A Tale of TWO TEST TYPES

Good assessments should measure what students have learned and what they need to learn.

W. James Popham
IN BRIEF

The author discusses the roles of classroom and accountability assessments in promoting instruction. He points out that classroom testing ought to be used primarily to promote improved learning rather than to determine how much students have learned, which is the purpose of accountability testing. Three key requisites are presented for each of these essential test types.

“...it was the best of times, it was the worst of times…”
Or so we are told by Charles Dickens in his classic characterization of the French revolution, A Tale of Two Cities. A similar paradox could be used today to describe today’s educational testing. Good tests are fostering first-rate instruction; bad tests are triggering terrible instruction. Sadly, this is true for both classroom assessments and accountability assessments.
Two Test Types

When I refer to classroom assessments, I’m thinking of the teacher-made tests that have abounded in our classrooms for centuries. These days, however, classroom assessments also include the ever-increasing number of tests developed for classroom use by states, districts, or commercial test publishers. But whether teacher-generated or externally developed, classroom assessments ought to be used chiefly to promote improved learning by students rather than to determine how much those students have actually learned.

Black and Wiliam (1998) have aptly distinguished between “assessment for learning” and “assessment of learning.” If the purpose of schooling is to help children acquire the skills and knowledge they will need in later life, then it should be apparent that classroom assessments’ raison d’être should be to improve the effectiveness of students’ learning. They should be assessments for learning.

Accountability assessments have a different role. Their function is to determine how much students actually have learned. Currently, the dominant accountability assessments are—without dispute—the annual achievement tests that each state is required to administer under the No Child Left Behind (NCLB) Act. Although there is considerable variability among state examinations, their underlying purpose is to let educational policy-makers and the general citizenry judge how well teachers have taught their students. Accountability tests, therefore, are dominantly assessments of learning.

The task before today’s educators is to promote the adoption of good classroom assessments and good accountability assessments so that both strongly contribute to children’s learning. But this means that educators must first be able to distinguish between what’s good and what’s bad in both of these test types.

Good Accountability Assessments

In order for an accountability test to provide accurate evidence of teachers’ instructional effectiveness, yet also contribute to their instructional decision-making, that test must possess three indispensable requisites (Commission on Instructionally Supportive Assessment 2001).

Requisite 1: A manageable number of assessment targets. A good accountability test should measure students’ mastery of only a modest number of important curricular aims rather than a litany of less important ones. Thus, a good accountability test in reading would attempt to measure students’ mastery of perhaps a half-dozen really significant reading skills. This less-is-more assessment strategy stems from recognition that teachers can focus their attention on only an intellectually manageable number of targets. Therefore, a good accountability test will measure students’ mastery of only those targets.
Requisite 2: Lucid assessment descriptors. A second necessary attribute of a good accountability test is that it must be accompanied by clear descriptions of whatever skills or bodies of knowledge are going to be assessed. If teachers understand the nature of the skills or knowledge to be assessed, they can devise instructional strategies to promote students’ mastery. It is important, however, that any description of what is to be measured must be carefully crafted so as to be not only succinct but *teacher palatable*. Lengthy, detailed descriptions simply won’t be read by busy teachers.

Requisite 3: Instructionally useful reporting. The final requisite of a good accountability test is that it must contain a sufficient number of items for each assessed curricular aim so that a reasonably accurate estimate of every student’s mastery of every assessed curricular aim can be reported to teachers, students, and parents. That sort of reporting also is necessary for teachers to evaluate the effectiveness of their own instructional activities.

Sadly, few state tests come even remotely close to satisfying these three requirements. In almost every state, accountability tests chosen to implement NCLB are *instructionally insensitive*—essentially incapable of detecting instructional improvements. Whether the tests are augmented versions of nationally standardized achievement tests or “standards-based” tests allegedly aligned with a state’s official curricular aims, students’ performances are influenced by their socioeconomic status.

Good Classroom Assessments

Before getting into the attributes of good classroom assessments, I first need to deal with classroom tests that are referred to as *formative assessments*—which they really are not. It was Michael Scriven (1967) who first drew a distinction between *summative* and *formative* evaluation in the late 1960s. Summative assessments try to determine what students have already learned, while formative assessments have a role in assisting students to learn.

While many tests recently distributed by commercial publishers are described as formative assessments, a closer inspection of these tests reveals that they are usually nothing more than mirror images of upcoming accountability tests. It would be more accurate to describe them as early-warning summative tests, and not formative tests.

