

HOW TO NURTURE THE CONSTRUCTION OF INSTRUCTIONALLY SUPPORTIVE TESTS: AN EXPERT PANEL'S VIEW*

W. James Popham
University of California, Los Angeles

“High-Stakes Testing” is an expression that, to educators at least, seems certain to function as a controversy-engendering lightning rod during the first part of the 21st century. It is a rare month that goes by these days without the appearance of a major newspaper story focused on the merits or demerits of high-stakes educational tests.

The most vocal proponents of high-stakes educational tests are those who believe that credible evidence of educational quality can best be obtained via such assessments. In particular, advocates of high-stakes tests believe that standardized achievement tests (either national, off-the-shelf tests or customized, state-specific tests) can supply the only *bona fide* evidence of educators' effectiveness.

Critics of today's high-stakes educational tests, however, have been increasing not only in numbers, but in the zeal with which they oppose these tests. *Newsweek* recently indicated that there are now more than two dozen states in which parent-teacher groups have been formed explicitly to oppose high-stakes educational testing.

Even the strongest supporters of today's high-stakes tests will concede that, at least in certain localities, test-based pressures on teachers to increase students' performances have led to an erosion of educational quality. Curricular coverage has been reduced because content not included on high-stakes tests tends to be passed over by teachers. In addition, excessive test-focused drilling has diminished the positive affect toward learning that students might otherwise acquire. Finally, we hear increasing reports that some test-pressured teachers have engaged in patently reprehensible test-preparation or test-administration conduct.

Some analysts of today's large-scale testing terrain believe that many of the assessment-triggered shortcomings we currently see in our schools could be rectified if only more appropriate tests were employed. These individuals argue that traditionally constructed standardized achievement tests should be replaced by tests that not only supply accurate accountability evidence but, at the same time, can also become *instructionally supportive*.

* Presented at the annual meeting of the National Council on Measurement in Education, April 2-4, 2002, New Orleans, Louisiana.

A Five-Association Project

Having recognized that most of today's high-stakes tests are selected or built at the state level, five national associations of educators decided in the summer of 2001 to undertake a project intended to foster the creation of high-stakes tests that would be educationally beneficial, not detrimental. A coalition was formed by the American Association of School Administrators, the National Association of Elementary School Principals, the National Association of Secondary School Principals, the National Education Association, and the National Middle School Association. This coalition authorized and supported a project calling for an independent commission to create a set of recommendations regarding the construction of instructionally supportive high-stakes tests.

The Commission on Instructionally Supportive Assessment was appointed in July 2001 to recommend how high-stakes accountability tests could be fashioned so that those tests also contributed to improved instruction. Members of the commission were Eva L. Baker, University of California, Los Angeles; David C. Berliner, Arizona State University; Carol Camp Yeakey, University of Virginia; James W. Pellegrino, University of Illinois at Chicago; W. James Popham, University of California, Los Angeles; Rachel F. Quenemoen, University of Minnesota; Flora V. Rodríguez-Brown, University of Illinois at Chicago; Paul D. Sandifer (Ret.), South Carolina Department of Education; Stephen G. Sireci, University of Massachusetts, Amherst; and Martha L. Thurlow, University of Minnesota. I was privileged to function as the commission's chair. I am grateful for the diligent and effective work of the nine outstanding individuals who served with me on the commission. Our report was released in late October 2001.*

My presentation today focuses on the conclusions reached by our commission. Because the commission's members devoted considerable attention to the precise wording of our conclusions, I will draw directly on the language we employed in listing nine specific requirements that an accountability test must satisfy if it is also to enhance teachers' instructional decision-making. Indeed, I will even borrow the language we used as a rationale for each of the nine requirements. Having listed those nine requirements, along with our rationale for each requirement (in italics), I will close out this paper with a few personal observations about the commission's work. I turn, now, to the nine requirements for high-stakes tests set forth by the Commission on Instructionally Supportive Assessment.

Requirement 1: Prioritized Content Standards

Whereas:

- *educators in many states cannot adequately address within the amount of time available for instruction the large number of content standards that are supposedly measured by state tests;*

* The Commission on Instructionally Supportive Assessment. *Building Tests That Support Instruction and Accountability: A Guide for Policymakers*. Washington, DC: Author, 2001.

- *state tests often do not adequately assess all of the content standards, and frequently center on standards that are easiest to assess; and*
- *state tests rarely provide educators with the kind of information they need to improve instruction; therefore:*

A state’s content standards must be prioritized to support effective instruction and assessment.

