



#1 IN THE SERIES

Demystifying Statewide Standardized Assessments

*What Statewide Assessments Are
Designed to Measure*

By Michelle Croft, Hailly T.N. Korman,
and Titilayo Tinubu Ali

APRIL 2023

Overview

While K-12 students take many kinds of assessments (also called tests) for different purposes, the statewide standardized assessments used as part of annual federal accountability for schools often receive the greatest public scrutiny.

Statewide assessments measure what a student knows and can do. They are based on what schools are expected to teach for that grade level or content area within each state — and they can play a valuable role in improving education. Through comparable and consistent data, they allow decision-makers and educators to better understand how the education system is serving students, particularly historically marginalized students, including students of color, English learners, and students with disabilities. Test scores help educators identify students' strengths and areas of needed support to guide changes in instruction, inform large-scale instructional decisions (for example, identifying gaps in a district's curriculum), and measure how much a student has learned in a full academic year. These scores can also provide information about the effectiveness of instructional programs and other student supports to help state and district leaders and policymakers direct resources to schools and student populations.

Due to the complexity of assessment development and uncertainty around how scores are reported and can be used, statewide assessments can seem mysterious to some policymakers, educators, parents, students, and the general public. We developed six briefs to provide an overview of the test development process — from the initial stages of assessment design to the final process of scoring and reporting results — to help readers improve their understanding of how statewide assessments are developed and used. These briefs are organized by six topics:

1. What Statewide Assessments Are Designed to Measure

2. Ensuring That Assessments Accurately Measure Academic Standards
3. Developing High-Quality Assessments and Items
4. Ensuring Comparability Across Administrations
5. Making Assessments Accessible for Students With Disabilities and English Learners
6. Reporting Assessment Scores

WHAT ARE SUMMATIVE ASSESSMENTS?

Different tests have different purposes. State assessments are one type of summative assessment because they are administered to measure what a student has learned relative to what students should have been taught over the course of a school year. State summative assessments also can inform large-scale instructional decisions, such as identifying gaps in a district's curriculum, and measuring how much a student has learned in a full academic year.

While these briefs focus on statewide summative assessments, we also acknowledge the importance of other types of assessments that may be included in a state's assessment system, including formative and interim assessments.¹

Note: Throughout the six briefs, we use the terms "test" and "assessment" interchangeably.

What Statewide Assessments Are Designed to Measure

State standardized assessments measure students' skills and knowledge relative to what students should have been taught during the school year. The scores are used to identify schools in need of assistance and can help provide information for school and district leaders to set goals or allocate resources to support instruction.

Key Takeaways

- Through uniform assessment given to all students, statewide assessments measure what students know and can do relative to a state's standards.
- States must administer assessments annually in grades 3 to 8 and high school in math and English language arts and science in certain grade spans. States may also administer assessments in additional grade levels or content areas (e.g., social studies).
- The scores are disaggregated by different student groups (e.g., race/ethnicity, gender, disability status) and publicly reported to provide key stakeholders a picture of how the education system is serving all students. Data are used to identify school needs and provide instructional information.

In K-12 education, statewide assessments measure students' skills and knowledge of a set of state-approved academic content (called content standards) in specific subject areas (e.g., English language arts, math, and science). Statewide assessments are designed to be standardized, which refers to the consistency of the *types of content* a student sees on the assessment (e.g., whether algebra or geometry is included in a math assessment) and the testing *conditions* (e.g., whether assessments are timed or untimed, whether students can go backward and resubmit answers to items that were seen earlier in the assessment).² Having uniform opportunities for students to demonstrate their knowledge and skills for the same content under the same testing conditions provides consistent data that policymakers, school boards, families, and educators may use to evaluate how well school systems are meeting the needs of individual students as well as groups of students.

ACCOMMODATIONS

Each student taking the assessment should have equitable opportunities to demonstrate their knowledge and skills. Sometimes students' testing experiences need to differ slightly due to specific needs or unique circumstances. Accordingly, states work with stakeholders to identify and provide an approved set of accommodations that certain students may receive during testing.³ For example, a student with a broken arm may receive the assistance of an adult to input their answers into a test booklet or online testing system.

What types of content are assessed?

Statewide assessments are based on a state's academic content standards,⁴ which define the academic material that students should have been taught, skills they should have developed, and tasks they should be able to complete in each grade level, in the areas of reading,⁵ math, and science.⁶ States maintain published standards on official websites so that the most updated versions are accessible to the public.

What are the purposes of statewide assessments?

Statewide assessments are primarily designed to serve as part of the accountability system that helps determine how to properly allocate resources and provide a level of transparency to the public. Under the Every Student Succeeds Act, states are required to administer these assessments in grades 3 to 8 and in high school in math and English language arts as well as science in certain grade spans.⁷ With a few exceptions, states must administer the same assessment to all students within the same grade level⁸ and ensure that at least 95% of students participate.⁹ Having students across the state take the same assessment allows comparisons across schools and across districts. States must publicly report scores for schools and districts and break out the scores for different student populations such as by race/ethnicity, socioeconomic status, disability status, English proficiency, gender, and migrant status,¹⁰ giving policymakers detailed information to identify schools in need of support and improvement.¹¹ This helps ensure that resources can be appropriately directed toward students and schools.

