*, 13%.

Mr. Hernandez Gets Real (Part II)

Stories From the Classroom

Mr. Hernandez took an Elementary Statistics course during college and had covered a bit in his classroom assessment course, too. He actually used data analysis in his everyday teaching and knew of multiple applications of mean, median, mode, and range and other ways of understanding data. He wanted to show his students a real-world application of the data analysis tools that would engage them more fully. He planned a lesson around the data analysis he used as part of his job and created a formative quiz. The quiz wouldn’t affect students’ grades, but it would create some real-world data that he could use in his teaching.

When the students had completed the test (after Mr. Hernandez assured them that this was just to create some data to play with and they wouldn’t need their names on it), he mixed them up and had the students score them as he read out the correct answers. The students then converted the raw score to a percent correct score. Mr. Hernandez went to the board and had the students tell them all the percent scores. He listed them all on the board.

“What does this information tell me?” asked Mr. Hernandez. Shiloh raised her hand. “That we do not know very much?” After the laughter, Mr. Hernandez asked, “How did the class do overall? If I needed one score to tell me, which should I use?”

Shiloh responded, “Mean?”

“Okay class, find the mean, median, mode, and range of our data set and then tell me what decision I should make based on that information.” After a few minutes of computation, Hunter raised his hand and said, “I don’t think that in real life you should use your normal grading rules for this sort of quiz.”

"Interesting—why not?"

"The mean score is 68, so most of the class would get, like, a D!" The class agreed. Mr. Hernandez asked, "Are the median and mode higher, the same, or lower than the mean?" The class agreed it was much lower. "Why, and what does that tell me about the test?"

The class agreed that it was because one test score was really high and made the mean higher. Mr. Hernandez explained that many decisions in the real world, many real-world jobs, used descriptive statistics like these to make decisions all the time. That included teachers. He then asked the students what other types of jobs might also use statistics to make decisions and how those jobs might use them. There were a lot of good suggestions, which led perfectly into the final assessment.

"For our end of unit assessment I am going to have you gather, organize, analyze, and display some data. You will need to come up with two data collection questions, one that counts frequencies in terms of a few categories and one that uses scores on some measure like our quiz. Think of questions that different occupations might really want to know the answers to."

Sergio wanted to know if he could do favorite pizza toppings for his categorical question. "Who would want to know this information?" asked Mr. Hernandez. "Pizza Hut or maybe the lunch lady," Sergio said.

On the day the project was due, Mr. Hernandez got enough butcher paper for all the students to hang their completed projects. Students were allowed to discuss their results with other classes that came by to see the project. Even Riel said he enjoyed the assignment. Mr. Hernandez knew that it made a difference to Riel that the work he did was the kind of work real people did in the real world. And that made a difference. Really.

THINGS TO THINK ABOUT

1. Think about the word "realistic." What would a *realistic* classroom assessment actually be like?
2. When would it be important to design authentic assessments in the classroom? When would it not be as important?
3. How does *authenticity* in assessment change as the context changes?
4. Which aspects of authenticity (those nine dimensions) seem to you to be most important?
5. What would a school dedicated to authentic assessment principles look like?

Looking Back in This Chapter

Applying authentic assessment procedures in the real classroom can be challenging. Though the community of classroom assessment scholars agrees that authentic assessments are potentially a powerful, transformative tool, there is not yet agreement on which aspects of authenticity are most important. Nine different dimensions of authenticity were presented, with the realistic nature of the assessment being emphasized. The overlap between the modern assessment approaches of authentic, formative, and performance-based was discussed. Validity and reliability issues of authentic assessment include determining what is authentic and the necessarily subjective nature of the scoring. General strategies for operationalizing authenticity and illustrations of authentic assessment tasks and scoring methods were provided for literacy and reading, mathematics, visual arts, and other content areas and purposes. A scoring rubric for determining *how authentic a particular assessment is* was provided.

