

EDUCATION WEEK

Published Online: November 29, 2012

Testing Group Scales Back Performance Items

By **Catherine Gewertz**

A group that is developing tests for half the states in the nation has dramatically reduced the length of its assessment in a bid to balance the desire for a more meaningful and useful exam with concerns about the amount of time spent on testing.

The decision by the **Smarter Balanced Assessment Consortium** reflects months of conversation among its 25 state members and technical experts and carries heavy freight for millions of students, who will be taking it in two years. The group is one of two state consortia crafting tests for the Common Core State Standards with \$360 million in federal Race to the Top money.

From an original design that included multiple, lengthy performance tasks, the test has been revised to include only one performance task in each subject—mathematics and English/language arts—and has been tightened in other ways, reducing its length by several hours.

The final blueprint of the assessment, approved by the consortium last week now estimates it will take seven hours in grades 3-5, 7½ hours in grades 6-8, and 8½ hours in grade 11.

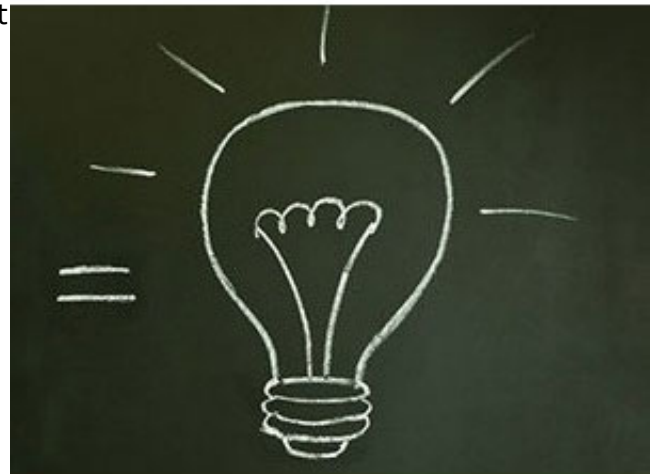
Earlier this fall, states' worries about too much testing time had prompted the group to offer a choice: a "standard" version of the assessment—6½ to 8 hours—or an "extended" one, which would run 10½ to 13 hours, with more items to facilitate more-detailed feedback on student performance.

Persistent doubts about that plan, however, led to further discussions and a decision to expand the shorter version by about 30 minutes and make it the only one offered, consortium officials said.

While many states saw value in having more performance tasks on the test, the amount of information they could yield didn't justify the additional testing hours, said Carissa Miller, the deputy superintendent for assessment, content, and school choice in Idaho, and the co-chairwoman of the SBAC executive committee. Including even one such task—which requires students to tackle longer, more complex math problems and write essays based on reading multiple texts—represents a major improvement in most states' assessment systems, she said.

[← Back to Story](#)

EDUCATION WEEK
Multi-User
LICENSING



**THE FULL
EDUCATION
PICTURE...**
...for up to 60% less

+ Click here for more info

"It's a precarious balance between having a test that we get all the measurement pieces we need, and having it be so long that it becomes impractical," she said. "Having even one very authentic performance task, [with] how much that will change instruction in states that have not had those kinds of things in the past. I think we really came to a sweet spot."

Drilling Down

A key push in the latest redesign was to ensure that the test yields enough detailed information to enable reports on student performance in specific areas of math and English/language arts, Smarter Balanced officials said. The U.S. Department of Education, in particular, pressed for that, said Joe Willhoft, SBAC's executive director. And the consortium's technical-advisory committee had persistent concerns about a pared-down test's ability to report meaningfully on student, as opposed to classroom- or district-level, performance, SBAC leaders said.

The final version will yield overall student scores in math and in English/language arts, by four levels of performance and on a yet-to-be-designed scale, Mr. Willhoft said. It will also produce student-level scores in three areas of math—concepts and procedures, communicating reasoning, and problem-solving/modeling/data analysis—and in four areas of literacy—reading, writing, listening, and research, he said.

In the earlier, "standard" version of the test, some of those areas were combined, making it hard to judge those aspects of students' performance. Adding more items and shifting their distribution allows the test to gauge students' skills in each area, Mr. Willhoft said, while time was managed by scaling back performance tasks and reducing the length of some reading passages.

Still, some experts see the resulting reports as being of disappointingly little instructional value.

W. James Popham, an assessment expert who serves on the Smarter Balanced technical-advisory committee, said tests can provide meaningful information only if teachers and students get more fine-grained feedback than an overall score in writing or in math "concepts and procedures."

"It's still too broad," he said. "No one can ferret out what students need help with. For Smarter Balanced to make a real contribution, it has to make certain that its other two pieces, the interim and formative assessments, are instructionally focused, so educators can do something with the results."

The evolution of the Smarter Balanced assessment showcases a persistent tension at the heart of the purpose of student testing, some experts say.

"Is it about getting data for instruction? Or is it about measuring the results of instruction? In a nutshell, that's what this is all about," said Douglas J. McRae, a retired test designer who helped shape California's assessment system. "You cannot adequately serve both purposes with one test."

That's because the more-complex, nuanced items and tasks that make assessment a more valuable educational experience for students, and yield information detailed and meaningful enough to help educators adjust instruction to students' needs, also make tests longer and more expensive, Mr. McRae and other experts said.

RELATED BLOG

The Right Balance

What Smarter Balanced did, he said, was to compromise on obtaining data to guide instruction in order to produce a test that measures the results of instruction. As a strong supporter of accountability, that's an approach Mr. McRae supports. It's also crucial to have data that guide day-to-day instruction, he said, but that should come from separate formative and interim tests.

That's what SBAC has in mind, said Mr. Willhoft. Its end-of-year, summative tests will measure results for accountability, and those can shape what schools and districts do long term, he said.

"I'm not convinced that the end-of-year summative assessment used for accountability could be imagined to be extremely instructionally useful," Mr. Willhoft said. It's the interim and formative pieces of its system, he said, that have the potential to affect day-to-day instruction in profound ways.

The plan is to have thousands of test items and tasks in an online "bank" teachers can draw from to custom-design interim tests on specific standards. Also available will be a bank of "formative" tools and strategies to help them judge and monitor students' learning as they go along, Mr. Willhoft said. That three-pronged approach—summative, interim, formative—makes up the "balanced" suite of tests many have sought, he said.

The final test design, with a mix of multiple-choice, constructed-response, technology-enhanced, and performance items, is a big improvement over the exams most states have now, said Deborah V.H. Sigman, California's deputy superintendent of public instruction and a member of SBAC's executive committee.

"We have a summative assessment that signals to the world that there are different ways to measure what students are learning and can do," she said. "That's a huge benefit."

Coverage of the implementation of the Common Core State Standards and the common assessments is supported in part by a grant from the GE Foundation, at www.ge.com/foundation.

WEB ONLY