A good classroom test is one that helps teachers and students decide how to better achieve the skills and knowledge the teacher has set out for students to master. Good classroom assessments, therefore, are chiefly assessments for learning, and here again there are three requisites of such tests.

Requisite 1: A preoccupation with instructional decisions. A classroom assessment for learning has a mission: to help kids learn better. Accordingly, such tests must be devised with instructional decisions constantly in mind. Rather than asking teachers to whomp up tests to tell whether or not a student has mastered a curricular aim, we must encourage teachers to generate classroom tests that supply both teachers and students with insights into students’ *evolving* mastery of the curricular aim.

Let me illustrate what such a classroom test might look like. Suppose a fifth-grade teacher (I’ll call him Mr. Melvin) is attempting to get his students to master a high-level cognitive skill in social studies. Recognizing that this particular skill is a challenging one, he has set aside six weeks to teach it. He also concludes that, in order for his students to master the skill, they must first master three subskills and a body of factual knowledge and terminology. Such sequenced enabling skills and knowledge needed by students *prior* to mastering an important curricular aim are referred to as *learning progressions* or *progress maps*.

If the classroom tests Mr. Melvin uses during his six-week unit are going to be legitimately labeled as assessments for learning, then they need to provide both the teacher and his students with insights regarding each student’s progress in mastering this prior knowledge. In some instances, he might give brief tests dealing separately with each of the enabling outcomes. In other cases, they could be combined in somewhat lengthier classroom assessments.

But the overriding question that Mr. Melvin’s classroom assessments must always be able to answer affirmatively is: Will the results of these tests inform both teacher and students about what is needed instructionally?

Requisite 2: Prioritized curricular aims. To create a legitimate assessment for learning takes plenty of hard thinking on the teacher’s part. Thinking through the nature of a progress map for a high-level cognitive skill is far from fool’s play.

Accordingly, just as we saw that a good accountability test should be focused on only a modest number of genuinely important curricular aims, the same goes for good classroom tests. If teachers can isolate a half-dozen really meritorious curricular outcomes to pursue during a school year, then those are the outcomes for which classroom tests should be developed.

I realize that what I’ve just recommended runs counter to the current NCLB-induced paranoia for teachers to “cover” every content standard. But “cover-everything” teaching doesn’t help kids learn what they truly need to learn. If teachers select what they regard as the most worthwhile of a state’s official curricular aims, and do a terrific job of getting their students to *deeply* master them, I assure you that those students will perform as well on NCLB tests as those who receive superficial, touch-all-bases instruction.

Requisite 3: Meaningful involvement of students. The third requisite of good classroom assessment is, of course, found in the use of such tests. There are enormous instructional payoffs in making students active, assessment-informed partners in the learning process. A teacher who is an adroit user of good classroom assessments will always provide students with readily understandable follow-up activities that can shore up shortcomings or build on prior accomplishments.
What’s to Be Done?

In my view, there is a crucial first step in improving the quality of both classroom and accountability testing. Educators, especially principals, need to beef up their knowledge about the basics of assessment if they are to push for improvements in educational testing. There are a number of easy-to-comprehend publications now available to help teachers and administrators learn what they need to know about educational tests (see box) and I recommend reading them.

If you will allow me to use a poor paraphrase of the concluding sentence in Dickens’ A Tale of Two Cities, if you set out to acquire a dollop of assessment literacy, “It will be a far, far better thing you do than you have ever done before.”

References
GEICO could save you $500 a year on car insurance.

It pays to do your homework.

NAESP members could receive a special discount on GEICO car insurance.

Visit geico.com for your free rate quote and be sure to select NAESP when asked for your affiliation.

GEICO offers you:
• Outstanding, 24-hour service online or on the phone.
• Fast, fair claim handling.
• Guaranteed claim repairs at GEICO-recommended shops.

To find out how much you could save visit geico.com or call 1-800-368-2734 today.

Average savings information based on GEICO New Policyholder Survey data through August 2005.

Discount amount varies in some states. Some discounts, coverages, payment plans, and features are not available in all states or in all GEICO companies. One group discount applicable per policy. Government Employees Insurance Co. • GEICO General Insurance Co. • GEICO Indemnity Co. • GEICO Casualty Co. These companies are subsidiaries of Berkshire Hathaway Inc. GEICO auto insurance is not available in Mass. GEICO, Washington, DC 20076. © 2005 GEICO