

HIGH SCHOOL ASSESSMENT IN A NEW ERA: WHAT POLICYMAKERS NEED TO KNOW



Under No Child Left Behind (NCLB) and now the Every Student Succeeds Act (ESSA), states are required to assess every student at least once in high school in reading or language arts, mathematics, and science, and some states have chosen to assess students more frequently.¹

The passage of ESSA gives states a new opportunity in how they approach high school assessment. Most significantly, it will allow states to use “nationally-recognized high school academic assessments” such as college entrance exams (ACT or SAT²) as their sole high school assessment for accountability purposes, and give states the option to permit individual districts to do so in lieu of the state high school test. In both cases, states will have to show that these assessments are technically sound and are aligned to the states’ academic content standards, although the rules for implementing this part of ESSA have not been finalized.

Given these changes, states now face important decisions about which high school assessments to use and for what purposes. In order to approach these decisions in a thoughtful manner and find the best solution for their state, there are several critical issues that policymakers need to understand.

The purpose of this brief, developed jointly by Achieve and the National Center for the Improvement of Educational Assessment, is to help state policymakers make sense of these issues and identify key action steps to consider as they make decisions about high school assessments. While this brief is focused on high school assessment, it is important to remember that the overarching goal is to create a coherent and aligned assessment system (across elementary, middle, and high school) that provides meaningful information to all students about their progress in becoming college and career ready—the end goal for students in every state.

The Issues

Purposes for High School Assessment

The most important issue for policymakers to consider before selecting a high school assessment is the purpose(s) for that assessment. There are a variety of purposes/uses for high school assessment. Some general categories include:

- School and district accountability – providing achievement and growth data for annual determinations of school and district performance
- Admission to college
- College and career readiness – placement in credit-bearing courses without the need for remediation
- Feedback to individual students – reporting on progress in mastering state standards (achievement and growth) and becoming college and career ready
- Determining student readiness for high school graduation
- Awarding course credit or grades
- Providing a signal to educators about the kind of instruction that is needed for students to be college and career ready by including items that require students to engage in complex cognitive tasks and apply their knowledge and skills

¹ The passage of the Improving America’s Schools Act (IASA) in 1994 established the requirement that all states have a high school English language arts (ELA)/literacy and mathematics assessment. While many states previously had high school assessments, every state was in compliance with that mandate in 2000.

² The SAT was developed by the College Board.

No single assessment can adequately support all of these purposes. For example, college entrance exams are a measure of college readiness, but generally do not provide specific feedback on student mastery of state standards or growth towards mastery. Individually state-developed assessments provide data on student mastery of standards, but are not typically accepted by higher education institutions for admission or placement purposes, although they have been used to determine the need for remediation. Some assessments may reasonably work for more than one purpose, such as PARCC and Smarter Balanced, which provide feedback on student mastery of the Common Core State Standards (CCSS) and are used by some higher education institutions for placement. However, it is unlikely that any assessment could support more than a few purposes, particularly when cost and testing time constraints are taken into account. Therefore, states will have to compromise by either:

- Adopting a system of assessments that work together to support most desired uses – this can be done in a variety of ways, but it can be technically challenging and may result in additional state tests, more testing time, and additional costs
- Selecting a single assessment appropriate only for the highest priority purposes

Either way, the purpose(s) for high school assessment should be the fundamental basis for choosing a high school assessment strategy.

Action Steps

1. Provide opportunities for all stakeholders to have a voice in identifying the purposes/uses for high school assessments.
2. Decide which purposes are the most important and keep the focus there.

Use for School and District Accountability

To be useful for school and/or district accountability, an assessment must meet a number of technical requirements. Ensuring the technical quality of a high school assessment can be a complex process. The role of policymakers is to require the vendor or state agency to produce evidence regarding the technical aspects of the assessment. They should ask:

- Is there adequate evidence that the assessment measures what it is supposed to measure?
- Is there adequate evidence that the assessment is aligned to state standards?
- Is there adequate evidence that the assessment accurately measures student performance of both low and high performers in addition to typical performers?
- Is there adequate evidence that the assessment can be appropriately used to measure growth as well as achievement?
- Are there procedures in place to ensure that the test content is kept secure?
- Can the logistics of the administration procedures be managed by all districts and schools? (This is especially important if the assessment is computer-based.)
- Is there adequate evidence that the assessment format (including available accommodations) allows all students (especially students with disabilities and English learners) to show what they know?

