



Transforming Statewide High School Assessment Systems: A Guide for State Policymakers

Achieve

November 2008

About Achieve

Created by the nation's governors and business leaders, Achieve is a bipartisan, non-profit organization that helps states raise academic standards, improve assessments and strengthen accountability to prepare all young people for postsecondary education, careers and citizenship. Achieve has helped more than half the states benchmark their academic standards, tests and accountability systems against the best examples in the United States and around the world. Achieve also serves as a significant national voice for quality in standards-based education reform and regularly convenes governors, CEOs and other influential leaders at National Education Summits to sustain support for higher standards and achievement for all of America's schoolchildren.

In 2005, Achieve co-sponsored the National Education Summit on High Schools. Forty-five governors attended the Summit along with corporate CEOs and K–12 and postsecondary leaders. The Summit was successful in making the case to the governors and business and education leaders that our schools are not adequately preparing students for college and 21st-century jobs and that aggressive action will be needed to address the preparation gap. As a result of the Summit, 34 states have since joined with Achieve to form the American Diploma Project Network—a coalition of states committed to aligning high school standards, assessments, graduation requirements and accountability systems with the demands of college and the workplace.

For more information, visit Achieve's Web site at www.achieve.org.

Copyright © November 2008 Achieve, Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information or storage retrieval system, without permission from Achieve.

Table of Contents

| | |
|--|---------|
| Why States Need To Transform High School Testing: A New Vision for Assessment..... | Page 4 |
| A New Vision for the High School Assessment System: Four Core Principles for College and Career Readiness..... | Page 6 |
| Core Principle One: “Proficient” Should Mean Prepared: Building College- and Career-Ready Anchor Assessments..... | Page 11 |
| Core Principle Two: Making Tests Matter Beyond K–12: Opening Doors to Higher Education and Employment Using High School Tests..... | Page 25 |
| Core Principle Three: Not Just More Assessment—Better Assessment: Measuring the Full Range of College and Career Readiness Skills and Supporting Richer Instruction..... | Page 32 |
| Core Principle Four: Streamlining Testing Time and Costs..... | Page 41 |
| Getting from Here to There: Key Questions to Inform State Action | Page 44 |
| A “Decision Tree” for State Policymakers..... | Page 48 |
| Acknowledgments..... | Page 53 |
| Endnotes..... | Page 54 |

Why States Need To Transform High School Testing: A New Vision for Assessment

Assessment has become an integral part of the education enterprise over the past two decades—and for very good reasons. Schools need common measures of student performance that can guide their work and that are well-aligned with the publicly articulated standards that students are expected to meet.

Yet current testing systems fall short of where we need them to be, especially in high school. This is in part due to the uneven quality and rigor of state assessments, but it is also a result of a relatively narrow vision of assessment that most states and districts have deployed.

First, state tests tend to measure too narrow a slice of what is most important for students to know and do. Achieve has conducted extensive analyses of state assessments and found that many high school assessments (particularly those used as exit exams) stop well short of assessing the advanced reasoning and problem-solving skills required for success in the college classroom or the workplace.¹ Though many state tests serve current purposes of testing well—such as holding schools accountable for reaching minimum standards of proficiency or certifying that students have basic skills to earn a diploma—the tests too often are not able to measure more advanced concepts.

As a result, students can score a “proficient” on these exams and still be unprepared for the challenges they will face after high school.

Second, in too few states do higher education leaders and employers actually use student results to admit, place or hire graduates. Until high school assessments open doors to postsecondary opportunities, students will continue to view them as irrelevant to their futures.

The third problem with current assessments is the limited information they provide back to schools and teachers to help guide and improve instruction. Statewide assessments are primarily paper-and-pencil tests given at the end of the school year. While these tests serve very important purposes, and should continue to play a major role in state assessment and accountability systems, they are not always as useful to classroom teachers as other forms of assessment.

The other—and perhaps most troubling and difficult—problem to address is that there simply may be too many tests. High school students take a lot of tests—some required by states, some required by colleges and still others required by districts. Too many of these tests are disconnected from one another and some of them are duplicative.

Overview and Context

In 2005, Achieve launched the American Diploma Project Network (ADP) to help K–12, postsecondary and business leaders collaborate to ensure that all students graduate with the skills and knowledge required for success after high school. There are now 34 states in the ADP Network working toward the goal of college and career readiness for all high school students—and, to date, 22 of these states have aligned standards with the demands of postsecondary education and careers and 20 states have adopted meaningful high school graduation requirements.

That is significant progress in just three years. Yet even the states that are furthest along are, at best, only halfway to the goal. Only ten states have tests in place that are capable of assessing whether students are ready for higher education and employment—and that are in

fact used by higher education. Only nine states have data systems that connect high schools and higher education and only four states have made college and career readiness the centerpiece of their accountability systems for high schools.²

States cannot afford to stop short of measuring and holding the system accountable for college and career readiness. Relying only on standards and graduation requirements to ensure that all students are ready for college and careers is not enough, especially because few state tests today tell high school students if they are prepared for their next step.

Achieve and the Education Trust recently published a major report resulting from a yearlong effort to distill lessons learned from states that want to make high school standards, curriculum, assessment and accountability policies more coherent and meaningful. That volume is intended not as a recipe for improving high schools but instead to offer ideas and key questions to consider in getting high school expectations right. The report, called *Measures that Matter: Making College and Career Readiness the Mission for High Schools*, offers suggestions and questions for policymakers to consider as they review the statewide framework for standards-based reform across five major areas:

- ⇒ Set a Clear Goal: Align High School Standards with the Demands of College and Careers
- ⇒ Assure that Students Enroll in a Course of Study Aligned with College and Career Readiness Standards
- ⇒ Provide High-Quality Curriculum and Teacher Support Materials
- ⇒ Measure Student Learning: A College- and Career-Ready Assessment System
- ⇒ Get Everybody Pulling in the Same Direction: An Information and Accountability System Focused on College and Career Readiness

States face lots of questions and choices—and, frankly, pushback—in the area of assessment. Achieve has developed this policy guide as a companion to the larger report in order to support states with additional advice and analysis.³

The goal of this guide on assessment is to help states weigh the policy choices they face in:

- ⇒ developing assessments of college and career readiness;
- ⇒ using those assessments to open doors for students to higher education success;
- ⇒ refining the system of assessments given in high school to provide tools for educators to enrich student learning; and
- ⇒ reviewing the system of high school assessments as a whole.

A New Vision for the High School Assessment System: Four Core Principles for College and Career Readiness

States need a new vision for high school assessment. In the testing systems we envision, all students will be asked to demonstrate in multiple ways that they meet college and career readiness standards—and the state will help students and schools get there. States will value interim and performance assessments as much as they value statewide paper-and-pencil summative tests. And students, parents, educators, administrators, policymakers and the public will understand why all the tests are valuable and what the results tell (and don't tell) about student learning and school success.

As states evolve their high school assessment systems to meet this new vision, they should include a combination of statewide and local measures, including:

- ⇒ large-scale “anchor assessments” that are pegged to college and career readiness standards for the end of high school; that can be used by higher education to place students into credit-bearing, freshman courses and that can be used to hold high schools accountable;
- ⇒ performance assessments, including constructed-response questions on the anchor assessments and student projects, demonstrations, and tasks that are best administered locally, to measure the full range of college and career readiness standards and potentially expand the measures used for accountability; and
- ⇒ interim assessments in core high school subjects and aligned to college and career readiness standards that are designed to measure student, classroom, school and/or district progress toward meeting college and career readiness standards and provide quick feedback to educators and administrators.

Many states have a K–12 testing system that includes some of these measures. But very few states have a system with all of these measures—and that aims for college and career readiness.

To realize this new vision for assessment, Achieve recommends that states focus on four core principles. This policy guide is organized around these four core principles. They are not the only ideas that should guide state decision making around testing—but we believe they are the elements most noticeably absent from high school assessment systems today.

Core Principle One: “Proficient” Should Mean Prepared: Building College- and Career-Ready Anchor Assessments

State assessments at the high school level must do a better job measuring real-world knowledge and skills that students will need to be successful after high school. In the new system that Achieve recommends, the most important and most visible large-scale assessments will measure academic preparation for postsecondary education, training and employment.

- ⇒ States should select, modify or build college and career readiness assessments to “anchor” high school testing so that “*proficient*” means *prepared* for students’ next steps. These “college and career readiness anchor assessments” should be given to all students, statewide, near the end of high school. The entire K–12 assessment system should be aligned with the anchor assessments, including any other assessments given statewide in earlier grades in high school, and the anchor assessments should be given strong weight in the state accountability system.

Core Principle Two: Making Tests Matter beyond K–12: Opening Doors to Higher Education and Employment Using High School Tests

Simply giving these tests statewide is not enough; in a new system, the results of the high school college and career readiness anchor assessments should open doors for students to higher education and employment.

- ⇒ States should use results from the anchor assessment not only to signal if students are ready for college and careers but also to place students into college classes or to waive job entrance tests. This will require *more collaboration, from start to finish, with college faculty and employers* in the development of, standards-setting on and use of high school tests.

Core Principle Three: Not Just More Assessment—Better Assessment: Measuring the Full Range of College and Career Readiness Skills and Supporting Richer Instruction

New statewide anchor assessments should have real meaning for students and schools. But even if they are excellent tests, they shouldn't be the only tests that matter.

- ⇒ States should work with districts to put in place a robust state and local system of assessment that promotes richer and more relevant teaching and learning. This will require *better summative tests, more performance-based assessments and high-quality interim assessments* aligned with state standards of college and career readiness.

Core Principle Four: Streamlining Testing Time and Costs

In the new system Achieve recommends, states and districts should work together to streamline the overall amount of testing students experience—determining which tests really matter, which are nice but not essential and which are extraneous and can be eliminated.

- ⇒ States should take stock of which tests students are already taking—at both the state and local levels—and what those tests measure, before adding new assessments. *New tests should fill in gaps, not create redundancies.*

Table 1: Common Elements of a College- and Career-Ready Assessment System

This table summarizes the four core principles for a new high school assessment and accountability system—and shows how a mix of different assessments and tests are needed to meet these principles.

| Measures | Core Principles |
|---|---|
| <p><i>What is the optimal mix of state and local assessment to measure college and career readiness?</i></p> | <p><i>What core principles should guide the development of this new system?</i></p> |
| <p>College- and career-ready “anchor assessments”</p> <ul style="list-style-type: none"> ⇒ Pegged to college and career readiness standards for the end of high school ⇒ College- and career-ready cut score signals readiness ⇒ Can be used by higher education to place students into credit-bearing, non-remedial courses ⇒ Statewide, large-scale test given to all students | <p>1. “Proficient” should mean prepared</p> <ul style="list-style-type: none"> ⇒ States should select, modify or build from scratch an anchor assessment in reading, writing and math in 11th grade or in advanced courses to all students. <p>2. High school tests should open doors to higher education and employment</p> <ul style="list-style-type: none"> ⇒ Higher education needs to collaborate with K–12 to develop/modify and truly use the anchor assessment in college-course placement decisions. ⇒ Cut scores should open doors to credit-bearing, freshman-level coursework, as well as to postsecondary or employer-based apprenticeships and training. <p>3. Testing systems should measure the full range of college and career readiness skills</p> <ul style="list-style-type: none"> ⇒ Anchor assessments should be robust, with constructed-response items and quality and quick scoring. ⇒ States also should help stimulate local performance measures—such as graduation projects, writing portfolios, or science experiments. ⇒ States also should identify or build high-quality, aligned interim assessments that diagnose student strengths and gaps. <p>4. Testing should be streamlined</p> <ul style="list-style-type: none"> ⇒ New tests should fill in gaps, not create redundancies. ⇒ States and districts should collaborate to identify tests that can be eliminated because they serve similar purposes or provide similar information. ⇒ States also can reduce testing costs and development time by collaborating across state lines to create and use tests in key subject areas and courses. |
| <p>Performance measures designed to ensure the full range of college and career readiness standards are measured</p> <ul style="list-style-type: none"> ⇒ Includes constructed-response (open-ended) questions on state tests ⇒ Also includes senior/graduation projects, performance measures, projects, tasks, etc. that are best administered locally | |
| <p>Interim assessments for core high school subjects</p> <ul style="list-style-type: none"> ⇒ Aligned to college and career readiness standards ⇒ Designed to measure student, classroom, school and/or district progress toward meeting college and career readiness standards | |

Why Have Different Types of Assessments in the Statewide System?

Clarifying the distinctions among assessments is an important first step in determining how to include such assessments in the assessment system focused on college and career readiness that we envision.

Summative assessment (also called large-scale, on-demand or standardized testing) is remarkably similar across states. The tests generally are given once at the end of a unit of time, such as the semester or school year, in order to evaluate students' performance against standards. They are typically given on a large scale, such as statewide (but they also can be national or district tests). Scores are reported in relation to standards-based achievement levels (e.g., "proficient"). Virtually every student enrolled in the tested grades is required to take these tests, with accommodations available for students with disabilities and English language learners. The scores are used in accountability determinations for schools that involve public reporting and consequences. Summative tests are the least flexible type of assessment because they need to be given under standardized conditions to ensure that the results are fair and comparable for all tested students.

Performance assessment means that the student constructs or generates a response or performance rather than selecting from a set of responses. The performance may be as short as writing down a single number or as extensive as a long essay, an oral presentation, a research project or a creative production. Because performance assessments may vary widely in terms of dimensions such as complexity and integration of knowledge and skills, the suitability of each performance assessment must be considered in the particular context of its intended use. While performance assessments have been recognized for their potential contribution to increasing the validity of assessment, they bring their own challenges in terms of traditional test development and psychometric qualities, including reliability issues. For example, performance assessments may be scored at an acceptable level of reliability if there are appropriate quality-assurance controls for scoring, but scoring will always be less reliable than machine-scored responses, it will be more costly and usually it takes more time. Another drawback is that, if only a small number of tasks are used, the test's validity may be severely limited. Despite some of these tradeoffs, performance assessments may better assess some important content and/or skills that cannot be measured well in a typical statewide summative assessment. They may provide a way to credibly credential student proficiency in addition to Carnegie units or graduation tests and they may do a better job supporting quality instruction and curriculum than summative testing.

Formative assessment is really a process used by teachers and students during instruction that provides feedback to adjust teaching and learning in real-time. The goal is not to measure whether students have mastered a cumulative amount of material; rather, the goal is to improve students' achievement of intended instructional outcomes daily or every few days. Formative assessment is embedded within learning activities and linked directly to current units. The assessments are small-scale (a few seconds, a few minutes, less time than a class period). They often are called "minute-by-minute" assessment or formative instruction. In addition, the tasks may vary from one student to another, depending on the teacher's judgment about the need for specific information about a student at a given point in time. Formative assessment is done by the classroom teacher for the explicit purpose of diagnosing where students are in their learning, where gaps in knowledge and understanding exist, and how to help improve student learning. Providing corrective feedback, modifying instruction to improve the student's understanding or indicating areas of further instruction are essential aspects of formative assessment. It does not make sense to aggregate formative assessment information beyond the classroom.