ON THE WEB

Selection of authentic task suggestions across many areas

<http://jfmuller.faculty.noctrl.edu/toolbox/examples/authentictaskexamples.htm>

Authentic assessment resources for classroom teachers

<http://www.uwstout.edu/soe/profdev/assessment.cfm>

Overview of authentic assessment with linked supports

<http://www.teachervision.fen.com/teaching-methods-and-management/educational-testing/4911.html>

Authentic assessment and multiple intelligences

<http://teachersnetwork.org/teachnet-lab/mbhs/scragg/multiple.html>

STUDENT STUDY SITE

Visit www.sagepub.com/frey to access additional study tools including eFlashcards, web quizzes, web resources, additional rubrics, and links to SAGE journal articles.

REFERENCES

- Abernethie, L. (2006). Authentic assessment as a neo-liberal technology of government. *Teaching and Learning Forum*, 1–9.
- Archbald, D. A. (1991). Authentic assessment: Principles, practices, and issues. *School Psychology Quarterly*, 6(4), 279–293.
- Archbald, D. A., & Newmann, F. M. (1988). *Beyond standardized testing: Assessing authentic academic achievement in the secondary school*. Reston, VA: National Association of Secondary School Principals.
- Baron, M. A., & Boschec, F. (1995). *Authentic assessment: The key to unlocking student success*. Lancaster, PA: Technomic.
- Borich, G. D., & Tombari, M. L. (2004). *Educational assessment for the elementary and middle school classroom*. Englewood Cliffs, NJ: Prentice Hall.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18, 1, 32–42.
- Burke, K. (2009). *How to assess authentic learning*. Thousand Oaks, CA: Corwin.
- Chang, C. C., & Tseng, K. H. (2008). Use and performances of web-based portfolio assessment. *British Journal of Educational Technology*, 40(2), 358–370.
- Chiu, C. H., Yang, H. Y., Liang, T. H., & Chen, H. P. (2010). Elementary students' participation style in synchronous online communication and collaboration. *Behaviour & Information Technology*, 29(6), 571–586.
- Darling-Hammond, L. (2012). *Powerful teacher education: Lessons from exemplary programs*. San Francisco, CA: Jossey-Bass.
- Darling-Hammond, L., Aness, J. A., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work*. New York, NY: Teachers College Press.
- Dez, M., Moon, J., & Meyer, C. (1992). What do we want students to know? . . . and other important questions. *Educational Leadership*, 49(8), 38–42.
- Donnan, P., McCormack, C., Battye, G., & Hart, I. (2008, July). *Doing group assessment: A web-based resource of good practice case-studies*. Paper presented at the meeting of the 31st Annual HERDSA Conference, Rotorua, New Zealand.
- Dorn, C. M., Madeja, S. S., & Sabol, F. R. (2004). *Assessing expressive learning*. Mahwah, NJ: Lawrence Erlbaum.
- Dutt-Doner, K., & Maddox, R. (1998). Implementing authentic assessment. *Kappa Delta Pi Record*, 34(4), 135–137.
- Ekbatani, G. (2008). *Measurement and evaluation in post-secondary ESL*. New York, NY: Routledge.
- Ekbatani, G., & Pierson, H. (Eds.). (2000). *Learner-directed assessment in ESL*. Mahwah, NJ: Lawrence Erlbaum.
- Engel, M., Pulley, R., & Rybinski, A. (2003). *Authentic assessment: It really works* (Master's thesis). Available from ERIC database. (ERIC ED479959)
- Fischer, C., & King, R. M. (1995). *Authentic assessment: A guide to implementation*. Thousand Oaks, CA: Corwin.
- French, D. (2003). A new vision of authentic assessment to overcome the flaws in high stakes testing. *Middle School Journal*, 35(1), 14–23.