Action Steps

1. Ask for independent evidence regarding the validity of the assessment, ability to accurately measure all levels of student performance, and capacity to measure student growth.
2. Require an independent alignment study to ensure that the assessment is aligned with state standards. This should include an analysis of the alignment of items and test forms with the standards.
3. Require a pilot of the administration procedures be conducted with at least a representative sample of districts and schools from across the state.
4. Require independent verification that all students can access the assessment (i.e., can show what they know).

Credibility with Stakeholders

The stakeholder groups for a state's high school assessment include high school students and their parents, high school teachers, counselors and administrators, institutions of higher education, education policymakers, and the public. Each of these groups wants and expects particular things from the state test. At times their needs overlap and intersect, but they are often in conflict. Therefore, what it means for an assessment to be credible will differ by constituency.

As it is with accountability, technical quality is important in building credibility with stakeholders. Policymakers and the public want to know that the assessment measures what it is supposed to measure and that the inferences made about individual student, school, and district performance based on the results from the assessment are justified.

One of the major concerns with statewide assessments at the high school level that are used for school accountability (particularly survey tests given to all students in a grade) is student motivation. This is a phenomenon that most commonly occurs in the United States. In most other countries, high school assessments are tied to specific courses and curriculum and are used for college entrance and/or high school graduation determinations or other certifications. There is now a concern in the U.S. that unless students have some "skin in the game" they are not likely to take state high school assessments seriously. For example, the willingness of students/parents and high school educators to participate in or support the state testing program is to some extent dependent on the degree to which higher education institutions (particularly those within the state) use the results from state assessments for admission, placement, and exemption from remediation. However, this version of "skin in the game" is applicable only to those interested in a college education. Other ways to increase the motivation for all students to perform well on high school assessment include counting performance on the assessment for high school graduation or as part of a grade in a class.

The impact on instruction is particularly important to teachers. They want useful and timely information that helps them understand where their students are in meeting important learning targets and must believe that the information they receive is worth the cost in instructional time to administer the assessment (this includes class time spent preparing for the test). Teachers also want the state test to signal what the state learning goals (in terms of the state standards) look like in practice and help them design instruction so that all students reach those goals. This will require test items that ask students to write about what they have read and show their thought process and problem-solving strategies in math as well as other complex cognitive tasks.

Students and parents will be more likely to find the test credible if individual student score reports are informative, timely, and easy to understand. Working with parents and students, Achieve has developed some examples of student score reports that students and parents value. These are available at: <http://www.achieve.org/samplestudentreports>.

College entrance exams tend to have "brand" credibility with students, parents, educators, and higher education because of their long histories, predictive validation evidence for first year college grades, and their use by a majority of colleges, universities, and scholarship agencies in the U.S. However, "brand loyalty" can be due in large part to

familiarity. For example, several states that had given the ACT to all 11th graders for a number of years recently switched to the SAT. This generated a great deal of pushback from superintendents and high school principals as well as students and parents, even though the SAT has “brand” credibility as a college entrance exam.

The multi-state assessment consortia (PARCC and Smarter Balanced) have partnered with leading postsecondary educators to grow the number of institutions that accept PARCC and Smarter Balanced college-and-career-ready scores for placement into credit-bearing courses. To date, nearly 300 institutions across the United States are using Smarter Balanced and/or PARCC scores to exempt students from remediation. This acceptance comes after a number of years working directly with postsecondary educators to design and develop the assessments. Like with any new assessment program, it will take future research to gather the appropriate evidence to assure additional institutions of their validity for that purpose. The number of institutions using Smarter Balanced or PARCC for placement may grow over time based upon this research and feedback from institutions already using the scores.

Action Steps

1. Ensure that all stakeholders understand and support the primary purpose(s) for the high school assessment.
2. Involve a broad group of stakeholders in the assessment selection process that includes representatives from public and private higher education institutions as well as teachers and content experts.
3. Involve students and parents in the development of student score reports in the case of state developed tests, or as part of the state procurement process if the state is considering adopting a college entrance exam or some other “off the shelf” assessment.

Alignment with State Standards and College and Career Readiness

Because states have identified college and career readiness as a central goal for all students, the need for assessments that measure student progress in becoming college and career ready has become critically important. High school assessments should let students (and their parents) know if they are prepared for college and/or to enter the world of work. They should also indicate how well schools and educators are doing in preparing all students to be successful after high school.

The consortia-developed assessments were designed to measure student mastery of the CCSS and provide an indicator of whether students are college and career ready. However, it is too soon to have validation evidence to support this claim because students who first took these assessments have not yet graduated from high school.