Interim assessment is the term we suggest for the assessments that fall between formative and summative assessment, including the medium-scale, medium-cycle assessments now used by the highest-performing school districts. Many of the assessments currently in use that are labeled “benchmark,” “formative,” “diagnostic” or “predictive” fall within our definition of interim assessments. Interim assessments (1) evaluate students’ knowledge and skills relative to a specific set of academic goals, typically within a limited time frame, and (2) are designed to inform decisions at both the classroom level and beyond the classroom level, such as the school or district level. Thus, they may be given at the classroom level to provide information for the teacher, but, unlike true formative assessments, the results of interim assessments can be aggregated meaningfully and reported at a broader level. As such, the timing of the administration may be controlled by the school or district rather than by the teacher. This means these assessments are less instructionally relevant than formative assessments. Interim assessments may serve a variety of purposes, including predicting a student’s ability to succeed on a large-scale summative assessment, evaluating a particular education program or pedagogy, or diagnosing gaps in a student’s skills.⁴

Core Principle One:

“Proficient” Should Mean Prepared: Building College- and Career-Ready Anchor Assessments

“Proficient” on high school tests should mean prepared for success after high school. In order to achieve this, states need to give more challenging statewide tests as part of their high school assessment systems—exams that reach far enough into the curriculum to measure the knowledge and skills students need to be ready for college and careers. This can be accomplished either by adding new assessments later in high school or by modifying existing assessments to ensure they measure the full range of state knowledge and skills. We refer to these tests as “college- and career-ready anchor assessments” because they should form the foundation for the high school assessment system.

These anchor assessments should be designed (or adapted) for the explicit purpose of determining if students have gained the academic skills they need in reading, writing and mathematics to enroll and succeed in credit-bearing, non-remedial, postsecondary courses and training after high school. Done right, the tests will signal if students have the academic preparation that is foundational to success in college and good jobs.

Most states have revised or are in the process of revising their high school standards to better align with college and career readiness expectations, such as those defined by ADP.⁵ Such standards should guide the development of the more rigorous anchor assessments or the adjustment of existing tests.

Three Approaches to College- and Career-Ready Assessment

According to Achieve’s research, only 10 states have English and mathematics assessments in high school that measure the advanced knowledge and skills valued by postsecondary institutions and employers and that actually are used by higher education to make judgments about placing students in college courses.⁶

These states are taking three different approaches:

- ⇒ rigorous comprehensive end-of-grade 11th grade tests adapted or developed by K–12 and higher education leaders;
- ⇒ end-of-course tests in advanced courses that are validated for use by postsecondary institutions or employers; and
- ⇒ college admissions tests adapted to align to state standards and become part of the statewide high school testing system.

Chapter Highlights

- ⇒ States should select, adapt or build from scratch a “college- and career-ready anchor assessment” in reading, writing and mathematics.
- ⇒ The anchor assessment should be given in the 11th grade or in advanced courses to all students and used to place students in credit-bearing college courses.
- ⇒ States can choose an end-of-grade test given to all 11th graders; a series of end-of-course exams (such as English 3 or Algebra 2); or an adapted national college admissions test.
- ⇒ Each of these three approaches has strengths and challenges. Combining approaches is possible but policymakers should be cautious about the increased testing burden on students and schools.
- ⇒ Other tests, such as placement tests or career readiness tests, may be useful for some purposes, but they should not be used as the anchor assessment.

The American Diploma Project Benchmarks: What Academic Skills Are Needed for Success in Higher Education and Employment?

The American Diploma Project Benchmarks, released in 2004, describes in detail the mathematics and communications/English skills and content that graduates should master by the end of high school to be ready for college and careers.⁷ College faculty across states and institutions, along with front-line hiring managers in high-performance, high-growth fields, have a fairly consistent view of the rigorous level of reading, writing, communications and mathematics skills that incoming freshmen need in order to be successful in first-year credit-bearing college courses.

In mathematics, high school graduates need to demonstrate the knowledge and skills typically learned in a four-year mathematics sequence including basic through intermediate algebra (Algebra 1 and 2), geometry, data analysis and statistics. They also need sophisticated mathematical reasoning and problem-solving skills that will enable them to confront real-world problems and apply their learning in a variety of contexts.

Graduates also need to be able to write and communicate effectively to different audiences; they need to be able to understand and analyze various types of complex texts. In addition, they need to work in groups, use new media and technology effectively and produce complex research projects. And they need to be able to apply sophisticated analytic and reasoning skills.

Table 2: Which States Currently Administer College- and Career-Ready Assessments That Are Used by Higher Education for College Placement?

Ten states currently administer college- and career-ready tests that also are used by higher education. Many other states have told Achieve they are planning to include or change to college- and career-ready end-of-course tests (see Table 4 on page 15). Adapted college admissions tests also are a trend gaining steam.

| Comprehensive End-of-Grade Tests | End-of-Course Tests | Adapted College Admissions Tests |
|----------------------------------|---------------------|----------------------------------|
| California* | California* | Colorado |
| Georgia** | New York | Illinois |
| Texas** | | Kentucky |
| | | Maine |
| | | Michigan |
| | | Tennessee |

* California's Early Assessment Program is attached to the state's grade 11 California Standards Test and the Algebra 2 end-of-course test.

** Georgia and Texas are shifting to end-of-course tests.

Table 3: Summary Table of the Three Approaches’ Advantages and Challenges

Each of the approaches has strengths and limitations. We recommend that state policymakers study the three approaches and select the approach that will best meet the state’s own needs.

| | Advantages of this Approach | Challenges with this Approach |
|---|--|--|
| Comprehensive End-of-Grade Test | <ul style="list-style-type: none"> • Tests all students at the same time • May be possible to adapt state’s current high school test • Probably not a substantial increase in testing time or costs | <ul style="list-style-type: none"> • Needs to be given at least in 11th grade to fairly assess college and career readiness (10th grade is too early) • May test students on material that was taught much earlier in their academic careers • Lacks content depth compared to end-of-course testing |
| End-of-Course Test | <ul style="list-style-type: none"> • Most closely aligned with state and course academic standards • Assesses students on material they’ve learned recently • Can inform decisions about subsequent classes for the student • Measures quality and consistency of courses • Allows for wider range of stakes and incentives | <ul style="list-style-type: none"> • May increase overall testing time (though state tests could replace local course finals) • Format does not lend itself to assess English classes (may still need comprehensive test) • Does not automatically have credibility with higher education; credibility has to be built through collaboration and research • Requires financial investment to create new tests (although at least one exam—ADP’s Algebra 2 assessment—has low cost because it has been developed with other states) |
| Modified College Admissions Test | <ul style="list-style-type: none"> • Widely known among colleges, families and the general public • Offers national comparability • Opens up college access to all students • May (but not always) cost less than state-developed tests | <ul style="list-style-type: none"> • Unclear alignment to state standards • May incur development and administration costs to augment tests • Does not always include the full range of advanced concepts and skills • May increase testing time in the school day • May not give enough feedback to students and schools about college readiness to inform 12th grade coursework |

Approach One: Comprehensive End-of-Grade (EOG) Tests

Comprehensive end-of-grade assessments are given to all students at the end of a grade—typically 10th or 11th grade—and cover a range of material in a subject that may have been taught in the current or previous grades. Twenty-eight states currently have EOG tests in high school. About half of the states give their tests in 10th grade; all states use them for student and/or school accountability.

Advantages of Comprehensive Tests

States interested in adapting their current comprehensive EOG tests to assess college and career readiness might do so because the current tests already exist as a “known quantity” among high school faculty and students. All students are tested at the same time with the same test, which is very important for ensuring that results are comparable for all students. Using comprehensive survey tests would not increase complexity, testing time or testing costs substantially. Existing development, administration, scoring and reporting procedures could be modified, but they would not necessarily need to be replaced or altered dramatically.

In many ways, this is the most straightforward approach to linking K–12 and higher education testing, yet this approach will likely work only with tests given in the 11th grade. Assessments given in the 11th grade are more likely to be able to tap into higher-level knowledge and skills. And, because 11th grade tests are taken closer to the end of high school, colleges are more inclined to honor scores for placement into credit-bearing college courses.

To Achieve’s knowledge, only three states—**California, Georgia and Texas**—have worked with higher education faculty and leaders to adapt their 11th grade EOG exams to measure the skills and knowledge desired by postsecondary education and to certify whether students are ready for credit-bearing postsecondary courses. What’s most notable about the approaches in California and Texas is the collaboration among postsecondary and K–12 agencies to modify the content of an existing test and to report college readiness results to high school students.

In **California**, the Early Assessment Program (EAP) has taken hold in a big way. The California Department of Education (CDE) and California State University System (CSU) worked together to convene faculty to review carefully the content of the 11th grade California Standards Test (CST) in English and the 11th grade CST and Algebra 2 end-of-course tests in mathematics; supplement the existing CST/Algebra 2 test with “Part B” college readiness items and an extended essay in English to ensure that, taken as a whole, the tests measure the skills and knowledge that are needed in higher education; set college-ready cut scores on the tests; waive placement testing for CSU if students test at the college-ready level; and develop new senior-year courses and modules for students who aren’t college-ready. (For more on the EAP, see the detailed sidebar on page 30.)

In **Georgia**, the Board of Regents of the University System of Georgia and the Technical College System of Georgia collaborated with the K–12 system on both standards and assessment. Higher education faculty participated in the process and came to an agreement on a single cut score that indicates college and career readiness on the Georgia High School Graduation Test. This cut score is set higher than the cut score for student graduation. A similar approach was used in **Texas**. In both Georgia and Texas, the 11th grade test now used is being phased out in favor of EOC tests. Higher education leaders in both states expect to stay involved as the EOC tests are developed and deployed.

Challenges with Comprehensive Tests

Comprehensive tests are straightforward, but they are not without their problems. Because the tests are cumulative—they are administered at a particular point in time, such as 10th grade, and not tied to the completion of a particular course—students may be tested on material they were taught much earlier in their high school careers. Indeed, the tests may even be unfair to students if they are being tested on material they haven’t studied in a few years. In many states, high school EOG tests are designed to test material that could be taught in 8th, 9th, 10th or 11th grades, depending on individual students’ course sequences and on local curriculum choices.

EOG exams by their design cannot go into as much depth in a particular content area as other tests, notably EOC tests, can. And if states only give an EOG test in high school at the end of grade 11, students in grades 9 and 10 will not receive much information about their progress toward meeting college and career readiness standards while there is still time to adjust coursetaking and catch up.

Approach Two: End-of-Course (EOC) Tests

End-of-course assessments are given at the end of a particular course taken in secondary school to measure the content taught in that course (e.g., an Algebra 1 assessment or an English 3 test). Sixteen states have EOC exams in place. Of these 16, only eight have tests in advanced courses. Another 18 states report to Achieve they are planning to put EOC tests in place in the next several years.⁸

Table 4: Many States Are Switching to End-of-Course Tests

If end-of-course tests are to serve as college and career readiness anchor assessments, they need to be given in advanced, college-preparatory courses—at least in Algebra 2 and advanced English (measuring both reading and writing). Eight states have tests at this level in place today. Eighteen more states are planning to do away with their current test or add end-of-course tests.

| States Currently Administering End-of-Course Assessments | |
|--|--|
| End-of-Course Tests in College- and Career-Ready Courses | End-of-Course Tests in Introductory Courses Only |
| Arkansas | Georgia |
| California | Louisiana |
| Hawaii | Maryland* |
| Indiana | Mississippi |
| New York | Massachusetts* |
| Oklahoma | North Carolina |
| Utah | South Carolina |
| Virginia | Tennessee |
| States Planning End-of-Course Assessments | |
| End-of-Course Tests in College- and Career-Ready Courses | End-of-Course Tests in Introductory Courses Only |
| Arizona | Delaware |
| Connecticut | Florida |
| Kentucky | Missouri |
| Maryland* | West Virginia |
| Massachusetts* | Washington |
| Michigan | |
| Minnesota | |
| New Jersey | |
| New Mexico | |
| Ohio | |
| Pennsylvania | |
| Rhode Island | |
| Texas | |

* In addition to their current end-of-course assessments, Maryland and Massachusetts are planning college- and career-ready end-of-course assessments in Algebra 2.

Source: Achieve, *Closing the Expectations Gap* research, 2008.

Advantages of End-of-Course Tests

EOC tests are growing in popularity because they can be tied closely not only to state academic standards but also to the actual courses students take and teachers teach. Rather than measuring cumulative learning across courses and grade levels, EOC tests measure specific learning within each course. Students take EOC assessments as they finish the courses, whether that's in 8th, 9th, 10th, 11th or 12th grade. The other two approaches (EOG comprehensive tests and adapted admissions tests) require students to take the exams at a specific point in their high school career, regardless of when they took certain courses or learned the material on the tests.

EOC tests also can be more sensitive to instruction than grade-level exams, which makes this approach attractive to educators. Because the exams are taken as soon as a student completes a course, teachers can isolate the learning, as well as the gaps in learning, that took place in their courses and adjust instruction for subsequent classes accordingly.

In addition, EOC exams help ensure that all students are exposed to an equally rigorous curriculum. States that are putting new coursetaking graduation requirements in place may find EOC assessments the most attractive testing approach because they can help education officials monitor the quality and consistency of instruction in high school courses. They also can help guard against artificial course title “inflation,” which happens when a course is given a common name, such as Algebra 2, but the content actually included in the course is substantially weaker than the name would imply.

States also can attach a wider range of stakes and incentives to EOC tests than to the other approaches. Some states—including **Georgia**, **North Carolina**, **South Carolina** and soon **Tennessee** and **Texas**—currently require or plan to require that EOC test results count for a portion of the corresponding course grade (usually 15–25 percent). This creates incentives for students to do well on the exams without making them an all-or-nothing graduation requirement. Other states—including **Arkansas**, **Indiana**, **New York** and **Oklahoma**—are using or are considering using EOC tests in introductory courses (such as Algebra 1 and Geometry) as high school exit exams and using EOC tests in upper-level courses to award placement into credit-bearing college courses.⁹

Challenges with End-of-Course Tests

There are challenges associated with EOC assessments, however. Using EOC tests may increase the overall amount of testing required by a state, depending on the number of courses for which the state develops an assessment. (However, it might not increase the amount of testing experienced by teachers and students if redundant tests, such as locally-developed course final exams, are replaced.) More tests also may mean more money needs to be devoted for test development and administration.

Furthermore, if EOC exams are used as part of students' course grades, then tests need to be scored and results turned around swiftly. This need for rapid scoring can sometimes limit the kinds of test questions used to mostly multiple-choice. States have found creative approaches to this dilemma, however; in **Indiana**, results from paper-and-pencil tests are returned within a week and teachers in **New York** score the Regents exams locally at the end of each term.

Using EOC tests to measure English language arts (ELA) is tricky. Few states have true EOC tests that measure particular English courses, such as World Literature or Expository Writing. This is in large part because the majority of states do not specify particular courses or curriculums for ELA that students must take to graduate. Instead, states usually require students to take four years of English in a course sequence commonly entitled English 1 through English 4—the course content of which is determined in a variety of ways at the district, school or classroom levels. This leads to a wide proliferation and wide variety of ELA courses in high school. It would not be practical to assess all of these courses and most states have not wanted to prescribe particular English courses. Thus, even EOC tests in English are more akin to a comprehensive test of reading and/or writing.

Perhaps the biggest challenge with EOC tests is to build and use the assessments in a way that ensures sufficient credibility with higher education faculty, campus and system leaders. Postsecondary faculty and system leaders should be integral partners with K–12 in test development, standard-setting and predictive validity research to determine how best to use the EOC assessments for placement.

