

COLLEGE-READY WORK MONOGRAPHS

# What's Next?

The Assessment Challenges Facing States



1954, 1983, 2002, 2010...

BILL & MELINDA  
GATES *foundation*



# What's Next?

## The Assessment Challenges Facing States

States now have an unprecedented—and welcome—opportunity to rebuild their assessment systems. With the arrival of next-generation academic standards, a state-led effort of the National Governors Association and the Council of Chief State School Officers, the federal government is investing \$350 million to improve state assessments.

The new common core standards establish what students at every grade level need to know on their way to graduating high school prepared for college or the workplace. But it will be state assessment systems that define how students and schools actually meet the new standards. And these assessment systems should provide students and teachers with the feedback required to improve and succeed.

We at the Bill & Melinda Gates Foundation hope states develop assessments that are every bit as good as the new standards. And working with our partners, we intend to help.

**Carina Wong**  
*Deputy Director, Education*  
*Bill & Melinda Gates Foundation*



## A View from One State

*Building—or buying—state assessments is difficult work, and, despite ambition and effort, it often produces mediocre tests. I've seen that frustration firsthand.*

*From 2003 to 2005, I was the bureau director for assessment and accountability at the Pennsylvania Department of Education. It was an interesting time in a place ready for change: we had a new governor with ambitious goals for both early learning and high school graduation, a newly reauthorized Elementary and Secondary Education Act in the form of No Child Left Behind (NCLB), and a minimal state accountability system—only a basic watch list for the lowest-performing districts.*

*Pennsylvania's academic standards were unfocused and its state tests disorganized. The eighth-grade test, for example, covered material from as far back as the third grade. The state needed a more rational, teachable system. But changing the standards would have been a long political process. Instead, we prioritized a list of “fewer, clearer, higher” standards for test items. We also adopted a new reporting system and looked at aligning the assessment system with district curricula.*

*Our approach may not have been elegant, but we had few choices. Because the department finalized tests far in advance of their release, changes were expensive and difficult to make. There was little capacity in the State Department to design an assessment system. Most of the testing work was contracted out to an independent company. A Technical Advisory Committee monitored*

test development from a psychometric perspective and provided some—but not enough—guidance on policy and instructional issues.

And as much as we wanted to improve tests, we also had to limit the scope of our changes. The targets had to remain stable so teachers could teach to the standards and the assessments. And we had to align the timing of each test to the pace at which teachers followed the curriculum. We wanted teachers to trust the tests, but every change signaled to them that the tests weren't good enough.

We were able to make only a few structural changes to the tests, like adding open-ended items, and we made them more transparent, clarifying their targets and the content's relative weight. The new test focused on core standards, reducing by a third the number of standards assessed. But ultimately we were making mostly cosmetic changes. Most disappointing to me, our tests inadequately measured higher levels of learning, like problem solving and mathematical reasoning.

*“There’s a big push to differentiate instruction, but the tests don’t differentiate. If we’re being asked to present our information in a different way, then the state’s going to have to present its tests in a different way.”*

—Middle school teacher,  
quoted in *Primary Sources:  
America’s Teachers on America’s Schools*,  
Scholastic, 2010.



# Opportunities & Challenges

*The group I now manage at the Gates Foundation focuses on standards, assessments, and the tools that teachers need to succeed. We know that an ever-growing tangle of policies will frustrate educators and generate assessment systems that don't address what's needed. That's why we encourage the coalitions that are producing standards and assessments that effectively ensure our ultimate goal: high school graduates who are ready for college or careers. We want to help willing states create the best-possible assessment systems. Our perspective identifies five essential commitments these states must make:*

## 1) Build next-generation assessment systems with:

- **Valid assessments that measure complex learning**  
Most teachers don't believe state tests accurately measure students' knowledge and abilities. But states can begin to address these teachers' concerns. There should be summative assessments that resemble challenging classroom work. Tests should also measure college readiness and reflect the values of higher education; when tests are coupled with university articulation agreements, for example, they gain credibility. States should seek to build tests that better measure academic content and research-proven strategies for student success in college. Finally, states should work to adequately assess deeper aspects of complex learning, including adaptive problem solving, critical reasoning and decision making, communicating, learning how to learn, collaborating, and assessing risk.