Individual state assessments are designed to measure student proficiency on state standards, and the degree to which they measure college and career readiness is dependent on the nature of the particular state’s content standards as well as the assessment design.

College entrance exams were not originally designed to assess specific state standards, although some states augmented the ACT to create a measure of their standards under NCLB. The SAT has been revised to be reflect college and career readiness standards according to the College Board, but only a few alignment studies have been completed.

Action Steps

1. Ask for an independent alignment study to determine the degree to which the high school assessment is aligned with college- and career-ready and state standards.
2. Request an ongoing analysis of how well scores on the state high school assessment predict success in college and career training programs.

Students' Ability to Show What They Know

State high school assessments must provide an opportunity for all students to show what they know and can do. This is particularly important for students with disabilities and English learners (ELs). To accomplish this, some changes in how the test is administered to those students (called accommodations) could be allowed to enable particular students to take the assessment without changing what the assessment measures. For example, an EL student might have portions of the math test read to her. Since the assessment is measuring math knowledge and skills, this accommodation is thought not to change what is being assessed.

The consortia-developed high school assessments employed a “universal design process” so that accessibility for students with disabilities and ELs was built in from the beginning. The same is typically true for new state assessments, although it would be important to have an “accessibility analysis” conducted by an independent entity.

Accommodations for students with disabilities and ELs can be problematic with college entrance assessments. The types of accommodations for students with disabilities and ELs provided by college entrance assessments may not match the accommodations that states and local education agencies have in their policies and procedures. Several federal laws (e.g., the Individuals with Disabilities Education Act [IDEA], ESSA) govern how parent/educator teams determine appropriate accommodations for students with disabilities and ELs. Both vendors require that requests for accommodations be approved by their staff, and ACT and College Board may deny those accommodations. Students can still choose to take the ACT or SAT using accommodations denied by the ACT or the College Board, but the scores will not be official scores submitted to colleges or universities for admissions decisions. Some states are working with these vendors to develop agreements about acceptable accommodations.

Another issue to consider here is the selective availability of test prep materials. Both college entrance exams have associated preparation materials and commercial prep programs. Because of the cost, students from wealthy families are more likely to have access to these resources than their low-income peers. The test companies are aware of this issue and have begun to address it themselves. For example, SAT has developed an agreement with Khan Academy to make prep materials available to students free of charge, and ACT has posted additional free prep materials on its website. However, wealthy families are still more able to access the commercial prep programs. There is some question about the effectiveness of these materials and programs in raising scores, but it is important that all students have access to these preparatory materials and programs to ensure fairness and provide meaningful, comparable scores for students, schools, and districts. Of course, the best test prep is always rigorous coursework and high quality instruction.

Action Steps

1. Have an “accessibility analysis” of the proposed high school assessment conducted by an independent entity. This could be part of a larger alignment study.
2. Involve Special Education and EL teachers in the assessment selection process, particularly in the review and identification of acceptable accommodations.
3. If a state is considering administering a college entrance exam as part of its accountability system, provide access to preparation materials for all students, regardless of income.

Cut Scores and Performance Levels

Cut scores and performance levels may seem like technical issues for testing experts. However, they are important issues for policymakers because they have the potential to impact stakeholder credibility and the utility of high school results.

Cut scores refer to the score a student must get on an assessment to achieve a certain rating (somewhat like grades in a class). Performance levels are the number of overall ratings (“grades”) or categories of performance based on the results from the assessment. State high school assessments typically have more than one performance level (e.g., students’ overall scores could put them in the partially proficient, proficient, or advanced range).

Cut scores are particularly important if a state adopts a college entrance exam as its high school assessment. Both ACT and SAT have college readiness benchmarks – that is, the score where a student is deemed ready to succeed in college³. A state’s elementary and middle school assessments may have cut scores that are not in sync with the ACT or SAT college readiness benchmarks. This can be an issue for individual students as well as the entire assessment system. For example, a student (and her parents) could believe she is on track to be college and career ready based on her middle school assessment results and then find out she is not college ready when she takes the college entrance exam that serves as the high school assessment. Instead of moving seamlessly into college, she now faces the prospect of taking remedial classes for no credit. At the system level, student performance could look like it is improving or declining dramatically from middle school to high school (depending on where the elementary and middle school cut scores are set) when the change is really due to a lack of alignment of the cut scores across levels. This can cause stakeholders to lose faith in the assessment system itself since the scores do not seem to make sense from one level to the next. On the other hand, a state could set a lower cut score for “mastery” on the college entrance exam than the benchmarks on the ACT or SAT. However, this may cause stakeholders to question why the state set a cut score on the high school assessment that is below the cut score for college readiness when the state’s goal for all students is college and career readiness.

