

WHAT DOES A HIGH-QUALITY ASSESSMENT ACTUALLY LOOK LIKE?

Advice for Policymakers in Choosing New State Tests Truly Aligned with New State Standards

(April 2014)

PURPOSE

Nearly every state has updated its K-12 education standards to focus them on the skills and knowledge students will need for success after high school—whether it’s pursuing some form of college or training or going directly into the workforce and seeking a living-wage job. Now, to match the aspirations embedded in these standards, state leaders need to put in place an assessment system up to the task—one that focuses much more on problem-solving and writing, provides timelier feedback to educators, and is fairer for all students. Most current state assessments—with their reliance on multiple choice—cannot do all these things well, and most states have banded together in one of two consortia to co-design shared, higher-quality assessments that are better and less expensive than what states could create on their own.

But what does a “high-quality” assessment really look like? How can state leaders be sure the assessments they ultimately choose—whether one of the assessments developed by the state-led consortia, an off-the-shelf option or a customized state-specific test—truly reflect the new standards and take advantage of all we know about the best assessment practices? The Council of Chief State School Officers (CCSSO) recently released [Criteria for Procuring and Evaluating High-Quality Assessments](#) to help state leaders maintain a high bar for assessment quality. This document summarizes the CCSSO criteria, as well as resources from other assessment experts on these issues—with the goal of helping policymakers understand the issues to focus on, to expect vendors to explain and to prioritize in selection decisions.

Of course, great teaching and strong supports for students matter most in how well students are learning, but clear standards and useful assessments that actually measure the standards provide the needed foundation for success. They signal what’s important and how well students are doing. That’s why states’ decision about which new assessment system to use demands scrutiny and specificity.

WHICH CRITERIA SHOULD STATE LEADERS PRIORITIZE TO GAUGE WHETHER A PROPOSED ASSESSEMENT IS ACTUALLY HIGH QUALITY?¹

1

Does it Measure the Important Shifts and Content in the State’s Standards?

The assessment should measure the full intent of the standards, including the main “shifts” that make the new standards—with their focus on college and career readiness as the goal—most different from older state standards.

The major content, skills and knowledge new assessments should measure are summarized below; (for more detail about what content really needs to be different in new tests, see page four).

¹ These criteria also include ideas and content from Linda Darling-Hammond, et al.’s [Criteria for High-Quality Assessment](#), Student Achievement Partners’ [Assessment Evaluation Tool](#) and Brookings’ [Standardized Testing and the Common Core Standards](#).

To effectively measure English language arts/literacy standards, high-quality assessments should:

- Include a balance of literary and informational texts of appropriate complexity for the grade assessed.
- Include texts from multiple genres (e.g., mystery, biography) and forms (e.g., poem, short story, news article) appropriate for the grade.
- Demand that students read carefully and deeply and use specific evidence from texts to obtain and defend correct responses on virtually all reading and writing items.
- Include different forms of writing tasks (e.g., exposition, argument, narrative), and require students to respond to texts, draw on textual evidence, and support valid inferences from texts.
- Require students to demonstrate real-world research and inquiry skills by having them find, process, synthesize, organize, and use information from various sources (e.g., book excerpts, websites).
- Measure the speaking and listening communication skills students need for college and career readiness.

Make sure the assessment tests writing with tasks that require students to use evidence from texts.

To effectively measure mathematics standards, high-quality assessments should:

- Focus strongly on the content most necessary for success in later mathematics. This means that assessments focus primarily on the major work of the grade or course: arithmetic in the elementary grades, and ratio, proportional relationships, pre-algebra and algebra in grades 6-8.
- Reflect a developmentally appropriate progression of content from grade to grade and course to course.
- Measure, with a balance of test items, a student's understanding of mathematical concepts (e.g.,

Make sure math tests require students to solve some problems, such as multiplication, quickly while also expecting students to work through other problems using multiple steps and multiple skills and concepts.

fractions), fluency to complete a process accurately and quickly (e.g., multiply single-digit numbers), and ability to apply knowledge to real-world problems (e.g., calculate a five percent pay raise).

2

Does it Require a Range of Cognitive Demand?

To effectively measure the state's college- and career-readiness standards, the assessment cannot *only* assess students' basic knowledge or lower-level cognitive skills. The assessment should allow students to demonstrate a range of higher-order, analytical thinking skills required by the state's standards. High-quality items and a variety of question types should be strategically used to appropriately assess the

To measure the new standards well, some experts have suggested that as many as half of the test items in both ELA/literacy and math should tap higher-level thinking skills, such as hypothesizing, analyzing and synthesizing as opposed to simply recalling or summarizing.

standard(s), such as multiple choice, short and extended open-ended, technology-enhanced tasks, multi-step problems in mathematics, and performance tasks in ELA/literacy.

3

Will it Show Whether a Student is Making Progress toward the Goal of College- and Career-Readiness?

The assessment in grades 3-8 should be designed to measure mastery of the standards and students' readiness for the next grade. High school assessments should measure whether students are on-track to succeed in college and careers, and to be used by colleges for determining placement of first-year students into credit-bearing (versus remedial) courses. It should provide an accurate indication of achievement for students at all levels, including the very high and very low. Once the assessment has been administered for multiple years, it should be validated as to whether it truly predicts success in college and careers.

Scores at or above "proficient" should reliably predict that students are—or are on track to being—ready to succeed in college or a career by the time they graduate from high school.