The timing of when students take EOC tests also can be an issue. For example, if students take Algebra 2 in the 10th grade—two years before entering college—higher education institutions will still want to determine whether results on that test could be used for placement two years later. (One approach to resolve this particular dilemma, which is most likely to occur with advanced students, is to require continued coursetaking in advanced mathematics through the senior year.)

Despite these challenges, overall, EOC tests may be the approach that is most likely to measure state standards, most closely linked to high school curriculum and instruction, and best-suited to align high school and college curriculum and instruction. With EOC assessments, high school teachers and students will get better information on their progress toward reaching college and career readiness, while state colleges and universities can use the tests in advanced courses for determining postsecondary placement.

New York’s Regents Exams: Serving Multiple Purposes with End-of-Course Exams

New York’s Regents exams are the best-known end-of-course tests in the nation. The state first administered Regents exams in 1878.¹⁰ Until 2000, only students earning the optional college-preparatory Regents Diploma were required to pass the Regents exams. Other students could earn a lesser Competency Diploma, issued at the local level. The state Board of Regents has mostly phased out that option because it did not prepare graduates for work and postsecondary education (though some students can still earn a Local Diploma if they don’t achieve a high enough passing score on the required mathematics Regents test).¹¹

The Regents exams are based on standards for particular Regents-level courses that are widely published. The exams are given during a one-week window each January and June, depending on the course. The assessments include both multiple-choice and constructed-response formats, though the distribution among these item types varies with the assessment.

All Regents assessments are scored at the school by teachers who receive very detailed scoring rubrics. A random sample of assessments is scored at the state level to ensure consistency in scoring. Scoring the Regents is considered to be both professional development and a professional responsibility.¹²

Beginning with the class of 2005, New York requires all students to pass five of the Regents exams as one condition of earning the Regents Diploma. There currently are 18 Regents end-of-course exams measuring different courses in the following subjects: English, mathematics, science, social sciences and world languages. Students take Regents exams only in the subjects in which they are enrolled.

Table 5: Which Regents Exams Are Required for Graduation?

| Subject | Regents Credits Required | Required Exam(s) | Grade in which exams are usually taken (may vary, particularly in accelerated programs; most are normally taken in June, unless otherwise noted) |
|----------------|--------------------------|---|--|
| English | 4 | <ul style="list-style-type: none"> English Language Arts | <ul style="list-style-type: none"> Grade 11 |
| Social Studies | 4 | <ul style="list-style-type: none"> Global History & Geography U.S. History & Government | <ul style="list-style-type: none"> Global - Grade 10 U.S. - Grade 11 |
| Mathematics | 3 | <ul style="list-style-type: none"> Math A/Integrated Algebra I | <ul style="list-style-type: none"> Grade 8/9 (if in an accelerated/honors course or block scheduling) Grade 10 Widely taken in January, sometimes in June |
| Science | 3 | <ul style="list-style-type: none"> Living Environment or Earth Science are usually completed to meet the requirement | <ul style="list-style-type: none"> Grade 8/9 or upon completion of coursework |

As Table 5 indicates, to pass Regents exams required for the diploma, students must score at least a scaled score of 65 (out of 100) on exams in Integrated Algebra (also known as Math A), Comprehensive English, U.S. History & Government, Global History & Government, as well as one exam of choice in science. (Math A covers concepts from algebra, geometry, and statistics and probability, while Math B focuses on a whole range of topics including trigonometry and advanced algebra. Math B is taken over the course of three semesters and it is often considered one of the most difficult of the Regents exams.) Students who want to earn an Advanced Regents Diploma also must pass the Math B exam, a second science exam and a world language exam.¹³

Only the mathematics and English Regents exams are used to meet the school accountability requirements of the federal No Child

Left Behind Act (NCLB). The cut scores for NCLB and graduation are the same—a minimum score of 65.¹⁴

What’s most notable about the Regents exams is their long-standing connection to higher education entrance and placement. Students who want to attend a State University of New York (SUNY) college or the City University of New York (CUNY) submit Regents exam scores as part of the admissions and course-placement processes. The actual score required is both higher than the 65 needed for high school graduation and varies among SUNY and CUNY and among different subjects:

- ⇒ At CUNY, students who score at least 75 on the Math A exam are considered ready for college mathematics (or courses with a mathematics prerequisite); students who score at least 75 on the Comprehensive English exam is considered ready for all other non-remedial classes. Students can still substitute other test results, such as SAT or ACT scores, to prove college readiness.¹⁵

- ⇒ For SUNY schools, a score of at least 85 on Math B exempts students from the general education mathematics requirement;¹⁶ a score of at least 85 on U.S. History exempts students from the required entry-level survey course (although they must still take one additional history course); and a score of at least 85 on a foreign language exam exempts students from the general education foreign language requirement, at the discretion of the campus.¹⁷

Table 6: Same Regents Exams, Different Scores for Different Uses

| Student Goal | Regents Exams Required | Minimum Passing Scale Score (1–100) |
|---|--|-------------------------------------|
| Regents Diploma | Five tests <ul style="list-style-type: none"> • Math A • One science • English language arts (ELA) • Two social studies | 65 on all tests |
| Regents Diploma w/Advanced Designation | Same five as Regents Diploma plus: <ul style="list-style-type: none"> • Math B • One more science (one of the two must be Living Environment) • Language other than English | 65 on all tests |
| Placement into College-Level Classes at CUNY | <ul style="list-style-type: none"> • ELA • Mathematics | 75 on both tests |
| Placement into College-Level Classes at SUNY | <ul style="list-style-type: none"> • ELA • Mathematics | 85 on both tests |

** For students entering 9th grade in 2008 and beyond, local diplomas will no longer be available, so all students must earn 65 or above on all required Regents exams.*

Though it’s not ideal that the public two-year system and four-year system have different placement standards, the collaboration of the K–12 and higher education systems around assessment is one of the strongest examples in the nation of focusing on college and career readiness. The placement policy has helped to send a common signal to students, teachers and parents about the importance of meeting college-ready standards in high school.

Approach Three: Adapted College Admissions Tests

The ACT and SAT are used nationwide as part of the college admissions and placement processes. To date, six states—Colorado, Illinois, Kentucky, Maine, Michigan and Tennessee—have incorporated the ACT or SAT into their high school assessment systems, requiring all students—not just those deemed college-bound students—to take them.¹⁸

What is measured on the ACT and SAT?

Begun in 1959 as a project of the American College Testing Program, the ACT assessment is a curriculum-based achievement test that measures student knowledge and skills in four areas: English, reading, mathematics and science reasoning. An optional writing test measures skill in writing a short essay. The ACT consists of four multiple-choice tests—one in each area—with each scored on a scale of one to 36. The composite score reported by ACT is the average of the four test scores. The ACT is explicitly curriculum-based, designed to assess what students have learned in their high school courses.¹⁹

Many colleges use ACT scores as surrogates for or supplements to course-placement procedures—in addition to their primary use as tools for evaluating students as part of the college admissions process.

Sponsored for more than 70 years by the College Board, the SAT Reasoning Test is intended to measure critical-thinking skills required for academic success in college. Originally called the Scholastic Aptitude Test, the SAT currently consists of three sections: critical reading, which has sentence completion and passage-based questions; writing, which has multiple-choice questions and a written essay; and mathematics, which is based on “topics from up through a third-year college preparatory course.”²⁰ Each section of the revised three-hour-and-45-minute SAT is scored on a scale of 200–800, with two writing subscores—one each for the multiple-choice and essay parts of the writing sections. (Each section is scaled so that 500 represents the average score with a standard deviation of 100 points.) Unlike most admissions and placement tests, the SAT mathematics test includes some items with student-constructed numerical answers. These are encoded in a special grid for computer grading.

The College Board suggests that the SAT be used primarily for its intended purpose of admissions but notes that SAT scores “may be helpful for identifying students in need of remedial training.”²¹ Many colleges use SAT scores as criteria for course placement or to exempt students from other tests that are specifically designed for placement purposes.

Advantages of Adapted Admissions Tests

College admissions tests are accepted widely by colleges, high schools, families and the general public as markers of whether students will succeed in a broad range of college courses. They are an important part of the college application process; scores are used to tell students if they are “college-eligible” and to which colleges they could apply, as well as how they compare to other students. In addition, though the ACT and College Board do not endorse this practice, many colleges use minimum ACT or SAT scores to decide whether to place students into credit-bearing courses or bypass students from certain developmental and even entry-level courses.

College admissions tests have another advantage over the other two approaches as well: The results are common across the country and therefore portable for students to use at postsecondary institutions. This is perhaps their greatest advantage. They also are comparable among states and have a long track record of use.

Many advocates believe that giving all students the opportunity to take these tests—because they are so tightly ingrained in the existing college admissions system—will give them a leg up on the college preparation process and may encourage students who did not view themselves as “college material” to pursue that path.

Challenges with Admissions Tests

The clearest challenge with college admissions tests is that they are not designed to be given as high school tests that have multiple purposes, such as determining if all high school students are ready for credit-bearing coursework or holding high schools accountable for meeting college and career readiness standards. This is largely because admissions tests are not designed to measure the curriculum of any particular school or state.

Achieve's 2007 *Aligned Expectations?: A Closer Look at College Admissions and Placement Tests*, a study of the major college admissions tests, found that the tests should not be included "as is" in state systems. Achieve analyzed more than 2,000 questions from college admissions tests (and also placement exams) to determine how these tests compare to one another and how well they measure the college and career readiness ADP benchmarks.

Achieve's analysis reveals that states will need to modify the tests to measure state standards. Generally, admissions tests are rigorous and balanced; in mathematics in particular, this is true with respect to both the content included and the rigor of the test items.²² Yet while admissions tests do some things very well, there are gaps in what they measure. Neither the ACT nor the SAT includes the full range of advanced concepts and skills reflected in the ADP benchmarks.

Other challenges include the amount of time it takes to get results back to students and schools; the limited number of new forms and administrations possible, especially when compared to state-developed testing programs that may make multiple forms each year; and the inability to release most or all of the test items from each administration.

It is clear that states cannot use the SAT or ACT "off the shelf" as the anchor assessment. States will need to adapt these assessments in order to test all students. Adaptations will range from augmenting with additional test items that measure state standards, as in **Illinois** and **Michigan**, to standardizing administration conditions and procedures across all high schools, as in **Colorado**. In **Maine**, state education officials conducted independent alignment studies of the SAT with Maine standards and concluded that they needed to add items in mathematics in order to meet state standards and comply with NCLB testing requirements.

Incorporating admissions tests into a state's assessment system as the college- and career-ready anchor assessment may not significantly increase the overall amount of testing experienced by students, as many students were already planning to take the tests. However, if states keep their current tests and add the ACT or SAT, this will increase the amount of testing that occurs during the school year and that takes away from instructional time. States could work creatively with districts and ACT or the College Board to adapt the standard Saturday test administration, but so far this has only occurred in **Maine**, which administers its modified SAT on a single Saturday.

States that are considering the ACT or SAT as the anchor assessment should take stock of what these tests measure, how well aligned they are to state standards and what modifications to state testing administration policies and procedures will be needed. If states choose this approach, they should work with ACT and the College Board to adapt the assessments to ensure greater coherence and alignment with state standards and ease of administration.

States that take this route to college and career readiness testing also may find that they need to supplement testing earlier in high school. If the adapted admissions test is given in grade 11, and is the first test given to all students since grade 8, students in grades 9 and 10 will not receive much information about their progress toward meeting college and career readiness standards while there is still time to adjust coursetaking and catch up.

Table 7: A Snapshot of Statewide Admissions Test Administration, Development and Cost Issues in Select States

| | Colorado | Illinois | Maine | Michigan |
|---|---|---|--|--|
| Test and Grade | Colorado ACT – 11th | Prairie State Achievement Examination – 11th | Maine High School Assessment – 11th | Michigan Merit Examination – 11th |
| Year of First Administration | 2001 | 2001 | 2006 | 2007 |
| Subjects Tested | <ul style="list-style-type: none"> • ACT Reading • ACT English • ACT Math • ACT Science | <ul style="list-style-type: none"> • ACT Reading • WorkKeys Reading for Information • ACT English • ACT Writing • ACT Math • WorkKeys Applied Math • ACT Science and state-developed items | <ul style="list-style-type: none"> • SAT Critical Reading • SAT Writing • SAT Math and augmented math items • State-developed science assessment | <ul style="list-style-type: none"> • ACT Reading • WorkKeys Reading for Information • ACT English • ACT Writing • ACT Math • WorkKeys Applied Math and Locating Information • State-developed math items • ACT Science and state-developed items • WorkKeys Locating Information and state-developed social studies items |
| Did the state modify or augment content? | No | Illinois developed a 45-question multiple-choice science test to augment the ACT science exam. | Maine augments the SAT Math with 18 items (mainly in data analysis) to align the test with state standards. Maine also administers a state-developed science assessment. | Michigan augments the ACT and WorkKeys with state-developed items in mathematics, science and social studies that more fully align the test to state standards. |
| Testing Time | 190 minutes on one day | 370 minutes over two days | <ul style="list-style-type: none"> • SAT: 240 minutes on regular Saturday administration • Mathematics and science augmentation: 165 minutes on one day | 460 minutes over three days |
| Are other statewide high school tests administered or in development? | 9th and 10th grade students are assessed with the end-of-grade Colorado Student Assessment Program. | No | Maine is developing performance-based assessments to measure 21st-century skills. | Michigan passed legislation requiring that the state develop end-of-course assessments in required high school courses/credit areas. |

Should States Use College Placement Tests as the College and Career Readiness Anchor Assessments?

Like the ACT and SAT, placement tests are a known quantity in higher education circles, though they are much less well-known or understood by students, high school faculty and the general public. Placement tests are usually used after students enter college to assess whether students possess the knowledge and skills in reading, writing and mathematics necessary to gain access to entry-level, credit-bearing courses. Students who fare poorly on placement tests are assumed to have gaps in their preparation and are advised—and often required—to take one or more remedial courses to learn (or relearn) important material considered prerequisite for credit-bearing courses. The most common placement tests given nationally are COMPASS and ACCUPLACER. Many institutions also develop their own placement tests.

Some states (including **Idaho, New Hampshire, North Carolina, Ohio and Washington**) are making college placement tests available for students to take voluntarily, sometimes at no cost, in high school. These exams can provide information to high school students about their readiness for credit-bearing, first-year college courses and allow teachers to work with students in their senior year to address learning gaps.

However, Achieve does not recommend substituting or using placement tests as the anchor assessments, which need to be comprehensive measures of college and career readiness, in the high school accountability system. Most of the college placement tests reviewed by Achieve are too narrowly focused on a subset of knowledge and skills. In mathematics and reading in particular, they reflect relatively low levels of rigor. If states were to incorporate existing placement tests into their formal high school accountability systems, this might inadvertently lead to a narrowing and watering down of the high school curriculum.

Should States Measure Career Readiness Distinctly from College Readiness?

Some state policymakers are interested in knowing whether and how to incorporate into the statewide assessment system additional tests—beyond the anchor assessments that measure academic preparation for success after high school—that are explicitly designed to measure readiness for specific industries or general “workplace readiness.”

There are literally thousands of certifications (and accompanying assessments) available in the United States. The certifications vary substantially in purpose, in what they cover, in the nature of their requirements and in how they are used. The category of industry certification, which includes any program designed to identify the skills and knowledge needed in a specific industry or job function, is by far the largest. A much smaller group of certifications, known as work or career readiness certifications, captures generic proficiencies and skills that are applicable to many occupations.

