

Next-Generation Assessment Systems

By Stanley N. Rabinowitz

An unprecedented confluence of factors—economic, political, and educational—is causing many states to rethink their student-assessment programs. But careful thought and expert guidance will be needed if they are to avoid the problems of the past and take advantage of promising new developments.

Most state assessment programs, regardless of their history or goals, were revised early in the last decade to meet the requirements of the federal No Child Left Behind Act: All assess reading/language arts and mathematics in grades 3-8 and high school, and science in elementary, middle, and high school. Collectively, they are increasingly homogeneous, largely multiple-choice, with some sprinkling of constructed-response and direct writing.

Change is in the offing. The upcoming reauthorization of the Elementary and Secondary Education Act, of which No Child Left Behind is the current version, is likely to result in changes to assessment requirements, allowing more flexibility and providing greater support for innovative assessment models. Dissatisfaction with current limited options, coupled with the Common Core State Standards Initiative's potential impact—assessments covering fewer but clearer objectives, and advancing more-rigorous content and skills requirements—will necessitate a broad reconceptualization of assessment. This inevitably will mean a shift away from state standardized testing as the only game in town, and a move toward the development of innovative state assessment *systems*.

This next-generation model will include differentiated roles for assessment at the federal, state, and school levels; the use of multiple measures; and assessments that support accountability programs focused on both growth and current status. It also will be likely to take greater advantage of technology, and will benefit from U.S. Department of Education initiatives and dollars, represented by the Obama administration's Race to the Top Fund and other grant programs.

What follows is an overview of key concepts states should consider as they move forward.

Multiple Measures. The growing call for more performance-based assessment is reminiscent of past practice, when programs such as the New Standards Project and the Vermont and Kentucky writing and math portfolios of the early 1990s seemed to usher in a new era. But significant changes in the political context of accountability, along with technological advances such as the use of computer simulations, should encourage states to look to the future, not the past.

To do this, we need to create a new vision of technical evidence for performance-based assessments, building on advances in such work as alternative assessments for students with disabilities. And we must ensure that teachers get the necessary training to incorporate performance-based instruction into their teaching, and to use classroom-embedded assessments to diagnose students' strengths and weaknesses.

Balance. When researchers call for the development of a *balanced* assessment system, they typically mean balancing roles and responsibilities. But other factors also must be balanced:

- **Cost.** What is the most cost-efficient way to get teachers, students, administrators, policymakers, and parents the information they need about individual students and the system itself?

- *Constraints.* How do we overcome technical limitations in the next-generation assessments and properly train teachers and others to develop and use new techniques and understandings?
- *Risk.* How do we support true reform without overburdening students, teachers, and other key constituencies?

Complementary Components. What constitutes the “system” in our next-generation assessment system? Having a system means there are several assessment components, each playing a complementary, not duplicative, role. Our system would have both formative (diagnostic, classroom-embedded) and summative (used for accountability, mostly on-demand) assessments. Each type would be designated for its primary role. And having both would allow teachers to use formative assessments honestly: to focus on how well students are advancing, without worrying that short-term deficiencies will affect their own evaluations or reflect badly on their students.

Valuing efficiency, our ideal system would include multiple-choice, constructed-response, and performance tasks, as appropriate, to best measure the knowledge and skills called for in challenging new content standards. Indicators would need to justify these various options’ inclusion: What is the *incremental validity* of each component (in other words, what would be lost if it were excluded)?

Roles and Responsibilities. In our system, each governing level would play a different primary role. The federal role would be limited but targeted, ensuring that all students achieved sufficient levels of reading, math, and science proficiency. States would build on that foundation by adding other content areas, such as social studies, and including additional indicators—those based, for example, on local economic conditions, resources, and values and beliefs. The primary local-level function would be curriculum and instruction, ensuring that all teachers are prepared to meet, and are supported in meeting, the needs of all students, including those most at risk, such as English-language learners, students with disabilities, and students living in poverty.

Assessments at each level would be consistent with that level’s “system” responsibilities. One beneficial outcome of having differentiated roles with specialized assessments would be that the technical requirements could vary based on the purpose and the stakes, allowing different types of validity evidence, for example, to drive local diagnostic tasks, as opposed to state and federal accountability assessments.

Choice. While most assessment and accountability systems expect all students to master all content at a “major” level, regardless of students’ interests or abilities, our system would incorporate an element of choice at the state, school, and student levels. Students could pick the content area or areas on which they wished to focus deeply, for example, and would be assessed accordingly, so long as they also demonstrated sufficient knowledge in other areas to support responsible citizenship and employment.

Some states have already moved in this direction by supporting magnet and charter schools, adding new content areas and indicators to state assessment and accountability programs, and requiring career “majors” for students at the high school level.

The bottom line when it comes to choice is that all schools and students should be great at something that society values, and sufficiently good at all else, to maximize motivation and success while keeping students’ options open for later-developing needs and interests.

Too often, assessment-reform plans are dismissed from the start because of the identification of what are thought to be barriers to their development and implementation. The real world certainly has real constraints. But before we say “we can’t,” we should ask ourselves the following questions:

- Is the barrier real, or is it just perceived?
- Is it real, or is the change just going against the traditional approach?
- Is it real, or is the change just difficult?
- Is it real, or does the change just require new technical tools?
- Is it real, or is the change just expensive?

We may find surprising opportunities if we dare to address perceived barriers to assessment reform honestly.

Our assessment and accountability systems should reflect what we value most for our students, schools, and society, and what we think it means to be a well-prepared student, worker, and citizen. Once these are clear, we should be willing to fight and to pay for their reflection in our system for measuring academic progress.

As states begin this important endeavor, they should take the following steps:

- Develop a vision statement that incorporates the values the system will represent.
- Devise an implementation plan with goals, key dates, milestones, responsibilities, and necessary resources.
- Secure sufficient funding to implement the plan.
- Develop a system for evaluation and feedback.

There is no time to waste. Much needs to be done, and the quality of American education is at stake.

Stanley N. Rabinowitz directs the Assessment and Standards Development Services program at WestEd, which manages the federal Assessment and Accountability Comprehensive Center. He also serves on the validation committee for the Common Core State Standards Initiative. This essay is adapted from his closing remarks at the 2009 Reidy Interactive Lecture Series.