- Frey, B. B., Schmitt, V. L., & Allen, J. P. (2009, November). *Assessing authentic assessment*. Presented at the annual Meeting of the American Evaluation Association, Orlando, FL.
- Gearhart, M. (2009). Classroom portfolio assessment for writing. In G. A. Troia (Ed.), *Instruction and assessment for struggling writers: Evidence-based practices* (pp. 311–336). New York, NY: Guilford Press.
- Gipps, C. (1995). Reliability, validity, and manageability in large-scale performance assessment. In H. Torrance (Ed.), *Evaluating authentic assessment*. Buckingham, England: Open University Press.
- Goos, M., Stillman, G., & Vale, C. (2012). *Teaching secondary school mathematics: Research and practice for the 21st century*. Crows Nest NSW, Australia: Allen & Unwin.
- Green, J. (1998). Authentic assessment: Constructing the way forward for all students. *Education Canada*, 38(3), 8–12.
- Gronlund, G. (2003). *Focused early learning: A planning framework for teaching young children*. St. Paul, MN: Redleaf Press.
- Gulikers, J. T. M. (2006). *Authenticity is in the eye of the beholder: Beliefs and perceptions of authentic assessment and the impact on student learning*. Maastricht, The Netherlands: Datawyse.
- Hancock, J., Turbill, J., & Cambourne, B. (1994). Assessment and evaluation of literacy learning. *Authentic Reading Assessment: Practices and Possibilities*, 46–62.
- Herrington, J., & Herrington, A. (1998). Authentic assessment and multimedia: How university students respond to a model of authentic assessment. *Higher Education Research and Development*, 17(3), 305–322.
- Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development*, 48(3), 23–48.
- Hirvela, A., & Pierson, H. (2000). Portfolios: Vehicles for authentic self-assessment. In G. Ekbatani & H. Pierson (Eds.), *Learner-directed assessment in ESL* (pp. 105–126). Mahwah, NJ: Lawrence Erlbaum.
- Hron, A., & Friedrich, H. F. (2003). A review of web-based collaborative learning: Factors beyond technology. *Journal of Computer Assisted Learning*, 19(1), 70–79.
- Jolly, J., & Kettler, T. (2000). Authentic assessment of leadership in problem-solving groups. *Gifted Child Today*, 27(1), 32–39.
- Kellaghan, T., & Madaus, G. F. (1993). The British experience with “authentic” testing. *Phi Delta Kappan*, 74(6), 458–466.
- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. United Kingdom: Cambridge University Press.
- Lajoie, S. P. (1995). A framework for authentic assessment in mathematics. In T. A. Romberg (Ed.), *Reform in school mathematics and authentic assessment*. Albany: State University of New York Press.
- Laurier, M. (2000). Can computerized testing by authentic? *ReCALL*, 12(1), 93–104.
- Lawton, D. (2000). Authentic assessment in South Brunswick schools. *Knowledge Quest*, 29(2), 30–38.
- Lindsay, A. D. (1991). *The Republic: The complete and unabridged Jowett translation*. New York, NY: Vintage Books.
- MacAndrew, S. B., & Edwards, K. (2002). Essays are not the only way: A case report on the benefits of authentic assessment. *Psychology Learning & Teaching*, 2(2), 134–139.
- Maden, J., & Taylor, M. (2001, February 27–March 3). *Developing and implementing authentic oral assessment instruments*. Paper presented at the Annual Meeting of Teachers of English to Speakers of Other Languages, St. Louis, MO.