Requirement 2: Unambiguously Described Content Standards

Whereas:

- *a state’s high-priority content standards will be measured by state tests;*
- *educators must understand what each of these content standards calls for from students; and*
- *many content standards are not worded with sufficient clarity for rigorous instructional planning and assessment design, therefore:*

A state’s high-priority content standards must be clearly and thoroughly described so that the knowledge and skills students need to demonstrate competence are evident.

Requirement 3: Standard-by-Standard Reporting

Whereas:

- *students, parents, educators, and policymakers need information about which content standards students are and are not attaining; and*
- *educators can do little to improve students’ achievement without information about their performance on each high-priority content standard, therefore:*

The results of a state’s assessment of high-priority content standards should be reported standard-by-standard for each student, school, and district.

Requirement 4: Classroom Assessments for State-Unassessed Content Standards

Whereas:

- *content standards that are not assessed by state tests are important and should be given instructional attention;*

- *educators need good assessment tools to monitor students' achievement and rarely have the time and resources to develop such tools; and*
- *assessments that are routinely administered by educators can and should be used to provide a complete picture of what students know and are able to do, therefore:*

A state must provide educators with optional classroom assessment procedures that can measure students' progress in attaining content standards not assessed by state tests.

Requirement 5: Monitoring Curricular Breadth

Whereas:

- *students benefit from a rich and deep curriculum; and*
- *state tests that measure high-priority content standards could narrow curricular coverage unless steps are taken to forestall such narrowing, therefore:*

A state must monitor the breadth of the curriculum to ensure that instructional attention is given to all content standards and subject areas, including those that are not assessed by state tests.

Requirement 6: Appropriate Assessment for All Students

Whereas:

- *all students must be given appropriate opportunities to demonstrate the degree to which they have mastered state content standards; and*
- *federal statutes require that such opportunities be offered in particular ways, therefore:*

A state must ensure that all students have the opportunity to demonstrate their achievement of state standards; consequently, it must provide well-designed assessments appropriate for a broad range of students, with accommodations and alternate methods of assessment available for students who need them.

Requirement 7: Sufficient Test-Development Time

Whereas:

- *tests developed too hurriedly neither support instruction nor supply accurate evaluative information for accountability programs; and*
- *there is a widespread misunderstanding that high-quality achievement tests can be developed in two years or less, therefore:*

A state must generally allow test developers a minimum of three years to produce statewide tests that satisfy the Standards for Educational and Psychological Testing and similar test-quality guidelines.*

Requirement 8: Pertinent Professional Development

Whereas:

- *most educators are unfamiliar with the instructionally supportive assessment system the Commission advocates; and*
- *it is imperative that educators become adept at using such an assessment system for instructional purposes, therefore:*

A state must ensure that educators receive professional development focused on how to optimize children's learning based on the results of instructionally supportive assessments.

Requirement 9: Ongoing Evaluation

Whereas:

- *any assessment system, no matter how well designed, can be improved; and*
- *a state's assessment system must perform both an accountability function and an instructional improvement function, therefore:*

A state should secure evidence that supports the ongoing improvement of its state assessments to ensure those assessments are (a) appropriate for the accountability purposes for which they are used, (b) appropriate for determining whether

* See American Educational Research Association. 1999. *Standards for Educational and Psychological Testing*, Washington, DC: Author. See also such documents as Pellegrino, J.W., N. Chudowsky, and R. Glaser (Eds.) 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press.

students have attained state standards, (c) appropriate for enhancing instruction, and (d) not the cause of negative consequences.

Two Reports

The commission's members concluded, early on, that our conclusions need to be concise if they were going to be read—by anyone, especially busy policymakers. Accordingly, we created a relatively succinct report aimed specifically at the state-level policymakers who would be deciding about the kinds of accountability tests to use in their states. This report was distributed widely by the convening associations in late October 2001.

In addition, however, because the commission was eager to assist those state-level officials who might wish to implement our suggestions, we prepared a second report providing illustrations of how the commission's nine requirements could be satisfied. We recognized the likelihood that a federal law would soon be passed calling for states' increased use of achievement tests in grades three through eight. Because many states would be required to carry out the necessary test-development by using the services of an external contractor, and because this would typically require a state's issuance of a request for proposals (RFP), we developed a second report containing *illustrative* (not recommended) RFP-language that could be used "as is" or altered if desired by state-level personnel. The commission's members wished to describe at least one set of procedures, as an example, that could be followed by state personnel if commission-recommended accountability tests were to be created. The second report's illustrative RFP-language showed how each of the commission's nine requirements could be addressed by state-level assessment personnel.