General Work Readiness

Credentials of general work or career readiness are designed to show employers that individuals have gained “soft skills”—those qualities and habits of mind such as work ethic, organizational skills, punctuality and self-discipline—as well as general or technical skills (such as basic information retrieval skills). Work readiness programs vary, but all are intended to certify that students are “career-ready,” without respect to particular occupations.

One of the best-known programs is ACT's WorkKeys program, which was developed in the early 1990s. There are nine Foundational Skills tests that measure cognitive skills and applied job skills in communication, problem-solving and interpersonal skills. Assessments that measure communications skills are administered in Reading for Information, Business Writing, and Listening and Writing; problem-solving assessments include Applied Mathematics, Locating Information, Applied Technology and Observation. In addition, an interpersonal skills assessment is a Teamwork test. Based on ACT-produced job profiles that capture the important skills needed for a particular job, each of the Foundational Skills assessments corresponds to four or five achievement levels so users can quickly see whether a test-taker has demonstrated knowledge and skills commensurate with the requirements for a particular industry or profession.

Unfortunately for state decisionmakers weighing how or if to include career readiness tests in state systems, there is little available research showing that employers rely on general work readiness certifications. And there is wide variation among work readiness credentials in terms of quality, rigor and utility in the labor market. It is impossible to make generalizations about the most effective approach for states to take. If states want to certify whether all students have soft skills before they graduate, they should consider which work readiness programs are gaining traction in their communities or regions and which would fit best within the overall system of state standards and assessments.²³

Industry Certifications: For All Students?

With a few notable exceptions—**Indiana**, for example, where a state-recognized certification is one of several ways students can meet the requirements for the Core 40 with Technical Honors diploma²⁴—industry certifications for high school students have been limited to students in career and technical education (CTE) programs or programs for youth who have dropped out.

Job-specific industry certifications typically are issued by an occupational or industry group to indicate the completion of specific training, coursework, apprenticeships or other preparation for a particular job. The development and governance of industry certifications are as diverse and complex as the American economy. Some are developed by industry associations, while others are developed by individual companies who offer proprietary training for their products.

States may want to incorporate or incentivize students to earn industry certifications, but policymakers face a number of challenges. First, there are thousands of certifications and most are not well-connected to mainstream high school teaching and learning. Second, how states can extend industry certifications beyond students in CTE programs remains to be seen. And, third, it may be impractical for the state to encourage, require or even pay for all high school students to earn industry certifications. Students who want to enter technical fields after high school may need job-specific certifications, but it is not clear that all students need or will benefit from these credentials.

Some states may find career readiness or industry credentials useful to connect the K–12 and postsecondary education systems with the state's economic and job-growth goals, either because strong local industries support them or because a particular assessment is well-aligned with individual students' goals. However, states should not substitute industry certifications for assessments of academic preparation for college and the workforce. Such a policy could leave students without the broad foundational skills in reading, writing and mathematics that they need in a wide variety of careers—in exchange for narrower skills relevant only to a handful of jobs.²⁵

Core Principle Two: Making Tests Matter Beyond K–12: Opening Doors to Higher Education and Employment Using High School Tests

It won't matter how good the state assessments are if education leaders and employers do not use the results. The tests should, in fact, open doors to higher education and employment. This can only be accomplished through close collaboration among state K–12 and postsecondary leaders.

For college- and career-ready assessments to have credibility, especially in the postsecondary community and also with educators, parents, students and the public, both the tests' content and cut scores have to be grounded firmly in what it takes to be successful in college and careers. They cannot be subject to the downward pressure on standards that we have seen with graduation exams.

If K–12 and public higher education use the same tests for college readiness, students will appreciate knowing that the assessments tell them whether they are ready for college-level work; educators will benefit by working together to strengthen and streamline curriculum and instruction; and parents will see the value, quite literally, in ensuring their children will not need costly remediation when they get to college.

Chapter Highlights

- ⇒ Higher education needs to collaborate with K–12 to develop/modify and truly use the college- and career-ready anchor assessment in college course placement decisions.
- ⇒ The tests' content and cut scores have to be grounded firmly in the skills to succeed in college and careers.
- ⇒ Higher education also needs to establish a common minimum placement standard for credit-bearing, freshman coursework across the majority of public two- and four-year institutions.
- ⇒ Meeting standards on the anchor assessment should open doors to credit-bearing, freshman-level coursework, as well as to postsecondary or employer-based apprenticeships and training.
- ⇒ Students should not be required to pass the anchor assessments to graduate, at least not in the near future. States will need to decide whether to keep their exit exams in addition to the college- and career-ready tests.

Collaboration To Develop and Use High School Tests

For these goals to be met, postsecondary leaders should feel confident that the high school anchor assessments measure the knowledge and skills students will need to be successful in college-level credit-bearing courses. Getting to this level of confidence requires higher education and K–12 leaders to work closely together to define the content standards and set cut scores for performance.

The stage for collaboration on assessment is being set by the initial collaboration on standards in many states. K–12 education agencies that have defined college and career readiness standards for the end of high school successfully have done so with the full participation of higher education and employers.

That partnership should be extended as new assessments are developed or existing state tests are modified. Postsecondary faculty, as well as campus leaders from two- and four-year institutions and system offices (and, ideally, employer representatives) likely will need to:

- ⇒ review test questions and, if important college- and career-ready skills are missing, identify how the test can be augmented to better measure those content and skills;
- ⇒ validate that the tests, overall, measure the content and skills most important for success in freshman college courses;
- ⇒ participate in determining the test's cut scores; and
- ⇒ engage in predictive validity studies to evaluate whether the high school assessments provide accurate information about student readiness for credit-bearing courses (though state leaders don't need to wait years to begin using the tests).

This is not the way most states are accustomed to managing their assessments. Higher-education leaders or faculty—let alone employers—are rarely invited into the K–12 assessment development process. This needs to change.

One of the best examples of K–12 and higher education collaboration is the partnership between the **California** Department of Education (CDE) and the California State University System (CSU).²⁶ Working together, CDE and CSU modified the state's existing 11th grade assessment to include college-ready skills and identify college-ready scores in English and mathematics. Students who score at the college-ready level and continue to take challenging courses in their senior year do not need to take the placement exams at CSU or get a certain score on the SAT/ACT. Instead, they are automatically placed in credit-bearing entry-level courses. Just as important, students who do not score at the readiness level can receive assistance during their senior year to prepare them for college. (For more on the EAP, see the detailed sidebar on page 30.)

Another notable example—that also includes serious efforts to engage employers—is underway in **Hawaii**. Hawaii K–12, higher education and business leaders have been working closely together via the statewide P–20 Council to advance student preparation by helping students “opt up” into a college- and career-ready course of study. Students who complete four years of mathematics through Algebra 2, four years of English including a semester of expository writing, and three years of lab science and who meet the state standard on the ADP multi-state Algebra 2 EOC exam will be rewarded with the following incentives and opportunities:

- ⇒ For those going on to higher education, students who meet the new diploma criteria will be able to meet the “rigorous curriculum” requirements of the state's new need-based college scholarship, called the B+ Scholarship, and gain automatic admissions at two of the state's four public baccalaureate institutions. Research is also underway to determine how to use the new ADP Algebra 2 EOC assessment to waive placement testing at Hawaii two- and four-year public colleges.²⁷

- ⇒ For those going directly to the workforce, students can accelerate through apprenticeship programs for the trades. Working together, the Hawaii Department of Education, Board of Regents, University of Hawaii System, Business Roundtable and Hawaii P-20 Council have secured commitments for students who meet these criteria to enroll directly in and proceed faster through the apprenticeship program led by the state carpenters' union or for electricians through the public utility—without taking the apprenticeship placement test. Employers also are working to identify alternate incentives, such as waiving entrance tests in some companies and apprenticeships tests in other fields.

Texas provides another model for collaboration. Texas higher education policymakers reviewed the 11th grade Texas Assessment of Knowledge and Skills (TAKS) to ensure it assessed the skills needed for college coursework. They then established a college-ready cut score higher than the score needed to graduate. The state even includes the percentage of students achieving that cut score in its annual school report cards.

The American Diploma Project Multi-State Algebra 2 Exam: States Raising High School Mathematics Achievement—Together

In May 2005, states in the American Diploma Project (ADP) Network began to explore the possibility of working together, with support from Achieve, to develop a common end-of-course exam in Algebra 2. These states were planning to require or strongly encourage students to take Algebra 2 (or its equivalent) to better prepare them for college and careers. State leaders recognized that using an end-of-course test would help ensure a consistent level of content and rigor in classes within and across their respective states. They also understood the value of working collaboratively on a common test: the potential to create a higher-quality test at a lower cost to each state and to compare their performance and progress with one another. With increasingly common end-of-high-school expectations among the states, collaborative efforts to develop assessments make both good education policy and economic sense.

Fourteen states—**Arizona, Arkansas, Hawaii, Indiana, Kentucky, Maryland, Massachusetts, Minnesota, New Jersey, North Carolina, Ohio, Pennsylvania, Rhode Island and Washington**—joined together to develop and use the common end-of-course exam in Algebra 2. This is the largest multi-state collaborative assessment effort ever undertaken.

In the spring of 2008, nearly 90,000 students across 12 of the 14 states in the partnership took the exam for the first time. In subsequent years, the number of test-takers is expected to grow.

The test development process may provide a model for collaboration across K-12 and higher education on assessment. The states involved high school mathematics teachers and postsecondary faculty in the review of questions as the test was developed. Studies also will be conducted at two- and four-year college campuses to establish the relationship between the ADP Algebra 2 exam and success in first-year credit-bearing college mathematics courses. Participating states also will examine the relationship between the ADP Algebra 2 exam and commonly used admissions and placement tests. All of these data will inform setting performance levels and cut scores and enable states to be more confident that proficient truly means prepared.

Collaboration Within Higher Education To Set Common Readiness Standards

Different higher education institutions and campuses have different missions and some institutions can be more selective in whom they admit, particularly the public “flagship” universities that are often the most selective. But having open enrollment policies or low entrance standards doesn’t mean that institutional standards are low.

Despite colleges’ varying missions, what college faculty want students to be able to know and do is fairly consistent across college campuses. That’s why it makes sense for higher education faculty and administrators to work together (with participation from K–12 educators) to establish a single statewide standard for college readiness at the open enrollment, nonselective and moderately selective two- and four-year public institutions. (“Flagship” or research universities may want a higher standard.) This standard should signal whether students are ready for credit-bearing coursework in core subjects like mathematics and writing.

In order for this to happen efficiently, leaders and faculty need to collaborate *within* higher education, *across* college campuses, and even *across* the two- and four-year systems.

As the K–12 and higher education systems work together to set a common college-ready cut score on the anchor assessments, higher education institutions ought to commit to using that score as a placement mechanism for students into credit-bearing courses. Reaching this level of commitment statewide will require the various higher education systems within a state to come to an understanding for both the transfer and articulation of credit.

Ohio provides a recent example of collaboration within higher education to set a common minimum placement standard. The University System of Ohio (USO) recently adopted a ten-year strategic plan that includes a clear definition of college readiness. As an interim strategy to enable higher education to immediately begin working with parents, students, teachers and counselors to help students get ready for college, the state’s Articulation and Transfer Council made, at the direction of the legislature, a recommendation regarding a statewide placement standard for entrance into the first college-level courses.

The USO chancellor accepted this recommendation, and the Board of Regents, in partnership with the Ohio Department of Education, will begin designing and implementing strategies that help students prepare for and meet this standard. The goal of Ohio’s statewide placement policy is to help students, parents and teachers determine whether a student is college-ready and to let schools know whether a student is ready for the first non-remedial course in English and mathematics. The policy is:

- ⇒ ACT score of 18 or higher for English (or an equivalent SAT or COMPASS of 69 or higher); and
- ⇒ ACT score of 22 or higher for algebra (or an equivalent SAT or COMPASS of 65 or higher for placement into non-remedial algebra)

Students interested in majors such as mathematics, physics and engineering will need to achieve a mathematics ACT score of 27 or higher (or COMPASS score) for placement and success in college-level calculus. While not required, an assessment of writing is strongly encouraged for placing students in courses of English composition.

This policy is not an admission standard for college. All high school graduates will be admitted to the community college of their choice. Some universities will likely have admission standards below this level as well. However, those students who do not meet the standard should expect to be enrolled in remedial classes.²⁸

By adopting this college placement standard, Ohio is giving all students a leg up on knowing what is required to avoid the need for remedial education, thus saving money and time toward a college degree.

Making College and Career Readiness Tests Count for Students

Anyone who has taught (or parented) high school students knows that if you give them a test or assignment and tell them it doesn't count, the test will get blown off or be given only halfhearted attention. It is in light of this reality that states need to decide what type of "stakes" to attach to new high school assessments.

Yet, given the current state of student readiness and school capacity, virtually no state is in a position to insist that all students hit the college- and career-ready level of performance on these tests to earn a high school diploma in the next few years. For these assessments to have credibility, especially in the postsecondary community, the test content and the cut scores have to be firmly anchored in what it takes to be successful in college and careers. States that have graduation exams know how challenging it is to set the standards on those exams at a high level while ensuring that all students receive the support they need to pass them. College- and career-ready anchor tests cannot be subject to the same downward pressure on standards that we have seen with graduation exams.

Rather than requiring students to pass college- and career-ready tests for graduation, states should consider attaching more positive incentives. If the collaboration on the tests is done correctly, students who score at the college-ready level can be guaranteed enrollment in credit-bearing, non-remedial courses in college. This will enable postsecondary systems and institutions to waive placement exam requirements, thus creating greater coherence between the two systems.

Another way to provide incentives for students on these college- and career-ready anchor assessments without making them high stakes is to allow the tests to count for a portion of the course grade. This is only feasible in states that choose to implement EOC assessments. This also could mean providing additional scholarship dollars as part of state need-based financial-aid programs for students who perform well.

In states that already require students to meet state standards to graduate, policymakers need to decide how the exit exam will match up with the college- and career-ready anchor assessment. Those states with a comprehensive EOG assessment in grade 11 may also want to consider requiring students to perform at a lower, but still meaningful, level on the anchor assessment in order to graduate.

However, in most cases, the exit exams are not challenging enough to be modified into college and career readiness tests, so states will need to determine whether both are needed. If all students are being given an assessment of college and career readiness, and this assessment opens doors to higher education, does the state also want to guarantee that students have met the minimum skills tested on the exit exam? In some states, the answer will be yes, and state education leaders will need to figure out how the exit exams relate to the new anchor assessments and how results will be factored into school accountability determinations. Other states may decide to phase the exit exam out over time as anchor assessments of college and career readiness take root.

California's Early Assessment Program: Assessing and Assisting Student Readiness

California's Early Assessment Program (EAP) is widely recognized as one of the most innovative programs for assessing high school students' readiness for postsecondary education. More than 60 percent of the nearly 40,000 first-time freshmen admitted to the four-year campuses of the California State University (CSU) require remedial education in English, mathematics or both—yet these 25,000 freshmen had taken the required college preparatory curriculum and earned at least a B grade point average in high school.²⁹

The California Department of Education (CDE) and CSU co-developed EAP in order to address these extremely high rates of remediation. The EAP builds on the state's existing 11th grade high school test in English and mathematics, the California Standards Test (CST), as well as the Algebra 2 end-of-course state test. After a collaborative process in which CSU faculty examined the CST test content and discovered that a good deal of the content needed to be college-ready was already included, CSU leaders had test developers create a special "Part B" with 15 additional multiple-choice questions in mathematics and 15 additional questions and a 45-minute essay in English.