- Meisels, S. (1996). Using work sampling in authentic assessment. *Educational Leadership*, 60–64.
- Meyer, C. (1992). What's the difference between authentic and performance assessment? *Educational Leadership*, 49(8), 39–40.
- Meyers, N. M., & Nulty, D. D. (2009). How to use (five) curriculum design principles to align authentic learning environments, assessment, students' approaches to thinking and learning outcomes. *Assessment & Evaluation in Higher Education*, 34(5), 565–577.
- Montgomery, K. (2001). *Authentic assessment: A guide for elementary teachers*. New York, NY: Longman.
- Moorcroft, T., Desmarais, K., Hogan, K., & Bekowitz, A. (2000). Authentic assessment in the informal setting: How it can work for you. *Journal of Environmental Education*, 31(3), 20–24.
- Moschkovich, J. N. (1998). Rethinking authentic assessments of student mathematical activity. *Focus on Learning Problems in Mathematics*, 20(4) 1–18.
- Mueller, J. (2005). Authentic assessment in the classroom and the library media center. *Library Media Connection*, 23(7), 14–18.
- Neo, M., Neo, K. T. K., & Tan, H. Y. J. (2011, June). Content restructuring with authentic learning strategies in a multimedia learning environment (MMLE). In T. Bastiaens & M. Ebner (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2011* (pp. 2233–2242). Chesapeake, VA: AACE.
- Newmann, F., Brandt, R., & Wiggins, G. (1998). An exchange of views on semantics, psychometrics, and assessment reform: A close look at “authentic” assessments. *Educational Researcher*, 27(6), 19–22.
- North, B. (2000). Defining a flexible common measurement scale: Descriptors for self and teacher assessment. In G. Ekbatani & H. Pierson (Eds.), *Learner directed assessment in ESL*. Mahwah, NJ: Lawrence Erlbaum.
- O'Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English language learners*. Saddle River, NJ: Pearson Education.
- Paris, S. G., & Ayres, L. R. (1994). *Becoming reflective students and teachers with portfolios and authentic assessment*. ERIC Reproductive Services No. 378166.
- Puckett, M. B., & Black, J. K. (2008). *Meaningful assessments of the young child* (3rd ed.). Saddle River, NJ: Pearson Education.
- Ratcliff, N. J. (2002). Using authentic assessment to document the emerging literacy skills of young children. *Childhood Education*, 78(2), 66–69.
- Schnitzer, S. (1993). Designing an authentic assessment. *Educational Leadership*, 50(7), 32–35.
- Shibliyev, J., & Gilanlioglu, I. (2009). Language testing and assessment: An advanced resource book. *ELT Journal*, 63(2), 181–183.
- Spinelli, C. G. (1998). *Teacher education reform: Promoting interactive teaching strategies and authentic assessment for instructing an increasing diverse population of students*. West Long Branch, NJ: 120 Opinion Papers.
- Stripling, B. (1993). Practicing authentic assessment in the school library. *School Library Media Annual (SLMA)*, 11, 40–57.
- Suen, H. K., Sonak, B., Zimmaro, D., & Roberts, D. (1997). Concept map as scaffolding for authentic assessment. *Psychological Reports*, 83, 734.
- Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 18(4), 433–449.
- Tombari, M., & Borich, G. (1999). *Authentic assessment in the classroom: Applications and practice*. Upper Saddle River, NJ: Prentice Hall.

- Torrance, H. (1993). Combining measurement-driven instruction with authentic assessment: Some initial observations of national assessment in England and Wales. *Educational Evaluation and Policy Analysis*, 15(1), 81–90.
- Torrance, H. (1995). Teacher involvement in new approaches to assessment. In H. Torrance (Ed.), *Evaluating authentic assessment*. Buckingham, England: Open University Press.
- Torrance, H. (2009). 4.2 Using assessment in education reform. In H. Daniels, H. Lauder, & J. Porter (Eds.), *Knowledge, values and educational policy: A critical perspective* (pp. 218–236). New York, NY: Routledge.
- Toscano, L. M. (2009). ELL authentic assessment defined. *Assessment*. Retrieved from <http://www.ibrarian.org/assesstask1.pdf>
- Wiggins, G. (1989). Teaching to the (authentic test). *Educational Leadership*, 46(7), 41–47.
- Wiggins, G. (1990). The case for authentic assessment. ERIC Digest. ERIC Document Reproduction Service No. ED, 328 611.
- Wilson, J., & Schwier, R. (2012, March). A model of authentic learning processes in instructional design. In P. Resta (Ed.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2012* (pp. 1285–1289). Chesapeake, VA: AACE.