Experts Urge States To Stay Course on High-Quality Assessments

By Catherine Gewertz on June 20, 2013 10:26 PM

Fort Washington, Md.

A bevy of heavy-hitting assessment experts has identified five things that make assessments high quality, and is urging states to hold out for such tests in the face of political and financial pressures that might weaken their resolve.

Their report, "Criteria for Higher-Quality Assessment," urges states and districts to demand these five things when evaluating or building assessment systems:

- That they examine higher-order thinking skills, especially those that are transferable and relate to applying knowledge to new contexts.
- That they provide "high fidelity" evaluation of those higher-order skills, such as through researching and presenting arguments.
- That they are internationally benchmarked to align assessment content and measurement practices with those used in leading nations.
- That they use "instructionally sensitive" items that reflect how well teachers are teaching and give them useful guidance on how to improve.
- That they are valid, reliable, and fair, as well as accessible to all learners.

These are the key factors, the experts said, needed for systems of assessment that not only measure the most important aspects of student learning in meaningful ways, but encourage good teaching and learning in classrooms, and produce timely, actionable feedback for teachers so they can adjust and customize instruction.

The criteria were crafted by many of the country's best-known thinkers on gauging student learning. The lead authors of the report are James Pellegrino, Linda Darling-Hammond, and Joan Herman. Three institutions released it jointly: the Stanford Center for Opportunity Policy in Education, or SCOPE; The National Center for Research on Evaluation, Standards & Student Testing, or CRESST, at the University of California-Los Angeles; and the Learning Sciences Research Institute at the University of Illinois at Chicago.

The assessment scholars chose the Council of Chief State School Officers' annual conference on student assessment here to release their report today. It appears to be intended in part to keep states on course toward the common assessments being designed by two state consortia at a time that some are getting political, financial and technological jitters, and in a few cases withdrawing from those consortium projects. It also appears to be aimed at the rising number of vendors that are getting into the market to assess the common standards.
In a conference session devoted to discussing the report, Joan Herman, a co-director of CRESST, said there are signs that some states could move away from the aspirational assessments currently being designed by two state consortia, and choose cheaper, quicker tests in their place. Given the consortia’s vision—to produce a suite of formative resources that support instruction, along with interim tests and summative assessments that include performance tasks and constructed-response items—such moves could take states back to the limitations of their current, predominantly multiple-choice tests, she said.

States are worried about whether the new tests will take longer, cost more, or run into technological snafus. Some worry that one or both consortia won’t deliver the tests on time, in 2014-15, as promised. Others are battling political headwinds as activists, lawmakers, and others push back against the tests—and the standards on which they’re based—as a federal intrusion on local education decisions, since the U.S. Department of Education funded the common assessment development, and strongly encouraged states to adopt the standards.

The consortia themselves are having to contend with these pressures, and could well be among the intended targets of the new report’s message. Both PARCC and Smarter Balanced consortia have revised or scaled back their test designs along the way in response to concerns about testing time and cost.

“There are threats every day” that states will back off their high expectations for new tests, Herman told the conference participants. In the face of the pressures on them, “how do we help states stay the course?”

Pellegrino, who led the session discussion with Herman, told participants that the crafters of the criteria have a “suite” of assessments in mind, not just one summative test. Formative strategies and tools to help teachers gauge and guide learning as it happens are an important part of that picture, as are interim tests, he said. But what appears on large-scale summative tests is important, he said, because of its power to shape what happens in the classroom. So statewide summative tests must include and encourage the types of activities that are complex and meaningful learning activities, he said.

One education researcher in the audience, Fritz Mosher, noted that embracing large-scale summative exams with the qualities outlined in the report would suggest that states and districts have embraced curricula defined by similar criteria. Chuckling that Mosher had “brought the C word into the conversation,” Pellegrino noted that most curricula and tests don’t reach that level. “We have a long way to go in our curriculum development process,” he said.

Participants brainstormed about how to keep states and districts focused on the potential benefits of higher-quality tests. Particularly important, some said, was including a number of performance tasks that require students to engage in longer, more complex activities to demonstrate their understanding of the materials. But those items are more costly to build and score. How to build crucial buy-in for good assessments as brushfires of anti-testing sentiment keep cropping up in states is a crucial question for which there were no easy answers.

Experience in recent decades with large-scale testing has left the public disillusioned by “fairly sterile multiple-choice” exams, Pellegrino said, so the task of convincing people that new tests will be fundamentally different—and better—is a tough one.

"Part of the backlash is legitimate,” he said. “[People ask], 'Why are my kids spending time answering meaningless questions'? We have a ways to go to demonstrate that the questions we are designing are relevant and valuable.”

[8]

Categories: Testing
Tags: assessment, assessment criteria, common assessments, CRESST, Learning Sciences Research Institute, PARCC, SCOPE, Smarter Balanced, testing

Print, Email, Tweet 19, Like

You must be logged in to leave a comment. Login | Register

2 of 3 6/21/13 10:08 AM
Ground Rules for Posting
We encourage lively debate, but please be respectful of others. Profanity and personal attacks are prohibited. By commenting, you are agreeing to abide by our user agreement.
All comments are public.