



Copyright 2013 National School Boards Association. All rights reserved. This article may be printed out and photocopied for individual or noncommercial educational use (50 copy limit), but may not be electronically re-created, stored, or distributed; or otherwise modified, reproduced, transmitted, republished, displayed or distributed. By granting this limited license, NSBA does not waive any of the rights or remedies otherwise available at law or in equity. By granting permission to use of our materials, NSBA does not intend to endorse any company or its products and services.



A New Vision for Accountability

*The Common Core
State Standards will require
more sophisticated tests,
but high-quality assessments
need not break the budget*

Linda Darling-Hammond

Student assessment has never been a hot subject with the media or the public, save for the scorecard-style reporting about winners and losers that grab headlines but rarely go below the surface to answer just what test scores represent. But that is changing. Student assessment is becoming a topic of conversation. It is spurring discussions about how and why we test students and what we do with the data from those assessments.

School boards will play critical roles in determining where that conversation goes in terms of public policy and student performance. And no one knows better than school board members the challenges that lie ahead.

Recent studies and expert panels underscore the importance of these evolving conversations and the high stakes we face for adequately preparing our students for a changing and complex world. Education is the economic issue of our day, and our assessments can drive instruction and learning so that we move forward.

But, just as the nation is poised for this progress, these same studies and experts sound alarms warning educators, policymakers, and the public not to back down from the challenges that lie ahead around student assessments.

Three challenges that come up most often are:

- Creating high-quality assessments that evaluate 21st century skills to replace current low-quality tests built on multiple-choice and fill-in-the-bubble questions.
- Investing wisely in assessment systems that can actually help improve teaching and learning.
- Making sure we use the assessments to support rather than punish students, teachers, and schools.

School board members are becoming engaged with these questions by the adoption of the Common Core State Standards in 45 states and the District of Columbia. The standards are a state-led effort to add rigor to learning goals and help more students become college and career ready.

Assessments being developed to support the Common Core, including those by the Smarter Balanced Assessment Consortium (SBAC) and the Partnership for Assessment of Readiness for College and Careers (PARCC), will be a big step toward measuring 21st century learning and skills.

Standards and assessments also are being revamped in states that have not adopted the Common Core standards, in an effort to better reflect the 21st century challenges our students will face.

Ensuring high quality

As school boards work to figure out how they want to use new assessments, they will face challenges to make sure that assessments are capable not only of gauging our progress toward deeper learning, but also helping us to get there. In our current high-stakes context, what is tested increasingly defines what gets taught.

Unfortunately, recent studies have found most current state tests measure very few of the 21st century skills U.S. students need. Even in the 17 states reputed to have the highest standards, a RAND Corporation report found that fewer than 2 percent of math items and only 20 percent of English language arts items assess higher-order skills.

In March, the 30-member Gordon Commission released a long-awaited report endorsing the Common Core assessment's emphasis on competencies such as critical thinking and problem solving, rather than on the rote recall of information and more basic skills. These are the same kinds of skill sets that are in greatest demand in today's workforce.

"Our conviction is that, while the field of measurement in education has established a splendid history primarily directed at the measurement of education, the future of assessment in education will depend on the field's capacity to pursue assessment for education," said Edmund W. Gordon, the commission's chairman and emeritus professor at Yale University and Teachers College, Columbia

University.

Echoing this concern, the Stanford Center for Opportunity Policy in Education (SCOPE) recently released a report signed by 20 education leaders and assessment experts titled "Criteria for High-Quality Assessment." The authors wrote the report out of concern that educators, policymakers, and assessment developers could back down on assessment changes that will enable tests to truly measure and promote complex learning-skills instruction for deeper learning.

As a measure of what high-quality assessment looks like, the report includes a set of criteria that states and districts should demand from next-generation assessments.

"There are a lot of options for assessment, and it's easy to lose sight of what matters most. These criteria, when followed, will make it much more likely that all assessment tools will work together to promote the instruction and learning that we want and need as a nation," said Joan Herman, co-director emeritus of the National Center for Research on Evaluation, Standards, and Student Testing at the University of California, Los Angeles, which co-sponsored the report.