Both the report for policymakers as well as the illustrative RFP-language report are available online from each of the five convening associations.*

Personal Reflections

Now, having devoted the bulk of this paper to a recapitulation of the conclusions of the Commission on Instructionally Supportive Assessment, I want to switch gears for a few paragraphs and look back at the work of the commission from a personal perspective.

First off, I am proud of the commission's work. I believe that if an accountability test were constructed in accord with the commission's nine requirements, such a test would not only supply accurate evidence by which to evaluate schools, but the test would also help teachers do a better instructional job.

* The Commission on Instructionally Supportive Assessment was convened by the American Association of School Administrators, the National Association of Elementary School Principals, the National Association of Secondary School Principals, the National Education Association, and the National Middle School Association. Available online at: www.aasa.org, www.naesp.org, www.principals.org, www.nea.org, and www.nmsa.org.

Second, I am immensely grateful to Carol, Dave, Eva, Flora, Jim, Martha, Paul, Rachel, and Steve for agreeing to carry out the commission's work—without recompense and under an astonishing tight timeline. These are smart, top-flight professionals, all of whom already had piles of existing commitments on their plates. Yet, because they believed in the potential payoff of the commission's mission, they took the time to tackle this project, and to do so with considerable energy and insight.

Third, I think that the five convening associations, namely, AASA, NAESP, NASSP, NEA, and NMSA, displayed both professional responsibility and political acumen by underwriting this project. And, most importantly, having promised that the commission would, once convened, be allowed to function independently, the five associations kept their word. They did not interfere with our deliberations or attempt to influence our conclusions. Finally, the five associations did a superb job of disseminating a professionally published report for policymakers and making both that report, as well as the report containing illustrative RFP-language, available on-line. Marcella Dianda of NEA gets special applause for her deft stewardship of the five associations' collaborative efforts.

Finally, what about the potential impact of the commission's two reports? Well, here we are, roughly a half-year from the release of the Commission's reports, and the ballots are still being counted. Will the commission's efforts make any difference at all?

I am guardedly hopeful. Our nation must now implement federal legislation calling for the state-level creation of many, many achievement tests that, almost certainly, will play a role in states' school-evaluation efforts. It's not that I think our commission's words, no matter how thoughtfully crafted, will cause policymakers to instantly opt for commission-endorsed tests instead of traditional ones. No, I have learned through the years that words-on-paper or, these days, words-delivered-electronically, rarely alter human behavior.

What does alter human behavior is *self-interest*. Regrettably, but realistically, people behave in ways that they think will be personally good for them. Fortunately, I think it is just possible that the commission's vision of instructionally supportive assessment may coincide with the self-interest of many key constituencies who have a stake in today's educational accountability arena.

- The five convening associations want their members' work to be appraised accurately with tests that give educators a realistic chance to be successful. Traditional accountability tests can't do this. Commission-recommended accountability tests can.
- Educational policymakers want to be associated with successful educational improvement strategies that lead to *demonstrably* improved learning by students. Traditional accountability tests have not and can not do this. Commission-recommended accountability tests can.

- Parents want to see schools help their children master significant skills and knowledge. And parents also want to see *credible* evidence that such mastery is taking place. Traditional accountability tests can't do this. Commission-recommended accountability tests can.
- Children want to learn what they need to learn in order to be successful and happy both in the present and in the future. Traditional accountability tests can't help them do so. Commission-recommended accountability tests can.

A confluence of those self-interests, on the part of educators, policymakers, parents, and children, just might lead to improved instruction in our schools. As I said before, I am guardedly hopeful.

Most people who went into education did so because they wanted to help children achieve their inherent potentials. If we are going to help children do so these days, in an altogether understandable era of educational accountability, we must create evaluative systems wherein it is in the personal self-interest of adults to fashion educational interventions that, coincidentally, turn out to benefit children. Assessment-based evaluative strategies that are carefully conceived to be instructionally supportive can benefit all concerned players. The accountability tests conceptualized by the Commission on Instructionally Supportive Assessment, if artfully installed, can be part of an improvement strategy that benefits kids.