



The Prepared Practitioner

April/May 2009, Bridging Educational Theory and Practice

Alan Colburn

An Assessment Primer

Although the title of this month's column may make your skin crawl—read on. Everyone should understand a few basics about the “A” word. Being able to distinguish formative from summative assessment and criterion-based from norm-referenced testing helps not only teachers, but also parents, the larger community, and students.

Often the first idea that comes to mind when considering assessment is a test: an exam that collects information about whether students understand knowledge or skills the teacher expects them to have learned. This traditional idea represents *summative assessment*, which comes at the end or “summation” of an instructional unit.

Formative assessment, on the other hand, is diagnostic—like a medical test. The purpose of this information is to understand what a student knows or can do now in order to figure out what should come next. Although both formative and summative assessments are checking for understanding, the results of formative assessments are used by teachers and students to adjust what they are going to do and to improve learning; the results of summative assessments are primarily used to assign a grade.

As pointed out by Popham (2008), the related term *formative evaluation* was coined in the late 1960s to evaluate an educational program “while the program is still malleable—capable of being improved because of an evaluation’s results...In contrast, when a mature, final-version educational program is evaluated in order to make a decision about its continuation or termination, this constitutes summative evaluation” (p. 3). Educators quickly expanded the use of these terms from evaluation to assessment and applied them to student knowledge as well.

Using a certain assessment instrument does not determine whether a teacher is assessing formatively or summatively. Rather, a teacher is assessing formatively when he or she uses assessment data as part of a planned process to adjust the teaching of the current unit. When a teacher uses data to determine what a student learned and to assign a grade, he or she is assessing summatively.

The distinction between norm-referenced and criterion-based tests is equally important. In a criterion-based test, each student’s performance is compared to preestablished criteria. If a student’s score is above some minimum, then he or she passes the test or is otherwise considered competent (e.g., a driver’s-license exam). In theory, everyone with a score above the minimum could pass the test.

In norm-referenced tests, on the other hand, scores are ranked and compared to one another (e.g., most high-

stakes standardized tests). The top scores get high numbers and the bottom scores get low numbers—regardless of how well individuals actually perform. Curved grading is norm-referenced: No matter what happens, someone will pass and someone will fail.

Criterion-based and norm-referenced tests are designed for different purposes. Criterion-based tests should measure competency, while norm-referenced tests are used to sort or rank. It would be great if everyone got a particular item correct in a criterion-based test. In a norm-referenced test, however, an item that everyone got right would probably be thrown away, as it does not help to rank or sort test takers. For this reason, norm-referenced tests should never be used to determine competency—but they often are used this way.

Everyone should understand a few basics about the “A” word.

A score on a standardized aptitude test, for example, does not measure an individual student’s aptitude; it measures aptitude compared to peers. Similarly, in international comparisons, U.S. student scores may rank in the bottom half, yet be only slightly lower than those in the top half, and these students may still be learning everything they need to learn.

Take for another example the last Olympics, at which the results of one swimming event were extremely close. When we consider that multiple competitors at the event broke the previous world record, we think about the event in a criterion-based way—each record-breaking swimmer exceeded a preestablished criteria. When we consider who was first and who was fifth, we think about the event in a norm-referenced way.

As science people, readers of *The Science Teacher* are comfortable with the idea of working with data. Many around us may feel differently. In a school system awash in assessment data, helping others understand where the data comes from and how it is meant to be used becomes critical. Together, we can turn the “A” word from “awful” to “awesome” (...or at least “all right”)!

Alan Colburn (acolburn@csulb.edu) is a professor of science education at California State University–Long Beach, and author of The Lingo of Learning from NSTA Press.

Reference

Popham, W.J. 2008. *Transformative assessment*. Alexandria, VA: ASCD.