Participating in Part B is voluntary and is geared toward students who are taking college-preparatory courses, yet test-taking is increasing rapidly. In 2008, nearly 353,000 high school juniors volunteered to take the EAP, which is 79 percent of eligible students. In August 2007, EAP test results were included for the first time on the K–12 score reports sent to all parents. EAP results indicate that students, by and large, aren't college-ready. In mathematics, only 13 percent were ready for college-level credit-bearing courses; an additional 42 percent were "conditionally-ready," while 44 percent were not ready. (Conditionally-ready students must take additional mathematics coursework in their senior year.) Only 17 percent were college-ready in English.³⁰

Because one of the goals of EAP is to help students identify gaps in their skills if they are not ready for college-level classes in mathematics and English, CSU and K–12 leaders haven't stopped with test development.

The EAP enables high school students to adjust their senior-year coursework if they need additional preparation for college. In order to make the senior year count, CSU designed the CSU Success Web site (see www.csusuccess.org) explicitly for students (though it also is used by teachers) to set out a step-by-step personalized road map to postsecondary education. The site includes preparation tools, testimonial student videos, access to each student's EAP scores and e-mail reminders about key college-going milestones and deadlines. K–12 and higher education leaders also worked to modify and improve 12th-grade high school courses by providing online English language arts and mathematics modules.

In English, California students are required to take three years of coursework in order to graduate from high school; students who want to attend CSU or the University of California (UC) need four years. CSU and K–12 faculty put a meaningful twist on the traditional senior year language arts class. Rather than the standard literature-based English class option for 12th graders, faculty redesigned the course and transformed it into an Expository Reading and Writing course. This new course helps high school seniors strengthen their critical reading, analytic and expository writing skills, which will better prepare each student to transition into college-level courses across the curriculum in college. High school teachers report they find the course material academically rigorous as well as engaging for their students and that the materials enhance their instruction in meaningful ways. Even better, students who have taken the newly designed course have scored higher on CSU placement exams than those who have not. As a result, the Expository Reading and Writing course was approved in 2006 by CSU and UC as satisfying their entrance requirements.

In mathematics, California requires students to complete two years of coursework before high school graduation; students who want to attend CSU or UC need three years (though four years are recommended). The EAP includes a new professional development series for high school mathematics teachers, which assists teachers in preparing students before the 11th grade CST and assists students through online preparation exercises (see www.collegemathsuccess.org). In part because of teacher shortages, the mathematics module has been slower to launch than the new Expository Reading and Writing course.

Teachers also receive substantial support. CSU reports it has held numerous EAP awareness sessions across the state for teachers of all backgrounds and experience levels and hosted online workshop and in-person full-day sessions at various locations and free of charge to high school mathematics teachers. This is noteworthy because the CSU system prepares between 55 and 60 percent of California's new teachers.³¹ Information about the content standards most needed for college readiness, what college proficiency levels look like and general EAP information is also now embedded in CSU's pre-service and in-service programs for teachers and administrators.

Core Principle Three:

Not Just More Assessment—Better Assessment: Measuring the Full Range of College and Career Readiness Skills and Supporting Richer Instruction

In the refined high school assessment system Achieve proposes, assessment will become not just a one-time event experienced mostly at the end of the school year. Assessment will become an ongoing tool for continuous improvement.

To realize this vision, state assessment systems will need to strike a better balance between tests that are mostly useful for holding students and adults accountable and tests that are mostly useful for improving teaching and learning.

Too many teachers don't see the connection between what statewide standardized tests measure or how their day-to-day curriculum will prepare students to meet standards on the state tests. Some advocates call for teacher-generated assessments or "multiple measures" in place of statewide summative assessments, while other advocates argue that the state tests ensure a common foundation of learning for all students and help to promote equity.

States don't have to choose either large-scale summative tests or classroom tests; they can build high school assessment systems on a foundation of higher-quality standardized tests coupled with assessments that teachers will find particularly valuable. Combining multiple kinds of assessments strategically in state systems will generate better information about students' college and career readiness and shore up support among educators. States can play a leadership role in making these sorts of assessments available—and ensuring a consistently high level of quality—to schools and districts.

Improving Large-Scale Assessments

Whether new tests are created or existing tests are modified, what is most important is that the statewide assessments measure essential skills and knowledge from the college and career readiness standards—and that the assessments do so in a way that guides instruction in the right direction.

Most statewide, summative, large-scale assessments are administered in increments of 50–120 minutes with 40–80 questions. All rely heavily—many exclusively—on machine-scoreable responses, which are primarily "bubble-in" multiple-choice or gridded formats. Some include constructed-response formats that ask students to write out their answers. Even most of these are short—generally only one or two sentences. Some states have items that may involve longer written responses, such as up to a page. Direct assessments of writing typically involve responses of one, two or three pages per prompt.³²

Chapter Highlights

- ⇒ New anchor assessments and other statewide summative tests should be robust and well-designed: They should include constructed-response items and have quality and quick scoring so that they are useful to educators and students.
- ⇒ Not all important college and career readiness skills can be tested on the anchor assessments—and so additional measures will be needed.
- ⇒ States also need to give weight to and help stimulate locally led performance measures. Performance assessments can include graduation projects, writing portfolios, science experiments or many different demonstrations of student learning.
- ⇒ States also should identify or build high-quality, aligned interim assessments that diagnose student learning strengths and gaps.

Valuing Constructed-Response Items on Large-Scale Assessments

Good multiple-choice tests can measure rigorous content and analytic skills, such as reasoning and problem solving; they needn't be limited to testing a student's ability to recall discrete bits of information. States should create more cognitively challenging multiple-choice items that tap higher-level content and a variety of skills. However, states ought not to rely exclusively on multiple-choice items because they cannot always probe students' capacity to analyze, evaluate, demonstrate or explain their thinking.

Currently, 29 states include short-answer items and only 24 states use constructed-response or extended-response items in their high school tests. Every state but one gives a direct writing assessment in high school.³³

States need to make far better use of open-ended questions, which are often more effective at measuring more applied skills, like problem solving in mathematics and writing. This means simultaneously increasing the quantity of these items while also making sure they are of higher quality than many are currently.

Using these items costs more than using only multiple-choice questions and they take more time to administer. Yet, the payoff downstream for students who become critical thinkers and problem solvers, and the payoff for schools in the increased quality of curriculum and teaching that such test formats encourage, makes such investments worthwhile.

Massachusetts is a good example of a state with high-quality constructed-response items. Its high school assessment system currently consists of ELA and mathematics Massachusetts Comprehensive Assessment System (MCAS) tests in 10th grade. Both assessments include a combination of multiple-choice and constructed-response items, giving students the opportunity to demonstrate a deeper and more sophisticated understanding of content in each subject. Massachusetts works closely with its test developer and with educators to create new assessments each year that are closely aligned with the state's curriculum framework and that measure the full depth and breadth of state standards. In designing each year's assessment, the state emphasizes a writing prompt and constructed-response questions in ELA and short-answer and extended constructed-response questions in mathematics. For example, out of 52 questions on the 2007 mathematics MCAS, five were short-answer and eight were extended constructed-response, amounting to 25 percent of the assessment and over 40 percent of the scale points.³⁴ Massachusetts annually releases all of the MCAS constructed-response test items, along with scored samples of actual student responses.³⁵ The questions and writing prompts, scoring guides and corresponding student work help Massachusetts uphold its commitment to transparency in its assessment policies, provide valuable information to educators about the expectations for student performance on the MCAS tests, and allow educators to evaluate the type of content knowledge and skills with which students are struggling or succeeding.

States also are experimenting with innovations in extended-response items in other subjects. **Connecticut** incorporated a staged science lab experiment into the annual summative test. Students were given an experiment to conduct, followed by multiple-choice questions about how they designed and conducted the experiment, as well as open-ended items that explored the scientific method and what students learned by conducting the experiment. **Washington's** high school science assessment incorporates similar methods and includes extended-response items also designed to explore student understanding of the scientific method.

Improving Scoring and Reporting

In addition, as states modify existing or build new college- and career-ready assessments, they also should take steps to improve how and when results are reported back to educators and students.

There are many ways that results and reports from large-scale tests can be improved.

- ⇒ States should strive to return results before the end of the school year, especially if the results are returned early enough to influence what courses students take next.
- ⇒ At least at the classroom or school level, test results should be available for each content “strand” in the standards, such as “statistics and probability.” This is difficult to do with large-scale assessments, which often don’t have enough items to report strand results at the student level, but it may be facilitated with EOC tests. Tests of Algebra 2, for example, ought to have enough items to report aggregated results on “operations on numbers and expressions” and “equations and inequalities.”
- ⇒ Results are typically reported at the student, school, district/county and state levels. States’ web-based reporting should become a tool to enable educators and district administrators to easily look at results at the classroom, course-title or department levels. This analysis can be done in many states today, but the data are not easily accessible electronically in a format that enables educators to aggregate the data as they need.
- ⇒ Teachers should have a larger role in scoring statewide assessments. This approach is not only excellent professional development but it also helps to demystify what is expected of all students.
- ⇒ All or nearly all open-ended test items should be published with corresponding examples of student work at each performance level—and better descriptions of why the work did or did not meet standard.
- ⇒ Score reports—at least those for middle school and early in high school—should indicate whether students are on track to meeting college and career readiness standards by the end of high school. This may require states to vertically moderate their elementary, middle and high school standards and tests.

Incorporating Performance Assessments

Some of the essential skills that college faculty and employers value in high school graduates are very difficult to measure via on-demand tests—even good ones. For example, it is difficult via on-demand testing situations for students to demonstrate that they can engage in teamwork or perform contextualized tasks that involve extended analysis, research or communication.

Accordingly, states should supplement their summative high school tests with performance assessments. Examples of performance measures that can capture a broader range of student skills include laboratory experiments, research papers, team projects, essays, portfolios, demonstrations, presentations and exhibitions. Using performance-based measures will increase the range of skills that are the focus of instruction and decrease the temptation to teach only the subset of skills and knowledge that is included in the summative tests. States that have gone down this path also describe it as powerful professional development for teachers, helping them to understand more deeply what kind of work is “good enough.”

That said, getting performance assessments right at the state level is challenging. States have struggled with the cost, the workload burdens on teachers and the inconsistency of expectations in scoring. As a result there are few examples of sustained large-scale efforts to implement performance assessments statewide. **Connecticut, Kentucky and Oregon** are among the most notable examples.

States can follow these rules of thumb as they consider when and why to use performance measures:

- ⇒ The content or skills to be assessed require complex construction by students or the integration of content/skills with other “habits of mind.”
- ⇒ It’s essential to get evidence of students’ reasoning or thinking processes.
- ⇒ There is an important context or setting that isn’t easily handled in large-scale testing situations, like a science lab.
- ⇒ Assessing the content/skill takes more time than available in a typical large-scale assessment.
- ⇒ The state wants to assess performances other than writing or choosing the best answer from among a set number of options.
- ⇒ The state wants to support richer instruction and deeper engagement by students.³⁶

As states set out to design performance assessment tasks, they should weigh all other operational considerations (such as administration, scoring, scaling, equating and reporting). States also should consider the costs and benefits of practical issues such as available time, money, expertise, likely education benefits, and the need for the measures to have credibility with policymakers, K–12 educators, higher education and the public.

As mentioned above, one of the biggest challenges in incorporating performance assessment into a statewide system is addressing concerns about consistency and quality. To implement performance measures—especially if the results are going to be used for accountability—states will need to establish clear performance criteria and scoring rubrics, disseminate examples of student work, assist with professional development and establish auditing mechanisms to help with reliability and enhance the measures’ credibility among educators and policymakers.

Local scoring—the process of examining and scoring student work—can become an embedded mechanism for ongoing professional development. Participating in the process will increase educators’ familiarity with student work as well as educators’ reliability in deciphering between strong, mediocre and weak work. The benefits of this experience can, and should, spill over into the classroom.

How States Are Using Performance Measures Today

One emerging strategy that embeds performance assessments in high school testing systems is to require high school students to complete graduation projects. The idea is to require each student to demonstrate a significant performance in order to graduate but rely on schools (sometimes along with community advisors and experts) to design, administer and score the performances.

The tradition of these exhibitions, research projects or presentations is that the exhibition focuses on a high level of performance and competence. The projects often are judged externally with criteria that include whether students demonstrate higher-level skills and knowledge and show the initiative, responsibility, maturity and other “habits of mind” often valued in higher education and the workplace. These graduation projects tend to be used to demonstrate depth and expertise in selected areas—sometimes integrating content and skills from across multiple disciplines—rather than showing general mastery of all the college and career readiness standards. The idea is to go deeper, not broader.

Idaho, North Carolina, Pennsylvania, Rhode Island and Washington require all students to complete an extended project as part of statewide high school graduation requirements, and **West Virginia** encourages all high schools to require rigorous senior projects for

graduation. **Hawaii, Louisiana and Texas** require graduation projects for students earning an honors or advanced diploma. In **Kentucky**, students earning a technical/vocational diploma must complete a culminating project related to their career cluster or major.³⁷ In all these cases, it is the local high school that administers the assessment, gauges student performance and decides if it is “good enough” to meet graduation requirements.

States also are stimulating, requiring and/or scoring performance assessments statewide. States are developing formal or informal partnerships with districts and nonprofits in order to encourage performance assessment linked with state standards.

Kentucky has the longest-running statewide performance assessment of any state. High schoolers must complete writing portfolios, which are included in the statewide school accountability index.³⁸ The writing portfolios are scored locally, while the state department of education plays a central role in auditing and providing extensive scoring training and professional development.

The **New Jersey Performance Assessment Alliance (NJPA)** is a five-year grant-funded project designed to develop, administer, and score valid and reliable performance prompts that will complement statewide standardized tests. The prompts measure mastery of state standards that are not currently assessed by summative tests through multiple-choice or constructed-response items. All prompts require students to apply their skills and knowledge in a novel way. Participants in NJPA have administered tasks in language arts literacy, mathematics and science at grades 3, 6, 8 and 11. More than 100 school districts, including hundreds of teachers and thousands of students representing a cross-section of New Jersey, have participated to date.

Of particular note is the use of a performance task along with the state’s new EOC test in Biology/Life Science. The EOC test is mostly multiple-choice, with some constructed-response items, and was operational for the first time in spring 2008. The New Jersey Department of Education (NJDOE) and NJPA field-tested a performance assessment prompt of approximately 90 minutes with all students who took the Biology EOC test in 2008. Student responses to this prompt are scored and reported by NJPA to the NJDOE. The state has made prompts, student response booklets and full sets of teacher guidelines indicating how the prompts align to standards available online.³⁹ At this time, though student scores on the prompt will not form part of the student score for the Biology/Life Science EOC test, by coupling the EOC test with the performance task, the state is signaling that both kinds of assessment are needed to measure student achievement.

Other states—most notably **Oregon, Washington and Wyoming**⁴⁰—include a process referred to as “collections of evidence” in their assessment systems. Since 2000, **Oregon** has employed a “collections of evidence” approach to supplement the state’s multiple-choice summative tests with performance tasks. Students submit to the state a series of work samples in reading, writing and mathematics that represent their growth to standards and quality of work over time. This collection supplements the state’s multiple-choice standardized test. The work samples can include research papers, experiments, speaking presentations and work experience. To support local school systems, the Oregon Department of Education (ODE) provides educators with sample tasks, prompts and examples of score student work; the mandatory official scoring guide also gives schools specific, consistent criteria for assessing students’ performance, along with grade-level anchors and scoring rubrics. To promote both transparency and consistency, the ODE also has created a practice scoring Web site to provide teachers, students and parents an opportunity to practice scoring student papers and receive feedback on their scoring accuracy.