- **New test designs**

A single high-stakes assessment at the end of the school year—or at the end of high school, when there is no time for remediation—is unlikely to measure college readiness. Coursework should count—instructional units could have modular assessments that add up to proficiency credits. End-of-course tests could measure what students learn over two years rather than only one year, a practice common in Europe and employed by the International Baccalaureate model.

- **Distributed accountability**

Students who are not proficient must get enough instructional time to truly improve their performance. To achieve reading and writing proficiency, students need constant and

consistent supports. The common-core standards intentionally push literacy beyond English-language-arts instruction. If students' literacy scores comprised summative pieces from other classes, like social studies and science, teachers would have incentives to emphasize literacy in those subjects as well.

- **Alignment with teacher- and student-accountability systems**

States need to define how new tests fit into their accountability systems—to determine which tests “count.” Also, states and districts moving toward better models of measuring effective teaching and using value-added systems must honor fairness to teachers.



## 2) Link standards, assessments, and instruction by supporting:



- **High-quality formative assessments**

NCLB focused on summative assessments *of* learning, and Race-to-the-Top (RTTT) encourages formative assessments *for* learning. Ideally, formative assessments are embedded in the curriculum and actually guide the design of the summative assessments; the two forms of assessment should be intertwined. States will need to ensure the quality and alignment of these new formative assessments.

- **Delivery platforms**

States should support instructional platforms that provide examples of high-quality formative assessments aligned to state tests. Often, states fail to provide

the information districts need: real-time data that teachers can use to adapt their instruction. States can help develop these instructional platforms in district systems and link them to state databases.

- **Syllabi for courses**

There are various ways to organize and sequence the new college-ready standards. States and districts can help by developing syllabi across multiple courses with a clear assessment framework built in. These syllabi would communicate to teachers and students exactly what should be taught, what the district expects them to teach or to learn, and how students' learning will be measured.

### 3) Create common assessments through several collaborations:

- **State partnerships**

States have the power to join assessment consortia, and to decide which, if any, assessment systems they will change. Most states have already joined at least one consortium that is competing for federal funding through RTTT. Collaborating on development and pooling states' purchasing power would create efficiencies. But what should the nature of state collaboration be? How would procurement rules need to shift, and can we develop common specifications? What kinds of waivers would states need from NCLB or its suc-

cessor during the design phases? These are tough questions. Our goal is to help states develop good answers.

- **Common proficiency levels**

If the states adopted a common framework for performance, we would know whether proficiency in Massachusetts equaled proficiency in Florida. Common assessments would allow reliable comparisons between states, even if those states administered different tests. But if we don't adopt common proficiency levels, we risk continuing the wild variation


in our national idea of "how good is good enough."

- **Shared test items and an item bank**

It's more financially efficient to purchase common test items that create "anchors" in the different assessment systems. States will have to agree on the qualities, types, and range of the anchor items.







## 4) Create real pathways to college by:

- **Starting in middle school**

We know some students need more time to get college- and career-ready. States should consider which initiatives should begin early to ensure all students meet standards.

- **Creating proficiency-based pathways**

New test designs and technology can converge in proficiency-based pathways that enable students to achieve the standards when and however it works best for them.

- **Making agreements with higher education**

Ultimately, state institutions of higher education must agree on the new assessments, and on articulations, like the Advanced Placement system, that allow students to go straight into credit-bearing classes or receive college credit when they enroll.

- **Expanding early college and dual enrollment**

These options build a clear pathway into college during the high school program.



## 5) Maintain the system with:

- **Ongoing validation**

Both the standards and the assessments need to be validated. States and the various assessment consortia must develop an evidence-based plan for revising the standards and assessments.

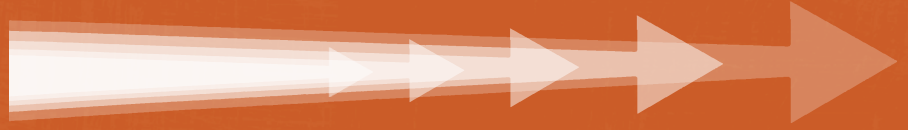
- **Intermediaries monitoring the quality of implementation**

States will probably need intermediary organizations to ensure that they are maintaining their assessments properly.