The number of performance levels is also important. There is only one national cut score on the ACT and SAT indicating the likelihood of college success.⁴ However, ESSA requires that high school tests separate student performance into at least three levels, and states have typically established four or five levels of performance (e.g., PARCC has five performance levels: Did not meet expectations, Partially met expectations, Approached expectations, Met expectations, Exceeded expectations). States that use one of the college entrance exams as their high school assessment will need to establish the additional cut scores needed either independently or with the cooperation of the ACT and the College Board.

Action Steps

1. Include a review of the alignment of elementary and middle school cut scores with high school cut scores as part of the high school assessment selection process.
2. Develop communication materials and strategies that help policymakers and the public understand how any differences in where cut points are set across levels impacts overall results.
3. Be prepared to communicate the rationale for any differences between cut scores on the high school assessment and college readiness scores on college entrance exams.
4. Be prepared to establish additional cut scores if your state adopts the ACT or SAT as the high school assessment.

State Oversight and Control

Another important issue is the degree to which states have control over the high school assessments. That is, who controls the:

- Design of the assessments
- Administration procedures
- Student data that is collected as part of the assessment process

State-developed assessments potentially offer the most control. For example, states can ensure participation by a broad range of stakeholders, including K-12 educators, higher education personnel, and content experts from the state in the development and review of test items. They can work directly with the vendor to limit testing

³ Defined as scoring a B or better in the first credit-bearing course in a relevant course (ACT), or as achieving at least a C in a set of first semester, credit-bearing college courses (SAT)

⁴ ACT’s College Ready Benchmarks are set by subject area.

time. Educators and parents can be involved in identifying appropriate accommodations for special populations. And, states that develop their own assessments can retain control over the student data - how it is used, where it is stored, how long it is stored, and who has access to that information. However, developing a specific state assessment takes at least two years from releasing an RFP to implementation (if not longer) and is an expensive process. Many states joined one of the national assessment consortia to avoid the cost of developing their own state assessment.

Consortia-developed assessments offer a lesser degree of control over the design and development process, but still have some opportunity for state input and participation by local stakeholders. Since states that utilize one of the consortium assessments typically contract with a vendor to administer and score the assessment, they can retain a high degree of control over the use and storage of student data. An important recent development is the ability of states to customize one of the consortia assessments to meet their specific needs. This will offer much more state control over the assessment content and administration procedures.

College entrance exams and other national tests offer less state control. The item development process, administration procedures, and the set of available accommodations are controlled by the testing company (in order to have nation-wide comparability). Both ACT and the College Board have data security policies and procedures to maintain the confidentiality of students, schools, and districts. States need to be aware of the policies and procedures and evaluate the degree to which such policies are aligned with the state's own policies.

Action Steps

1. With stakeholder involvement, determine the desired degree of state control over test development, administration procedures, and data utilization and privacy.
2. As part of the assessment selection process, be clear about what is under the state's control and what is in the purview of the vendor.
3. Have or adopt data security policies and procedures.

The Bottom Line

Today's high school assessment environment is volatile and messy. Some high school students and their parents have pushed back against what they see as irrelevant and meaningless state tests by "opting out" of the assessment. Additionally, until recently, very few in postsecondary education paid much attention to the results from state tests because they were not sufficiently rigorous and did not measure the knowledge, skills, and higher order thinking thought necessary for postsecondary success. Further, college admissions officers are cautious in adopting new measures and have been slow to use state tests because these tests have not been comparable across states, change frequently, and lack evidence that they can offer usable information for admissions decisions. But, as states have been implementing college- and career-ready standards for all students, they have been moving away from existing state tests to select or develop state assessments that reflect college and career readiness. Some states have also been using the ACT or SAT in their state testing and accountability systems for a number of years in an effort to raise expectations for all students in advance of the nearly universal move to college- and career-ready standards and assessments now in place across states.