Rhode Island has designed its new high school graduation requirements to rest upon a comprehensive system of locally designed performance assessments. Beginning with the graduating class of 2008, high school students must demonstrate proficiency in applied learning skills in ELA, mathematics, science, social studies, technology and the arts. It is up to schools to decide whether students are assessed through portfolios, exhibitions, EOC assessments or Certificates of Initial Mastery. Rhode Island has developed an online tool kit to provide schools with detailed assistance in the planning, development and implementation of a local assessment system that is aligned to the state’s grade-level standards and allows students to cultivate personal interest and deeper knowledge in the content areas. Although schools have latitude to develop performance assessments that best fit their communities, the state monitors their rigor, validity, reliability and alignment with standards through a “peer support and review” process.

Although the graduation policies in Oregon and Rhode Island are relatively new, each state sees performance assessments as a promising means to expand the knowledge and set of skills that high schools assess and to influence the quality of teaching and learning. Measuring college and career readiness through performance assessments will require significant guidance and monitoring from the state, but these two states' experiences demonstrate the possibilities of this approach.

Ohio: Piloting a Next-Generation Assessment System and Common Performance Measures

In 2007, Achieve and McKinsey & Company released a study that benchmarked Ohio's education policies against international best practices.⁴¹ One key recommendation is that, while Ohio's assessment system is in good shape, it should be strengthened to include more measures of college and career readiness, including performance measures and classroom-based tasks.

Ohio leaders also have determined that its current 10th grade high school assessments have several unintended consequences. These unintended consequences include:

- ⇒ Teachers feeling pressure to teach to the test, to focus on students who are “barely proficient” and near the cutoff for passing, instead of focusing on stretching students to their potential.
- ⇒ The tests focusing on basic, rather than college- and career-ready, skills and having low-level tasks instead of questions that emphasize higher-order thinking and reasoning skills

In response, the Ohio Department of Education has launched a pilot program to design Ohio's “next generation” assessment system. The pilot program has several key design principles:

- ⇒ Ohio's assessment system should involve teachers more as partners.
- ⇒ Assessment data should inform education decisions for all students and be driven by the curriculum, rather than the other way around.
- ⇒ The system should measure all high school standards, not just those that can be tested on paper-and-pencil assessment.
- ⇒ The system should signal college and career readiness for 11th and 12th graders.
- ⇒ Multiple measures are important to gain a complete picture of student learning.
- ⇒ The system should help assess and meet the legislative intent of Ohio's new college- and career-ready graduation requirements.
- ⇒ The system should assess “21st-century skills” as well as academic skills—including learning and innovation skills; information, media and technology skills; and life and career skills' flexibility and adaptability.

To meet these goals, the pilot program will engage a representative sample of volunteer school districts (who will be selected through an application process) to develop a curriculum-embedded performance assessment model collaboratively in English, mathematics and/or science for the 11th and 12th grades. In addition, participating districts will receive professional development and support to design course syllabi and performance tasks, use scoring rubrics and align the tasks with college-ready expectations. The assessment model will be consistent across the participating districts and will be piloted during the 2009–10 school year.⁴²

Interim Assessments: Tools and Support For Instructional Improvement

Even the best state systems that include multiple data points on high school performance often don't provide the kind of data that educators and students need—at least not when they need it most. Students, parents and classroom teachers need and want to understand what to do with the data and feedback generated from assessments. A lot of the concern about “too much testing” may really be about whether or not the tests are seen as useful to educators and students, especially if the tests are perceived as having high stakes but low instructional value.

As we have described, few state tests provide results soon enough or in a format that is useful enough for educators to adjust instruction during the school year. Statewide “summative” tests are by their very nature designed to measure learning at the end of the year. They are not the best tools for teachers to diagnose academic weaknesses and adjust instruction during the year. Yet educators are being held accountable for results and they need additional tools and training.

What many clamor for are classroom-based assessments whose primary purpose is to inform, support and improve instruction so that all students can meet state standards. This is why we suggest that states take a leadership role to ensure that interim assessments are available to all teachers.

What Is the Appropriate State Role?

We recommend that states and districts focus on developing or identifying high-quality interim assessments—those that are given at the end of marking periods to determine if students are on track to meet standards. True formative assessments—the kind that teachers give daily and/or every few days—should continue to be created and used only by teachers. States can encourage greater formative assessment literacy through teacher training and professional development but should focus development efforts on interim assessments.

It should not be left to chance whether schools and districts have access to these measures, but, unfortunately, that remains the case today, leaving many without these vital assessments. In the system we envision, states take responsibility for ensuring that schools and districts have access to these richer assessment tools while the responsibility for administering, scoring and using them remains at the district level.

There are a variety of interim assessment products on the market created by textbook and testing companies, but these are of varying quality. Just as we've seen with summative assessments, interim assessments will only be useful if they are high-quality—meaning they are well-aligned with state standards and able to provide useful information back to districts, schools and teachers.

The state role here should not be to mandate a single statewide interim assessment; states should take responsibility to ensure that all schools and districts have access to at least one full set of high-quality interim assessments. Districts should be free to use their own if they are of high quality. States may want to take a firmer approach with low-performing schools and districts by requiring those systems to use the state-provided tools unless they can prove that their own tools are equal or superior.

State departments of education also can serve as a quality-control clearinghouse by, for example, evaluating districts' use of interim assessments; vetting potential vendors of interim assessments; and providing information to districts on the characteristics of the interim assessments that are commercially available, focusing on the quality of the items and alignment with state standards.

Testing Innovations that Support Learning

States don't have to only consider 20th-century technologies as they develop or adapt new tests to measure 21st-century skills and knowledge. States can use traditional paper-and-pencil tests to measure learning, but they also can adapt tests into online formats, reducing time for scoring and providing students with faster and more frequent opportunities to take the assessment. Online assessments

will produce and report test results quickly—in some cases immediately or within 24 hours—potentially making it easier to integrate assessment and instruction throughout the school year.

In **Oregon**, at the discretion of school districts, students can take the multiple-choice portion of the 10th grade assessment reading, mathematics, science and social studies tests online. They can attempt reading, mathematics and science up to three times and social studies twice between October and May of 10th grade.

Supporting Teaching and Learning in Michigan

In 2005, when Michigan policymakers enacted the Michigan Merit Core, which is a series of minimum high school courses required for all students to graduate, they also decided to adopt the ACT and WorkKeys as the statewide high school assessment and to stimulate greater consistency and quality of teaching across the state by creating and disseminating curriculum-embedded tests.

The state department of education has been funded to create “model” end-of-course exams. One or two examples of each exam in a subject area will be built, placed on a secure Web site and made available to local districts. There are no testing services provided by the state and the original legislation leaves it up to the local districts as to whether they use the model state exams.

A new partnership in Michigan is forming to promote higher-quality and more consistent use of interim assessments. Members of the Michigan Assessment Consortium (MAC) include the state department of education, intermediate service districts, local school districts, and nonprofit and university partners. The members plan to collaboratively develop, implement and use a variety of student assessments to improve student achievement and provide professional development for local educators on how to use appropriately a variety of assessments.

Backers of the consortium believe it will help streamline testing and improve quality. As the stakes for student achievement have increased due to federal and state laws, some districts have responded by independently creating and using interim benchmark and end-of-course assessments. Yet the technical quality of these assessments may be questionable and much of this work is redundant. Other districts may not have changed their assessment practices and instead are using tests that are not aligned to Michigan’s content standards. MAC provides an opportunity for districts to work together to develop quality assessments aligned to Michigan standards more efficiently and at a lower cost.

MAC hopes to develop interim assessments based on model instructional units that are aligned with the state’s High School Content Expectations. These state standards will be culled down to determine the most important concepts and ideas and then sequenced so that educators will have a model for teaching in a particular subject area. This prioritization of the state standards makes instructional unit-based interim assessments an ideal tool. Interim assessments aligned to state standards also may be developed for K–8.

Computer Adaptive Testing (CAT) takes state-of-the-art testing even further. CAT software determines what questions to ask of an individual student based on his or her test performance during the testing session. CAT software can estimate a student's level of proficiency on a given test more quickly. The assessment begins with a question of medium difficulty, then offers a harder or easier question depending on whether the student answered the initial question correctly or incorrectly. The process continues, presenting questions of varying levels of difficulty, until CAT can estimate with sufficient precision the student's level of performance on the assessment.⁴³

Some computer adaptive testing programs use test items beyond multiple choice, including constructed-response questions, more complicated questions that assess complex problem-solving skills and more innovative items that ask students to organize or move objects on a screen or prepare an essay.

Going online with CAT has helped students in **Virginia**, who have to pass EOC tests for graduation, receive help in mathematics more quickly and flexibly. The state department of education worked with a computer-testing company to design the Algebra Readiness Diagnostic Test System (ARDT) as a web-based application powered by CAT.⁴⁴ The purpose of the ARDT is to provide mathematics intervention services to middle-level students (those in grades 6–9) who may be at risk of failing the Algebra 1 EOC test. ARDT includes a diagnostic pretest to identify student's mathematics strengths and weaknesses, an intervention service that addresses the weaknesses through 2 ½ hours of additional weekly instruction and a post-test.

The assessment is administered in schools on computers connected to the Internet. Results from the diagnostic test are available immediately and provide information correlated to the state standards and assessment reporting categories. Students targeted to take the diagnostic tests include those who were not successful in their previous intervention/remediation program, performed below average in their previous year's mathematics program or did not pass the state mathematics assessment.

Still, Achieve notes that decisionmakers should be cautious when considering computer adaptive testing, particularly if they are thinking of using CAT as the mainstream state assessment. Because students will be tested on a wider range of questions than on a paper-and-pencil test, a sufficiently large item bank must be developed, which can be challenging and costly.⁴⁵ Access to computers, access to and speed of Internet connections, and the quality of hardware may vary across school districts, which could interfere with test administration. State and school leaders need to be concerned about whether the existing allocations of computers to classrooms is sufficient to support the amount of testing planned and whether teachers or testing proctors are trained to deal with unexpected technical issues that may arise during testing.

Core Principle Four: Streamlining Testing Time and Costs

Students already take a lot of tests, especially in high school. As states evolve their assessment systems, they need to be sensitive to the fact that some educators and students are already feeling “over tested.” In addition to tests developed and administered by the state, students take tests mandated by school districts (such as interim assessments and annual EOG tests) and others as part of the college application process (such as PSAT and SAT; Plan, Explore and ACT; and Advanced Placement and International Baccalaureate exams).

The time spent preparing for these tests is not insignificant—both for students and educators—and most would agree the various assessments are not well-aligned with one another, in either their content or their purposes.

It would be unfortunate if states simply layer more tests on top of existing tests. As much as possible, the total amount of high school testing should be streamlined. A substantial increase in testing without good-faith efforts to streamline the system will frustrate policymakers, educators, students and the public.

The goal is not more testing; the goal is smarter testing.

Reducing Testing and Eliminating Redundancy *Within States*

To streamline testing within each state, state leaders will want to take a hard look at how many tests students are currently taking—at both the state and local levels—and what those tests seek to measure. What information do the current tests provide and to whom? Will new tests provide additional or similar information? If similar, which tests can be streamlined or removed?

Once the state maps out its plan for refining the overall high school assessment system, the next step is to examine the net effect on the testing burden of incorporating college- and career-ready anchor assessments *and* local- or state-led performance measures *and* interim assessments.

If the plan, taken as a whole, seems too expansive, the state can consider which planned assessments can be dropped without harming the overall goal of increasing college and career readiness. Or, the state may decide that the rationale for expanding testing is compelling enough to override concerns about the testing burden and, in that case, the state will want to develop careful communications to describe the urgency of implementing the full proposed suite of assessments.

For example, states that want to adapt a national college admissions test as the college- and career-ready anchor assessment should think carefully about what to do with their other tests, such as existing 10th grade tests or EOC tests. While there may be good reasons to have both types of tests—higher education may be more comfortable with the admissions tests than with the EOC tests, for example—this expanded testing will be expensive and potentially too burdensome on schools.

Chapter Highlights

- ⇒ Though this guide advocates for states to broaden their vision of assessment, simply adding more tests, without reducing and streamlining when possible, will frustrate educators, students and the public.
- ⇒ States and districts should collaborate to identify tests that can be eliminated because they serve similar purposes or provide similar information.
- ⇒ States can reduce testing costs and development time by collaborating across state lines to create and use tests in key subject areas and courses.

Another option is to eliminate tests given to nearly all students that may have served fine purposes but don't fit in the future vision of the state assessment program. For example, some states pay for all students to take the PSAT (or Plan and Explore) and offer administration during the school day, without considering how the tests align with state standards or how effectively the results can be used by schools and students to increase college readiness.

Getting a handle on the volume of current testing also will require surveying or auditing local districts to understand which additional tests they administer and why. In addition to interim assessments, many districts give norm-referenced assessments that are not aligned to state standards, and sometimes the tests are given multiple times during the school year!

Taken together, this process will help point out gaps and redundancies in the assessment system.

Once the results are analyzed and then compared against the state's plans to augment, reduce or expand statewide testing, Achieve recommends that states work with districts to streamline the overall amount of tests that students are taking. This will require give and take, but it is in the best interests of students and schools that, as new assessments are added, others are taken away.

States that are planning to build or add EOC tests may discover, for example, that many districts already give final exams at the end of the same courses. It is in the best interest of students and teachers that states work with their districts on a strategy that enables the state EOC tests to replace the local exams rather than create unnecessary duplication.

Once school districts know that the state plans to take responsibility for additional areas of assessments, districts are likely to want to achieve cost savings by eliminating unnecessary tests (such as commercial standardized tests that will be replaced by better interim assessments or replacing district-created and district-scored course final exams with the state's EOC tests).

In addition, the state may decide to take a stronger hand in promoting the use of quality performance measures and interim assessments. But these new measures don't have to be mandatory for all school systems. For example, some high-performing schools and school systems could be exempted from using state-led performance measures or interim tests. Other school systems could be exempted if they can demonstrate to the state that their own measures are equivalent to or of higher quality than the state's tests. Or, the state may decide to disseminate performance tasks or new interim assessments to all districts statewide but only require low-performing districts to use the state measures.

Reducing Testing Costs and Redundancy *Across States*

States also can reduce testing costs and development time by collaborating across state lines to create and use tests in key subject areas and courses. For example, the 14 states currently participating in the ADP multi-state Algebra 2 EOC exam are working together to create a higher-quality test faster and at a lower cost to each state. Because of the long-term potential for large numbers of students to take the Algebra 2 EOC exam, at its own expense, the test contractor Pearson developed the test items to meet the specifications agreed upon by the nine original states. The base price for the testing program is determined by the total number of tests ordered across the states participating in the ADP Algebra Consortium for a school year. The per-student cost to administer and score the test even decreases as more students participate.

