

other contexts. Using both types of criteria to assess students in any one of the three tasks about the Civil War should show us whether students are seeing the period in historical perspective, as well as making links to their own experience.

### How Can We Assure Expert Judgments?

The ability to use the criteria to determine the quality of students' work is what we mean by expert judgment, and it is far from the subjective process some fear. Because the criteria are known in advance to the students (and teachers use them to design learning experiences that lead up to the appraisal), assessment becomes a matter of gathering evidence in the student's performance to support a judgment whether each criterion is met.

Some criteria we might identify to assess students' ability to perform any of the three tasks about the Civil War noted above might include:

1. *Accurately uses information from the historical period (no evidence of anachronisms).* When students use information to create a picture of life in a specific historical period, the teacher can see the depth of their understanding; obviously, the presence of 20th century devices (televisions, fax machines), for example, would reveal problems in a student's grasp of the period.

2. *Uses sufficient detail to create a sense of what it was like for people who lived at the time under study.* This criterion calls upon the teacher's and students' sense of "how much is enough?" The teacher should talk through the need for the performance to satisfy an audience's need: Who will read what the student produces? What context-setting information will the audience need? How much description and how many examples

## What's the Difference Between Authentic and Performance Assessment?

CAROL A. MEYER

Performance assessment and authentic assessment are often used interchangeably, but do they mean the same thing? Although both labels might appropriately apply to some types of assessment, they are not synonymous. We must be clear about the differences if we are to support each other in developing improved assessments.

### Two Examples

To distinguish between the two terms, let's look at a familiar form of assessment with which we have a wealth of experience. Following are two examples of a direct writing assessment in which students produce writing samples.

*Case 1:* Every May school district X conducts a direct writing assessment. For four days, all students at selected grade levels participate in a standardized series of activities to produce their writing samples. Using a carefully scripted manual, teachers guide students through the assessment with limited teacher directions and extended student writing time (up to 45 minutes) each day: Topic Introduction and Pre-writing (Day 1), Rough Drafting (Day 2), Revising and Editing (Day 3), and Final Copying and Proofreading (Day 4). The assessment clearly supports the Writing-as-a-Process instructional model.

*Case 2:* School district Y also conducts a direct writing assessment annually in May. Each student has a

conference with his or her teacher to determine which paper from the student's portfolio to submit for assessment purposes. The papers in the portfolio have not been generated under standardized conditions but, rather, represent the ongoing work of the student for the year. All the papers were developed by the student, with as much or as little time allocated to each of the Writing-as-a-Process stages as he or she saw fit.

Is Case 1 an example of a performance assessment? Yes. The students are asked to perform specific behaviors that are to be assessed: to prove that they can write, the students produce a writing sample. Is Case 2 an example of a performance assessment? Yes, also. The portfolio contains numerous examples of actual student performance, although much of the structure associated with testing has been removed.

Is Case 1 an example of an authentic assessment? No. While the students are asked to perform the specific behavior to be assessed, the context is contrived. In real life, individuals seldom write under the conditions imposed during a standardized direct writing assessment. Is Case 2 an example of an authentic assessment? Yes. Performance is assessed in a context more like that encountered in real life; for example, students independently determined how long to spend on the various stages of the writing process, creating as many or as few rough drafts as they saw necessary to complete their final copies.



As we can see, performance assessment refers to the kind of student response to be examined; authentic assessment refers to the context in which that response is performed. While not all performance assessments are authentic, it is difficult to imagine an authentic assessment that would not also be a performance assessment.

### Criteria for Authenticity

To determine whether a given performance assessment is authentic, we must ask, "Authentic to what?" It is a seemingly simple question, but one whose answer may be complex. The following are just a few facets of authenticity: stimuli, task complexity, locus of control, motivation, spontaneity, resources, conditions, criteria, standards, consequences.

Some of these points may be more critical than others in a particular assessment. The assessor needs to make that determination. But in labeling an assessment as authentic, the assessor must specify in what respects the assessment is authentic.

Moreover, because authenticity has a multidimensional nature, some assessments are more authentic than others. Ironically, the most authentic assessment in many situations can probably not be contrived for purposes of testing, for then it would no longer be totally authentic. Educators and assessors must thus be explicit about which facets of authenticity are most critical.

### Proposed Definitions

Two definitions may help further clarify the distinction between the two terms.

In a performance assessment, the student completes or demonstrates the same behavior that the assessor desires to measure. There is a minimal degree, if any, of inference involved. For example, if the behavior to be measured is writing, the student writes. The student does not complete

multiple-choice questions about sentences and paragraphs, which instead measure the student's ability to proofread other people's writing, and require a high degree of inference about the student's ability to write.

In an authentic assessment, the student not only completes or demonstrates the desired behavior, but also does it in a real-life context. "Real life" may be in terms of the student (for example, the classroom) or an adult expectation. The significant criterion for the authenticity of a writing assessment might be that the locus of control rests with the student; that is, the student determines the topic, the time allocated, the pacing, and the conditions under which the writing sample is generated.

### Implications for Educators

What significance do these definitions have for educators? First, when we read materials or attend presentations, we must determine whether the authors or presenters are sensitive to the distinction between the two terms. We must be particularly cautious of generalizing from information provided by individuals who use the terms interchangeably.

Second, we must become informed consumers when purchasing tests, assessment programs, or other materials being marketed as either performance or authentic assessments. Buzzwords sell, unfortunately, so beware.

Third, when planning an assessment, we must carefully identify the purpose in order to determine whether performance assessment — authentic or not — is relevant. Only appropriate matches will improve assessment of student learning. □

**Carol A. Meyer** is an Evaluation Specialist in the Planning and Evaluation Department at Beaverton School District, P.O. Box 200, Beaverton, OR 97075.

will be enough to paint a vivid picture for the reader? Having the teacher and students explore beforehand the meaning of "sufficient detail" can be an effective way to make the criterion clear to students.

3. *Draws out relationships or comparisons between that period of history and the present.* This criterion addresses critical thinking needed to make relationships, draw inferences, and engage in analysis. Both the teacher and students should examine the appropriateness and accuracy of the comparisons. Are these only the most obvious? Are they the most significant? Understanding and making relationships is not a skill that needs to wait until middle or high school. Elementary school children can learn to identify similarities and differences.

4. *Uses affective language in dealing with the experiences of people — in history and today.* This criterion requires preparation, as do the others, in the learning experiences that build up to the assessment. Our traditional testing practices have not emphasized affective goals. What better way to begin to make history meaningful than to see it as affecting the way people feel about their lives? Both the teacher and students need to ask questions like: Does the affective language capture what it might feel like to live in a period of war, given the circumstances of the time? Does the student link the way people in that period might have felt with his or her own feelings in a similar experience?

By being explicit and open about the criteria and giving students many examples of excellent work, we give them guidelines for improvement.

### How Can We Provide Feedback?

Because the types of performances we've described do not reflect single

Copyright © 1992 by the Association for Supervision and Curriculum Development. All rights reserved.