The five criteria are:

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

The authors argue that these pieces are needed for systems of assessment that measure the most important aspects of student learning in meaningful ways; encourage good teaching and learning in classrooms; and produce timely, actionable feedback for teachers.

"Part of the backlash [against current tests] is legitimate," James Pellegrino, a report lead author and co-director of the Learning Sciences Research Institute at the University of Illinois at Chicago, said during a recent presentation of the report. "[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."

Paying for high-quality assessment

Unfortunately, U.S. assessments lost ground when the No

Many assume high-quality assessments that include open-ended items and tasks will cost too much.



Child Left Behind Act was passed and dramatically increased the amount of testing required. Most states abandoned open-ended performance tasks that measured writing, research skills, and explanations of mathematical thinking—limiting tests to multiple-choice items only, which are cheap to score by machine.

Another SCOPE report, “Developing Assessments of Deeper Learning,” shows that average spending on NCLB tests is about \$25 per pupil—less than one-quarter of 1 percent of overall education costs.

Ironically, though, researchers found that the combined costs of these tests, along with test prep activities, interim and benchmark tests, data systems, and staff time spent on trying to boost scores total more than \$50 per pupil. Although these funds could have supported higher-quality assessments that encourage more productive teaching and learning, they do not, because all efforts are directed on raising scores on low-quality tests that do not measure critically important skills.

Many assume high-quality assessments that include open-ended items and tasks will cost too much, but we found that this need not be the case. Cost analyses show that several factors can make higher-quality assessments feasible and affordable:

- Costs are dramatically reduced when states act together in consortia that unlock economies of scale.

- Online delivery and efficient scoring of open-ended tasks by computers (where the technology allows) and by teachers who are paid professional development stipends also can reduce costs.

- Furthermore, involving teachers in developing and scoring assessments produces the double benefit of improved instruction and more efficient use of resources.

With these innovations and the integration of formative and interim tools into coherent systems like those being developed by PARCC and SBAC, states and districts could spend considerably less than they do now for assessments that are considerably more robust in evaluating college- and career-ready skills. In the end, intelligent investments and a focus on what students and teachers need for the future assessments of deeper learning can become feasible and affordable.

Improvement, not punishment

Finally, the benefits of improved assessments will be realized when they are placed in a new accountability paradigm, as well—one that uses assessment data to guide instruction and inform school decision-making.

Using assessments as sledgehammers, rather than as tools for improvement, could undermine implementation of the Common Core and impede the changes in instruction that are needed to reach the goals of improved learning.

Increasingly, teachers have voiced concerns that using

assessment to single out and punish rather than to support and guide would be both unfair and counterproductive. A recent poll of 800 K-12 teachers by the American Federation of Teachers found that while 75 percent support the Common Core initiative, 83 percent support a moratorium on high-stakes consequences for students, teachers, and schools.

The public also is voicing concern, disapproval, and outright resistance to the use of test results in a punitive way. In April, hundreds of students across New York were joined by teachers and parents in calling for a boycott of standardized tests. “These tests are not being used for educational purposes. They’re strictly used to evaluate the teachers, evaluate the schools, and to punish the teachers in schools,” Jeanette Deutermann, a parent leading the effort, told a local CBS affiliate.

For higher-quality assessments to benefit teaching and learning, it is critical that, rather than pouring new wine into old bottles, we create a new vision for accountability—one focused on improving instruction. ■

Linda Darling-Hammond (ldh@stanford.edu) is the Charles E. Ducommun Professor of Education at Stanford University, where she is the founding director of the Stanford Center for Opportunity Policy in Education. Her research and policy work focus on issues of educational equity, teaching quality, and school reform. In 2008, she served as director of President Obama’s education policy transition team. Her most recent book, published in April, is *Getting Teacher Evaluation Right: What Really Matters for Effectiveness and Improvement*.