Table 8: The More States Collaborate, The More Costs Can Go Down: Per-Student Costs for the Algebra 2 Multi-State End-of-Course Test

| Number of Tests Ordered | Administrations: Fall 2008–Spring 2011 | Administrations: Fall 2011–Spring 2013 |
|-------------------------|---|---|
| | PRICE (\$) | PRICE (\$) |
| Fewer than 100,000 | \$26.06 | \$23.56 |
| 100,000 – 149,999 | \$21.06 | \$17.96 |
| 150,000 – 199,999 | \$19.58 | \$15.94 |
| 200,000 – 249,999 | \$18.06 | \$14.17 |
| 250,000 – 299,999 | \$17.06 | \$13.05 |
| 300,000 and greater | \$16.06 | \$11.97 |

Source: American Diploma Project Algebra I and Algebra II End-of-Course Exam Pricing Schedule, www.achieve.org/node/842

States in the New England Assessment Consortium (**Maine, New Hampshire, Rhode Island and Vermont**) collaborate for similar reasons and with similar results. The bonus of these partnerships is the ability for states to compare their performance and progress with one another.

Getting from Here to There: Key Questions to Inform State Action

States that consider the analyses and recommendations in this report and then emphasize and use better assessments in their high school testing systems should find over time that more students are prepared for successful transitions to college and careers.

Yet there is no single “right” way to refine the high school assessment system to meet the four core principles described in depth in the previous sections. Instead, each state needs to convene cross-sector leaders—such as K–12 and higher education system leaders, the governor’s office, legislative leadership and employers—to weigh carefully the assessment policy and resource options they face.

This chapter is intended to inform each state’s deliberations by considering key questions and offering illustrations—including a “decision tree”—to help winnow the options states face.

Key Questions To Consider

The following questions can help states work through assessment policy choices and tradeoffs.

Core Principle One: “Proficient” Should Mean Prepared

If your state doesn’t currently administer a test of college and career readiness to all of your students—or is dissatisfied with your current test—what assessment approach will fit best?

- ⇒ Which of the three approaches (EOG, EOC or adapted admissions test) will be the most effective measure(s) of college and career readiness?
- ⇒ Is it wiser to augment one or more of your current state tests or look for a better assessment for this purpose? Which strategy will enable the strongest alignment with your state’s high school standards?
- ⇒ Can your state build from the assessments it already has or do you need a wholesale change?
- ⇒ How important is it for the assessments to be taken in close proximity to when students learn the material? Does your state want a system that allows for more or less flexibility in terms of when students take the tests?
- ⇒ How important is it for the test results to be portable across state lines for students going out of state to college?
- ⇒ Which assessments will provide the best feedback to teachers and schools? To students and parents? Which assessments are best used to hold high schools accountable?

Core Principle Two: Making Tests Matter Beyond K–12

How does your state make sure that high school assessments have the necessary credibility with higher education and employers so that performance on the assessments allows students to access postsecondary opportunities?

- ⇒ What will it take to get higher education to use the high school college- and career-ready anchor test(s) for placement decisions? What will it take to establish a college-ready cut score? To establish a common college placement standard?
- ⇒ Have colleges in your state agreed on a common placement standard for college placement tests? How can that common standard guide the development of the high school college-ready cut score?
- ⇒ Who are the higher education leaders that need to be brought into the process?
- ⇒ If your state does not have a few key leaders who can facilitate this process, what policies can the state enact to ensure collaboration in this process?

What incentives will your state attach to the new high school college- and career-ready assessments so that students and schools are motivated, but not overly obsessed, about performing well on them?

- ⇒ What incentives will be attached to the state tests for students? How can your state make these tests “count” in some way for students to ensure they take the test seriously, without making the tests too high stakes, too soon?
- ⇒ Should your state continue to require students to pass other tests to graduate from high school? If so, how do the content and skills tested on those exams line up with the college- and career-ready tests?
- ⇒ How will the college- and career-ready tests be used for school and district accountability?
- ⇒ How will other tests given in high school be used for school and district accountability?

Core Principle Three: Not Just More Assessment—Better Assessment

Are your state’s high school summative tests (including but not limited to the college- and career-ready anchor assessment) assessing the full range of skills students need to be college- and career-ready?

- ⇒ Has your state done a gap analysis of high school assessments and standards to determine which skills are being measured well and which are not?
- ⇒ Is your state making ample use of constructed-response questions on statewide assessments or are you too heavily relying on multiple-choice questions?
- ⇒ Do score reports tell students and families if they are on track to graduate ready for college and careers?

Does your state need to make better use of performance-based measures?

- ⇒ Are any performance assessments included in your state's high school tests? Which standards are most appropriate to measure through performance-based assessments?
- ⇒ What kinds of performance measures are already being used by schools, school networks or school districts? Can these be expanded or scaled to other districts? Can universities help with R&D and piloting of new measures?
- ⇒ How will your state include teachers in the selection and development of performance components for statewide administration? Should the state set specifications and parameters for the skills that need to be assessed and leave some local discretion for designing the tasks and prompts? How will the state ensure fidelity to the standards and consistency of scoring across classrooms, schools and districts?

Has your state made supporting instruction a high enough priority in the assessment system?

- ⇒ Which school districts have invested in developing or identifying aligned, high-quality interim assessments? Are there particular districts or schools that do not have access to a full set of interim assessments, or where the quality of such assessments is low or unknown? Can some districts' tests be shared with other districts as exemplars?
- ⇒ Are there districts or charter schools that are using high-quality interim assessments? Is there a way to partner with them to make these tests available to others? Has your state determined whether commercially available interim assessments are well-aligned with the state's standards?
- ⇒ Has your state made available an online item bank that will help teachers develop their own formative assessments that are aligned with state standards?
- ⇒ What is your state doing to ensure that teachers get professional development on how to interpret results from interim assessments and use them to improve instruction? Are teachers getting help learning how to develop and use their own formative assessments? Are there any standards or requirements for ensuring that teachers begin this development in pre-service preparation programs?

Core Principle Four: Streamlining Testing Time and Costs

How can your state make sure that testing does not become too overwhelming or expensive?

- ⇒ How many tests and how much testing time will be added, once principles one through three are addressed? What's your state's appetite for an expansion of testing?
- ⇒ If leaders want to combine any of the three major approaches (EOG, EOC or adapted admissions tests), is the rationale for using multiple kinds of college- and career-ready assessments compelling enough to override concerns about the testing burden?
- ⇒ Which tests could be subtracted as new ones are added? Which district-administered tests (such as standardized norm-referenced tests) can be eliminated as high-quality state-led tests (such as better interim assessments) become available? Can any of the assessments be made optional (such as exempting high-performing systems from using state-developed interim assessments)?

- ⇒ Can higher education reduce the number of college placement tests by setting a common college readiness standard and guaranteeing placement for students who meet the standard late in high school?
- ⇒ Can states work together to reduce testing costs and become better consumers with the testing industry?

Charting Each State’s Own Path to a New High School Assessment System: A “Decision Tree” for State Policymakers

As each state considers the key questions posed above to unpack assessment options, we hope the following “decision tree” can be a tool to help guide state discussions. It’s important to remember that there is no single path or “right” answer. Each state’s context, history, traditions and capacity will ultimately determine its approach to assessing college and career readiness. By beginning with your state’s existing tests as a starting point and honoring your state’s own traditions, we hope a coherent picture of a next-generation high school assessment system will emerge.

The first view of the “decision tree” is a summary graphic. More detailed options and considerations follow as Table Nine.

A “Decision Tree” for State Policymakers

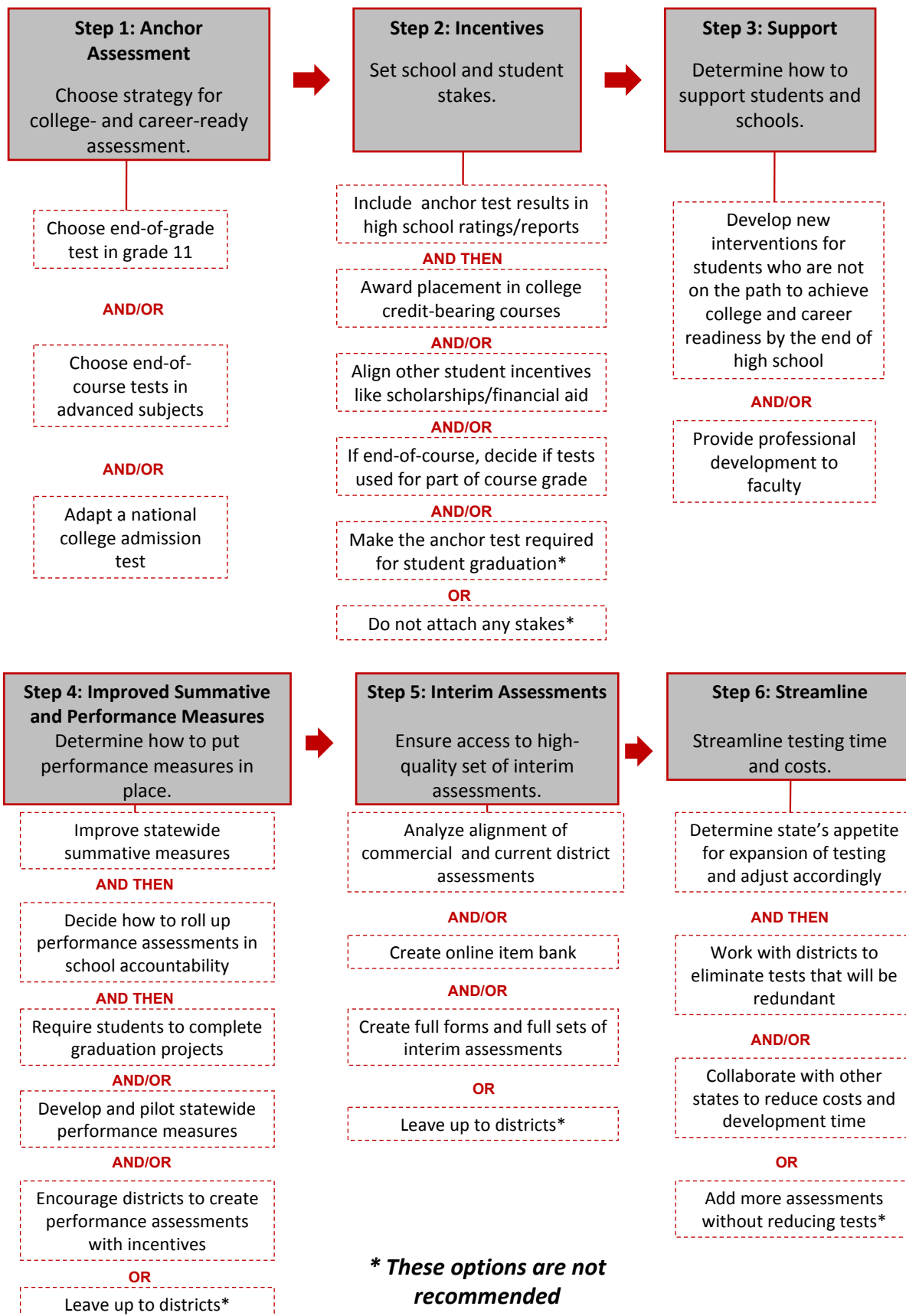


Table 9: Steps and Options to Address Core Principles

| | Key Decisions | Options | Key Considerations |
|---|---|---|---|
| <p>Step 1: Choose the college- and career-ready anchor assessment.</p> | <p>Choose strategy for the anchor assessment. Consider whether to adapt or build from the existing assessments or if wholesale change in the system is warranted.</p> | <p>Option 1: End-of-grade comprehensive assessments</p> | <p>Tests must assess rigorous 11th grade content that is validated by higher education as college-ready.</p> |
| | | <p>Option 2: End-of-course assessments</p> | <p>Develop end-of-course tests in advanced courses such as Algebra 2 or English 11 and validate by higher education to ensure sufficient credibility. Consider how many additional end-of-course tests to use and in what subjects.</p> |
| | | <p>Option 3: Adapted college admissions tests</p> | <p>Take stock of what these tests measure and how well they are aligned with state standards. Augment the assessments to ensure alignment with state standards. Achieve does not recommend using college placement tests, because its research has found that placement tests are too narrowly focused on a subset of knowledge and skills.</p> |

| | | | |
|---|--|--|---|
| <p>Step 2: Decide what incentives for students to attach to the anchor assessment.</p> | <p>In collaboration with higher education, determine a cut score that demonstrates college and career readiness.</p> <p>Second, hold high schools accountable for performance on the anchor assessment.</p> <p>Next, determine the range of student incentives to use.</p> | <p>First, incorporate the anchor assessment results in the high school accountability system.</p> | <p>High schools should be held accountable for a wider variety of indicators than today (such as coursetaking, graduation rate, more assessments). The college- and career-ready anchor test should be given primary weight among test results used to identify, support or intervene in high schools.</p> |
| | | <p>Option 1: Guarantee placement in credit-bearing entry-level college courses in mathematics and English</p> | <p>Waiving additional placement testing once students get to college is a good way to streamline testing overall and motivate and support college readiness. Tying additional “opening doors” incentives, such as financial aid, college admissions or waiving employer tests, also can motivate and support student success.</p> |
| | | <p>Option 2: Identify and award additional incentives (such as scholarships or apprenticeship advancement)</p> | |
| | | <p>Option 3: Factor results into course grade</p> | <p>This option only works with end-of-course tests. Test results must be returned before the end of the school year.</p> |
| | | <p>Option 4: Make passing the anchor assessment a graduation requirement</p> | <p><i>Not recommended.</i> States should keep the bar set at college and career readiness, which would be very difficult to do currently if passing the test were required for graduation. States with end-of-course tests in introductory subjects (such as Algebra 1 or English 9) could use these for high school exit. On comprehensive exams, a lower cut score for high school exit could be set in addition to the college- and career-ready cut score (with the goal that eventually these scores will converge).</p> |
| <p>Option 5: Do not attach incentives</p> | <p><i>Not recommended.</i> Students need to be motivated to perform on the anchor assessment.</p> | | |

| | | | |
|--|---|---|--|
| <p>Step 3: Develop supports for students and schools to achieve college and career readiness.</p> | <p>Create comprehensive supports and interventions for middle and early high school students. Once college and career readiness scores are available to high school seniors, support students who need more time and help to get ready.</p> | <p>Option 1: Develop new courses, modules and interventions for students who are struggling in middle and early high school and students who are not college- and career-ready on the anchor assessment</p> | <p>Use data from assessments given in 8th, 9th and 10th grades, along with course grades and other key early indicators, to intervene early with struggling students.</p> <p>For courses in the junior and senior years, encourage K–12 and higher education institutions to partner and collaborate to develop new courses and teaching units and support better teaching and student success. If resources are limited, use results from anchor assessment to focus resources on mathematics or English.</p> |
| | | <p>Option 2: Provide professional development to high school faculty</p> | |
| <p>Step 4: Improve summative tests and determine how to incorporate performance measures.</p> | <p>Statewide summative tests, including the college- and career-ready anchor assessments, should maximize usefulness for improving teaching and learning.</p> <p>In addition, performance measures should be part of the state system.</p> <p>Determine how to put the measures in place.</p> | <p>Improve summative tests to maximize their usefulness for teaching and learning</p> | <p>Use constructed-response items efficiently and appropriately. Involve teachers in scoring. Create better Web-based data analysis tools. Return results as quickly as possible and at strand levels if possible.</p> |
| | | <p>Decide if performance assessments should “roll up” to factor as one measure in high school accountability or if they have stakes for students</p> | <p>If the performance measures do <i>not</i> count for students or schools, then the <i>reliability</i> of the instruments matters less and <i>validity</i> could be increased to emphasize usefulness for teaching and learning. If the measures <i>do</i> count, then reliability must be increased even if validity decreases.</p> |
| | | <p>Option 1: Require all students to complete graduation projects</p> | <p>Determine level of state involvement in scoring graduation projects. Establish clear performance criteria. Disseminate student work. Provide professional development. Establish auditing to ensure reliability.</p> |
| | | <p>Option 2: Develop and pilot performance measures</p> | |
| <p>Option 3: Encourage districts to create performance assessments via incentives</p> | | | |
| | | <p>Option 4: Leave entirely up to school districts</p> | <p><i>Not recommended.</i> Not enough districts have capacity to do performance assessment well on their own.</p> |