It is important to understand that college entrance exams and measures of college and career readiness such as PARCC and Smarter Balanced and some state tests overlap – that is, they measure some of the same things. However, ACT and SAT were specifically designed to predict success in college and were not designed to measure the full extent of knowledge and skills that defines college and career readiness. The consortia-developed assessments and some state assessments were designed to assess the much broader range of knowledge and skills contained in the CCSS or standards adopted by most states that are key to success in both higher education and career training. And, unlike college entrance exams, these tests were intended to provide students with an

indication of progress toward becoming college and career ready as well as giving teachers a view of what college and career readiness “looks like” and the kind of instruction that is required to have all students reach that goal.

Within this complex context, states will have to decide within the next year whether to continue using their current high school assessments or replace or supplement them. States basically have four options:

1. Develop their own assessments that measure state standards
2. Utilize one of the consortia-developed assessments designed to measure the CCSS and college and career readiness
3. Select a college entrance exam that is designed to predict success in college
4. Employ some combination of these options⁵

Because ESSA empowers states to allow districts to use “nationally-recognized high school academic assessments” instead of the state high school assessment, districts may advocate for an additional option. That is, they may pressure states to let them choose from a menu of approved high school assessments (e.g., ACT or SAT or career and technical assessments that indicate readiness for specific careers). This is problematic if the results from these assessments will be used for accountability. Allowing districts to use different high school assessments for accountability purposes creates a lack of comparability and makes the results difficult to interpret. It may also advantage some schools/districts over others; one test may be “easier” or less rigorous (or at least be perceived to be) than another on the menu. A menu of assessments could be helpful if the goal is to provide individual students with a variety of pathways to obtain different college and career readiness credentials as long as the results from these different assessments are not aggregated as a measure of student academic performance.

There are tradeoffs in choosing any of the four primary options listed above in terms of what is being measured, the degree of state control, credibility, and cost (in terms of money and time). Individual state tests provide the most state control (including over testing time), but developing a state specific test is an expensive and time consuming process. Additionally, state developed assessments typically have much less credibility with higher education and high school educators and students than national college entrance exams or the consortia-developed assessments. However, some states have established agreements with their state institutions of higher education to use scores on the state test as admission requirements (e.g., California) and others are moving in that direction.

Consortia-developed assessments measure student progress towards mastery of the CCSS and college and career readiness and provide a comparison with student performance in other consortium states. In addition, some states accept the scores for admission to state universities and colleges, although they do not have as much credibility with college-bound high school students, their parents, and higher education as college entrance exams. Consortia-developed assessments are becoming more customizable for states, giving them more control over the development and administration of the assessment (which may limit comparability with other states). The consortia-developed assessments come with some political baggage that may make them unacceptable to some stakeholders.

College entrance exams have a lot of “name recognition” for educators and students, and are widely used for college admissions and placement. However, these tests do not provide complete information about where students are in meeting state standards because they were not built to measure those standards. In many cases, states will need to supplement these assessments in order to accurately measure their state standards. States have little or no control over the development of these assessments or their administrative procedures. In addition, they provide only one performance level which may not align with the performance levels defined at the elementary and middle levels. It will take some effort in order to add performance levels and evaluate the alignment of these cut scores across levels.

There is no perfect solution to the high school assessment question. Policymakers need to understand the issues and the tradeoffs involved in making this decision. In the end, the most important thing state policymakers can do is to help clarify the purposes for high school assessment and ensure that their state high school assessments serve those goals as part of a coherent assessment system that moves all students toward college and career readiness and provides meaningful information to students and educators on students’ progress towards that goal.

⁵ Some states are considering a “hybrid” model in which some consortium-developed items are incorporated into their existing state assessments or are used by the vendor in the development a new state test.

Acknowledgments

Achieve and the National Center for the Improvement of Educational Assessment (NCIEA) would like to thank the individuals and organizations that contributed to this brief. We thank Jacob Mishook, Associate Director, Assessment and Accountability at Achieve and Elliott Asp, Senior Fellow, State Policy Implementation and Support at Achieve for their leadership in writing this brief. Special thanks to Thanos Patelis and Joseph Martineau at NCIEA for developing the initial draft and for their careful review of, and thoughtful feedback, on the paper. Marie O'Hara, Associate Director, State Policy and Implementation Support at Achieve also provided helpful feedback.

Additional thanks to Kelly Van Beveren, Senior Communications Associate, Strategic Communications and Outreach for her editorial review and to Tom Flavell, Senior Associate, Online Communications, Strategic Communications and Outreach for his layout and design work.

Finally, we thank the Bill & Melinda Gates Foundation for their generous funding for this brief.

-Michael Cohen

President, Achieve

-Scott Marion

Executive Director, Center for Assessment

CC-BY 4.0 Achieve 2016. This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