- **Professional-development systems**

New forms of assessment will require professional development. States need to develop coherent, transparent strategies for reaching teachers. We believe technology can play an important role here.

# Moving Forward



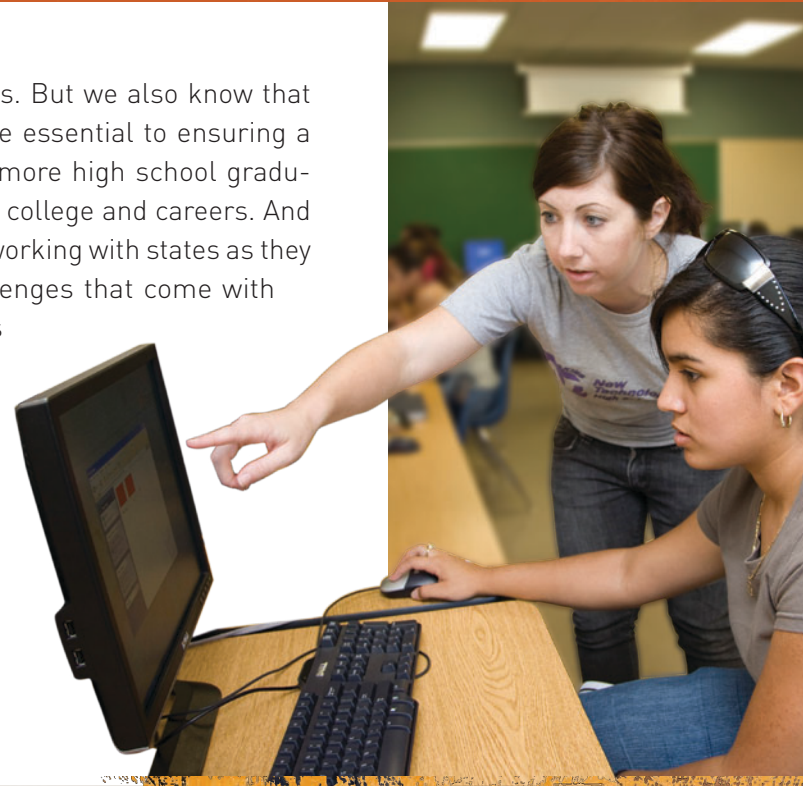
For the last several years, the Bill & Melinda Gates Foundation has been thinking strategically about how states can take common standards to the next level. It has tried to anticipate the challenges so it can help states build the next-generation assessment and instructional systems it thinks will be crucial to implementing the college- and career-ready standards. It plans to invest more than \$350 million in areas that include:

- helping states build frameworks that could support a common assessment strategy across states
- building prototypes of both formative and summative assessments in math and literacy that align to the standards, challenge students, and help teachers give meaningful feedback
- building prototype syllabi for courses that connect the standards, assessments, and instruction
- creating new intermediaries for validation, item-bank development, as well as professional development
- developing specifications for new technology-based instructional platforms that allow states to deliver high-quality assessments and assignments aligned to the standards, and that provide districts with real-time data that can influence instruction
- identifying new ways of thinking about the psychometric rules that guide tests so that higher-quality items can be used for large-scale assessments
- creating new scoring technology and new forms of diagnostic assessments

- developing improved academic supports for students by expanding our research base in this area and identifying new strategies to accelerate learning through technology
- scaling existing programs, and funding research focused on more effective learning by students
- partnering with states and districts to help them implement next-generation assessment and instructional systems with the features mentioned above

We know it's difficult for states to build and maintain substantially different as-

essment systems. But we also know that these systems are essential to ensuring a future where far more high school graduates are ready for college and careers. And so we commit to working with states as they take on the challenges that come with new assessments tied to common standards.





SHARE:    

©2010 Bill & Melinda Gates Foundation.  
All Rights Reserved. Bill & Melinda Gates  
Foundation is a registered trademark in  
the United States and other countries.

[www.gatesfoundation.org](http://www.gatesfoundation.org)

BILL & MELINDA  
GATES *foundation*