| | | | |
|---|---|---|---|
| <p>Step 5: Determine state role in providing leadership for and access statewide to interim assessments.</p> | <p>Ensure that all schools and districts have access to at least one full set of high-quality interim assessments.</p> | <p>Option 1: Analyze alignment of commercial and/or district-created interim assessments with state standards</p> | <p>Least costly/intrusive state role; publish alignment analyses and consider what to do if analyses indicate that no existing packages align with state standards.</p> |
| | | <p>Option 2: Create online item banks aligned to state standards</p> | <p>Most flexible approach; provide schools and districts with opportunity to “mix and match” items as needed. May not result in consistency of rigor or quality across districts.</p> |
| | | <p>Option 3: Create full forms and sets of aligned interim assessments</p> | <p>Most likely to promote consistency and quality of instruction, but potentially also the costliest approach. Align to standards and anchor assessments. Tie full sets to core units of instruction within key high school courses.</p> |
| | | <p>Option 4: Leave entirely up to school districts</p> | <p><i>Not recommended.</i> Not enough districts have capacity to do interim assessment well on their own.</p> |
| <p>Step 6: Take stock of overall testing system and streamline testing time and costs.</p> | <p>Step back and examine the net effect on the testing burden of incorporating the anchor assessment, performance measures and interim assessments into the overall high school testing system.</p> | <p>First, determine state’s appetite for expansion of testing and adjust plans accordingly</p> | <p>Consider which assessments can be dropped without harming overall goal of increasing college and career readiness. Or decide if the rationale for testing expansion is compelling enough to override concerns about the testing burden. Eliminate tests that don’t align with state standards.</p> |
| | | <p>Option 1: Work with districts to eliminate tests that will now be redundant</p> | <p>Identify current district assessments that will become redundant once the anchor, performance and interim assessments are in place. Help districts identify cost savings from eliminating unnecessary tests. Decide which districts are exempt from using state performance or interim measures.</p> |
| | | <p>Option 2: Collaborate with other states to reduce costs and development time</p> | <p>Save development time and cost and put pressure on testing contractors to improve product.</p> |
| | | <p>Option 3: Add more assessments without reducing tests</p> | <p><i>Not recommended.</i> A substantial increase in testing without good-faith efforts to streamline the system will frustrate policymakers, educators, students and the public.</p> |

Acknowledgements

Achieve would like to thank Jennifer Vranek of Education First Consulting and Achieve's state leadership consulting team for writing this guide.

Alissa Peltzman, Achieve's ADP Network Manager, and Matt Gandal, Achieve's Executive Vice President, provided considerable strategic and editorial guidance.

Thanks also to the following Achieve consultants and staff who provided advice and feedback during the development of the guide: Susan Bodary, Marc Frazer, Lisa Gray, Scott Jenkins, Bill Porter, Todd Webster, Kate Blosveren, Michael Cohen, Joanne Eresh, Kaye Forgione, Tracy Halka, Lesley Muldoon, Jean Slattery, Laura Slover and Robert Triscari.

Thanks to Tom Flavell of Achieve for layout and Sionainn Marcoux of Education First Consulting for creating the report's graphics.

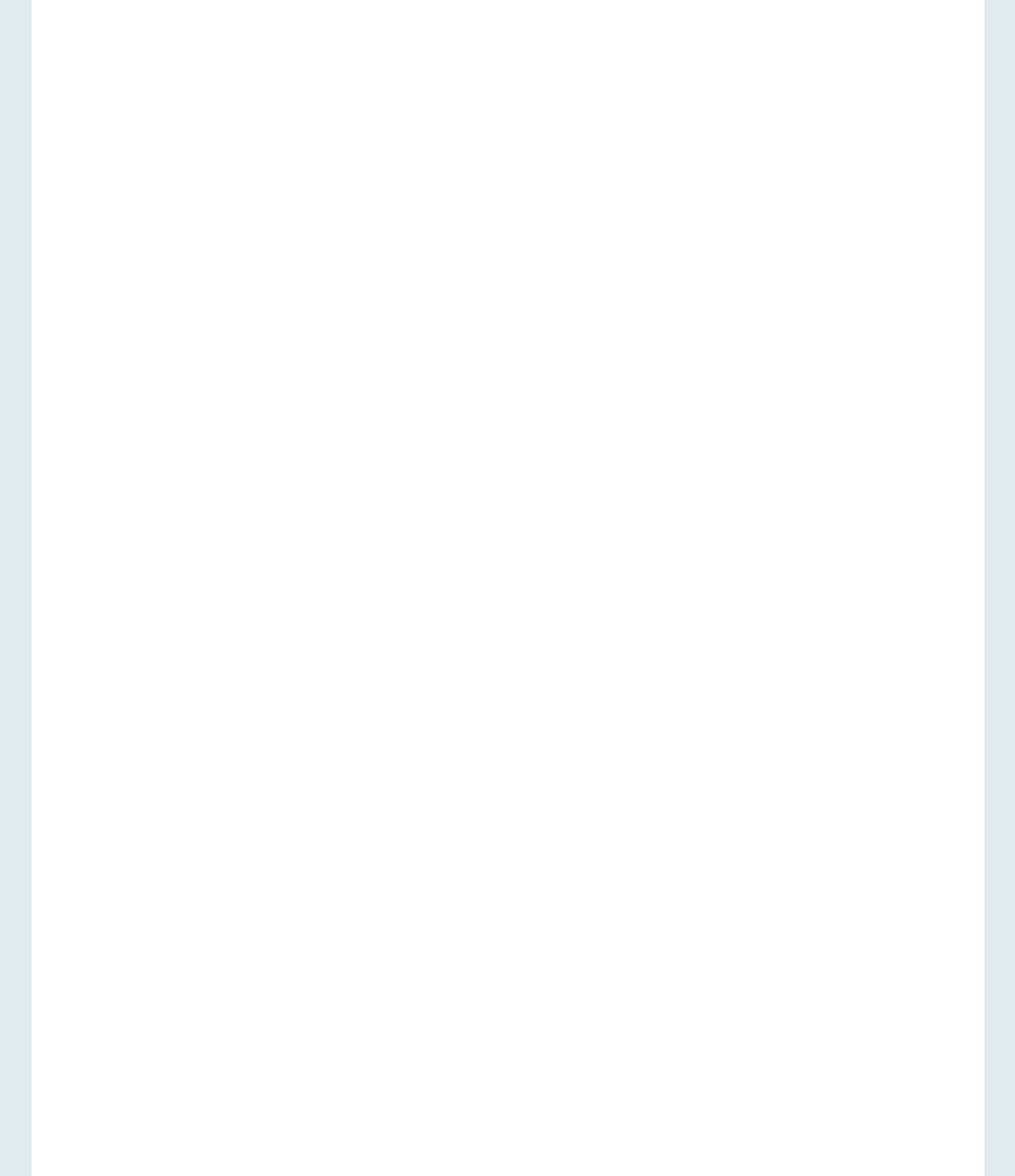
Finally, Achieve would like to thank the Bill & Melinda Gates Foundation for providing generous funding for this effort and for the broader work of the American Diploma Project.

Endnotes

- ¹ Achieve, *Do Graduation Tests Measure Up?: A Closer Look at State High School Exit Exams*, 2004.
- ² Achieve, *Closing the Expectations Gap 2008: An Annual 50-State Progress Report on the Alignment of High School Policies with the Demands of College and Careers*, 2008.
- ³ Achieve and Education Trust are releasing on their Web sites several other companion white papers, including a paper on career readiness credentialing and two others on the inclusion of special education and English-language-learner students in assessment systems.
- ⁴ Perie, Marianne, Scott Marion, Brian Gong and Judy Wurtzel, *The Role of Interim Assessments in a Comprehensive Assessments System: A Policy Brief*. Prepared for Achieve, the Aspen Institute and the National Center for the Improvement of Educational Assessment Inc., 2007.
- ⁵ Achieve, *Closing the Expectations Gap 2008*, 2008.
- ⁶ Achieve, *Closing the Expectations Gap*, 2008.
- ⁷ Achieve, *Ready or Not: Creating a High School Diploma that Counts*, 2004.
- ⁸ Achieve, *Closing the Expectations Gap*, 2008.
- ⁹ Education First Consulting, *End-of-Course Tests: A Report to the Washington State Board of Education*, January 2008.
- ¹⁰ Interestingly, in the 1920s and '30s, the Regents exams expanded to include many vocational fields and the testing system served somewhat to sort students into academic programs or vocational programs during and after high school. By the 1970s, the vocational Regents exams were discontinued and many of the subject-specific exams were consolidated. New York State Education Department, Bureau of Elementary and Secondary Testing Programs, *History of Regents Examinations 1865 to 1987*. Available at www.emsc.nysed.gov/osa/hsinfo/gen/hsinfo/genarch/rehistory.htm.
- ¹¹ Section 100.5(a)(5) of the Regulations of the Commissioner of Education Relating to the General Education and Diploma Requirements Regarding the State Assessment System, State of New York.
- ¹² New York State Education Department, *School Administrator's Manual 2008: Regents Examinations, Regents Competency Tests, and Second Language Proficiency Examinations*. Available at www.emsc.nysed.gov/osa/sam/secondary/home.html.
- ¹³ Section 100.5(b)(7)(ix) of the Regulations of the Commissioner of Education Relating to the General Education and Diploma Requirements Regarding the State Assessment System, State of New York.
- ¹⁴ Education First Consulting, *End-of-Course Tests*, 2008.
- ¹⁵ City University of New York and the New York City Department of Education, *College Now: Tests*. Available at collegenow.cuny.edu/nextstop/finish_hs/tests/#nys_regents.

- ¹⁶ Math A and B will be phased out after June 2010 in favor of Algebra 1, Geometry and Algebra 2 Regents exams.
- ¹⁷ Office of the Provost of the State University of New York, *Guidelines for the Approval of State University General Education Requirement Courses*. Available at www.suny.edu/provost/generaleducation/course-guidelines-v2.txt.
- ¹⁸ Achieve, *Closing the Expectations Gap*, 2008.
- ¹⁹ ACT Inc., “Facts about the ACT.” Available at www.act.org/news/aapfacts.html.
- ²⁰ College Board, “SAT Reasoning Test.” Available at www.collegeboard.com/student/testing/sat/about/SATI.html.
- ²¹ College Board, “Predictive Placement Validity Studies.” Advisement, Placement & Retention. Available at www.collegeboard.com/highered/apr/aces/vhandbook/predictive.html.
- ²² Achieve, *Aligned Expectations?: A Closer Look at College Admissions and Placement Tests*, 2007.
- ²³ Muller, Robert D. and Alexandra Beatty, Practical Strategy, LLC, *Work Readiness Certification and Industry Credentials: What Do State High School Policy Makers Need To Know?* Discussion paper prepared for Achieve, December 2007.
- ²⁴ 511 IAC 6-7.1-7 of the Indiana Administrative Code Regulations Regarding Core 40 Diploma with Technical Honors, 2005.
- ²⁵ Muller, Robert D. and Alexandra Beatty, Practical Strategy, LLC, *Work Readiness Certification and Industry Credentials: What Do State High School Policy Makers Need To Know?* Discussion paper prepared for Achieve, December 2007.
- ²⁶ California State University, *California Early Assessment Program Augmentation of STAR Standards*. Available at www.calstate.edu/eap/augmentation.shtml.
- ²⁷ Personal communication with Achieve, August 2008.
- ²⁸ Ohio Board of Regents, *The Strategic Plan for Higher Education 2008–2017*. Available at uso.edu/strategicplan/index.php.
- ²⁹ The National Center for Public Policy and Higher Education, *Policy Alert: Mixed Signals in California*, March 2008; California State University, *Fall 2006 Freshman Proficiency at Entry (Fall 2006) and One Year Later (Fall 2007) – Systemwide*, available at www.asd.calstate.edu/remrates/06-07/systemwide.htm.
- ³⁰ California State University, *Early Assessment Program Participation and Results*. Available at eap2008.ets.org/Viewreport.asp.
- ³¹ California State University, *Standards-Based Teacher Preparation in the California State University: Annual Report*, 2006.
- ³² Perie, Marion, Gong and Wurtzel, *The Role of Interim Assessments in a Comprehensive Assessments System*, 2007.
- ³³ EPE Research Center, *Quality Counts 2008: Tapping into Teaching*, 2008.
- ³⁴ Massachusetts Department of Education, *2007 MCAS Technical Manual*. Available at iservices.measuredprogress.org/documents/MA/Technical%20Report/04-09-08%202007%20MCAS%20Tech%20Rpt%20Final%20PDF.pdf.

- ³⁵ For example, see the 2008 released items at www.doe.mass.edu/mcas/2008/release/default.html and scoring guides and examples of student work at www.doe.mass.edu/mcas/student/.
- ³⁶ Unpublished discussion paper prepared for Achieve and the Education Trust by Brian Gong, the National Center for the Improvement of Educational Assessment, September 2007.
- ³⁷ Achieve, *Closing the Expectations Gap*, 2008.
- ³⁸ EPE Research Center, *Quality Counts 2008*, 2008.
- ³⁹ New Jersey Performance Assessment Alliance, “What’s Killing the Fish?: High School Science/Biology Teacher Guidelines,” 2006. Available at www.state.nj.us/education/aps/cccs/science/bio_guidelines.pdf. Sample Student Prompt Booklet, Student Response Booklet and Teacher Guidelines Booklet are available at www.nj.gov/education/aps/cccs/science/resources.
- ⁴⁰ Oregon Department of Education, *Assessment of Essential Skills*, available at www.ode.state.or.us/search/page/?id=2042; Oregon Department of Education, *Work Samples*, available at www.ode.state.or.us/search/page/?=219; Washington Office of Superintendent of Public Instruction, *Certificate of Academic Achievement Options*, available at www.k12.wa.us/assessment/CAAoptions/default.aspx; Wyoming Department of Education, *Standards, Assessment and Accountability: The Body of Evidence*, available at www.k12.wy.us/SAA/BOE.asp.
- ⁴¹ Achieve, *Creating a World-Class Education System in Ohio*, 2007.
- ⁴² Personal communication with Achieve, August 2008.
- ⁴³ Hamilton, Laura S., Stephen P. Klein and William Lorie, *Using Web-Based Testing for Large-Scale Assessment*, RAND Education: Report #IP-196, 2000.
- ⁴⁴ Virginia Department of Education, *Virginia Algebra Readiness Initiative*. Available at www.doe.virginia.gov/VDOE/Instruction/Math/ARI/index.shtml.
- ⁴⁵ Wang, Tianyou and Michael J. Kolen, “Evaluating Comparability in Computerized Adaptive Testing: Issues, Criteria, and an Example,” *Journal of Educational Measurement* 38(1), 19–49, 2001.



Measures *that* Matter