4

Was it Developed with Significant Input from both Educators and State Leaders?

States should have substantial input into the design of their new tests, including from state education agency leaders and especially from those who work most closely with students, such as K-12 educators, college/university faculty and experts in career and technical education. Educators should inform the test blueprints and should be involved in test item development and review prior to full implementation. The assessment should be field tested before full implementation, and states should have the option to participate in decisions about the operational assessments based on field test results.

Additionally, educators should participate in a research-based process to set the college- and career-ready cut scores.

Those closest to students should have a role during every step of the test development, implementation and evaluation process.

5

Are the Test's Purposes and Design Transparent?

Many sample items and practice tests in each grade/subject should illustrate the assessment's content coverage and depth. The assessment developers should make publicly available the documents needed to understand the assessment's purposes, expectations and uses. For example, assessment blueprints and item specifications enable states to know what standards will be measured, and what choices the assessment developers have made—in terms of content assessed, item types, the range of cognitive demand, etc.

Policymakers, educators, parents and students should all have access to sample test items and plans for scoring and reporting.

6

Does it Allow for Comparability across States?

The assessment should allow educators to compare students' performance across states—and even internationally—and quickly assess and strengthen the abilities of students when they move to a new area.

7

Has it Been Reviewed by External Assessment Experts?

The assessment developers should be willing to put their assessments to the test. Each assessment system should undergo a rigorous review during development and after implementation by independent state and national experts in assessment—both psychometricians and content experts.

State leaders should not just trust but verify. External experts should be asked to confirm how well a test actually meets its promises for content included, fairness and rigor.

8

Will it Provide Timely and Useful Information on Student Learning?

The summative assessment reports should be easy to understand and delivered in time to provide useful, actionable data to students, parents, and teachers. The reports should clearly indicate each student's annual progress toward being ready for the next grade (for grades 3-8) or college- and career-level work (for high school assessments), and could include appropriate instructional actions for students, parents and educators to take based on student performance. To be most instructionally valuable for educators and students, a state assessment system should include non-summative diagnostic or interim assessments with data that are shared back during the school year.

Both educators and policymakers need to see the test reports as helpful tools for measuring student learning and adjusting instruction.

9

Does it Protect Student Privacy and Data?

The assessment must meet all federal and state student privacy laws and requirements. All assessment data must be owned by the states, which should have clear agreements with testing vendors as to how the data are transferred and used.

Appendix: More Detailed Evidence for Common Core English Language Arts and Mathematics

Because the Common Core standards—and other state standards designed with the goal of ensuring students are learning what they need to succeed in college and careers after high school—expect slightly different skills and knowledge and expect students to master content at different grade levels than prior state standards, it's critically important for state leaders to ensure any new assessment truly measures these shifts. Specifically, here's the most important content state leaders should be sure to look for on any new test in elementary, middle and high school.

To effectively measure English language arts/literacy standards, high-quality assessments should:

- Include a balance of literary and informational texts of appropriate complexity for the grade assessed.
 - In grades 3-8, approximately one-half of the texts that students encounter on the assessments must be literary, and approximately one-half are informational. In high school, approximately one-third of the texts are literary and two-thirds informational.
- Represent the balance of the genre and text characteristics called for in the standards.
 - Assessments texts have the appropriate level of text complexity for the grade, as determined by both quantitative and qualitative analyses.
- Demand that students read carefully and deeply and use specific evidence from texts to obtain and defend correct responses on virtually all reading and writing items.
- Include different forms of writing tasks (e.g., exposition, argument, narrative), and require students to respond to texts, draw on textual evidence, and support valid inferences from texts.
 - In grades 3-8, approximately one-third of writing tasks are expository, one-third are opinion or argument, and one-third are narrative.
 - In high school, approximately two-fifths of writing tasks are expository, two-fifths are argument and one-fifth may be narrative.
 - In all grade levels, expository and argument writing tasks place a premium on writing to sources, requiring students to use evidence from texts to present careful analyses and well-defended claims. As students progress through grades, imaginative narrative prompts should be increasingly text-based, with narrative description (text-based, chronological) writing being dominant in the one-fifth of high school writing that is narrative.
- Include a significant focus on real or simulated research tasks that require students to write to

sources and analyze evidence from diverse sources and formats.

- Simulated research tasks comprise a significant percentage of the total number of points on reading assessments. Research tasks must always require that students write to sources and require students to analyze evidence from diverse sources and formats.

To effectively measure mathematics standards, high-quality assessments should:

- Focus strongly on the content most necessary for success in later mathematics. This means that assessments focus primarily on the major work of the grade or course.
 - In grades 3-5, at least 75 percent of the total assessment points align to arithmetic, which includes both mathematical understanding of and fluency with operations involving whole numbers and fractions. Topics like statistics, probability and congruence, which are not major work in elementary school, should be assessed in later grades.
 - In grades 6-8, at least 65 percent of the total assessment points align to ratio, proportional relationships, pre-algebra and algebra.
 - In high school, at least 50 percent of the total assessment points align to widely applicable prerequisites for postsecondary work, including algebra, functions and rich modeling applications.
- Measure, with a balance of test items, a student's understanding of mathematical concepts (e.g., fractions), fluency to complete a process accurately and quickly (e.g., multiply single-digit numbers), and ability to apply knowledge to real-world problems (e.g., calculate a five percent pay raise).
- Connect mathematical content with mathematical practices or processes, such as modeling with mathematics.