In addition to accountability, the scores have instructional value because they provide information about what students know and can do relative to the state's content standards. At the individual level, educators can use the scores to determine if a student is meeting the state's grade-level expectations. Schools can use the scores to make decisions to improve student learning outcomes. Scores also can be used to look at groups of students to see areas of strengths or needed support. For example, if reading informational texts needs improvement, the school may decide to purchase supplementary curriculum in that area. ✨

Having uniform opportunities for students to demonstrate their knowledge and skills for the same content under the same testing conditions ... [evaluates] how well school systems are meeting the needs of individual students as well as groups of students.

Endnotes

- 1 Formative assessments are “a planned, ongoing process” that provide evidence of student learning to improve student outcomes. Formative assessments help educators design activities and instructional material to better align with student needs during learning. Interim assessments are another form of assessment that are administered periodically during the year and — depending on the assessment — can serve a formative function (i.e., for learning) or a summative function (i.e., to measure how much a student has learned). Formative Assessment for Students and Teachers (FAST) State Collaborative on Assessment and Student Standards, *Revising the Definition of Formative Assessment*, Council of Chief State School Officers, 2018, <https://ccsso.org/sites/default/files/2018-06/Revising%20the%20Definition%20of%20Formative%20Assessment.pdf>.
- 2 See “Ensuring Comparability Across Administrations” brief for more detail about standardization.
- 3 See “Making Assessments Accessible for Students With Disabilities and English Learners” brief for more detail about accommodations.
- 4 State standards are designed to require uniform expectations for learning. The use of the word “standard” for both state standards and standardized tests is coincidence but also a reflection of the shared goal of uniformity.
- 5 Some states refer to their standards as English language arts/literacy to include writing and listening.
- 6 20 U.S.C. § 6311(b), <https://www.govinfo.gov/content/pkg/USCODE-2010-title20/pdf/USCODE-2010-title20-chap70-subchapl-partA-subpart1-sec6311.pdf>.
- 7 Ibid.
- 8 Three exceptions in ESSA allow different assessments to be administered. A small percentage of students with significant cognitive disabilities participate by taking specially designed alternate assessments (ESSA § 1111(b)(2)(E)). For students who take algebra in the eighth grade, the state may allow the student to take the high school end-of-course algebra assessment in eighth grade and then administer a more difficult assessment when the student is in high school (ESSA § 1111(b)(2)(C)). The other exception is if the state approves a district to use a locally selected, nationally recognized high school assessment instead of the state assessment (ESSA § 1111(b)(2)(H)). If a district opts to use the nationally recognized assessment, all schools within the district would administer the nationally recognized assessment. The state must also provide evidence of comparability to the state assessment so that scores can be compared across the state.
- 9 ESSA 20 U.S.C. § 6311(c)(4)(E), <https://www.law.cornell.edu/uscode/text/20/6311>.
- 10 Ibid.
- 11 20 U.S.C. § 6311(b)(2), <https://www.law.cornell.edu/uscode/text/20/6311>. States may also have their own accountability system incorporating assessment scores in addition to the state’s federal accountability system.

About the Authors



MICHELLE CROFT

Michelle Croft is a senior analyst at Bellwether in the Policy and Evaluation practice area. She can be reached at michelle.croft@bellwether.org.



HAILLY T.N. KORMAN

Hailly T.N. Korman is a senior associate partner at Bellwether in the Policy and Evaluation practice area. She can be reached at hailly.korman@bellwether.org.



TITILAYO TINUBU ALI

Titilayo Tinubu Ali is a partner at Bellwether in the Policy and Evaluation practice area. She can be reached at titilayo.ali@bellwether.org.

About Bellwether

Bellwether is a national nonprofit that exists to transform education to ensure systemically marginalized young people achieve outcomes that lead to fulfilling lives and flourishing communities. Founded in 2010, we work hand in hand with education leaders and organizations to accelerate their impact, inform and influence policy and program design, and share what we learn along the way. For more, visit bellwether.org.

ACKNOWLEDGMENTS

We would like to thank the many individuals who gave their time and shared their knowledge with us to inform our work, including the Consortium on Assessment Policy and Advocacy (CAPA) for their support of this project.

We would also like to thank our Bellwether colleagues Andrew J. Rotherham, Harold Hinds, and Alexis Richardson. Thank you to Alyssa Schwenk, Abby Marco, Andy Jacob, Zoe Campbell, Julie Nguyen, and Amber Walker for shepherding and disseminating this work, and to Super Copy Editors.

The contributions of these individuals and entities significantly enhanced our work; however, any errors in fact or analysis remain the responsibility of the authors.



© 2023 Bellwether

- Ⓒ This report carries a Creative Commons license, which permits noncommercial re-use of content when proper attribution is provided. This means you are free to copy, display and distribute this work, or include content from this report in derivative works, under the following conditions:
- ① **Attribution.** You must clearly attribute the work to Bellwether and provide a link back to the publication at www.bellwether.org.
- Ⓓ **Noncommercial.** You may not use this work for commercial purposes without explicit prior permission from Bellwether.
- Ⓒ **Share Alike.** If you alter, transform, or build upon this work, you may distribute the resulting work only under a license identical to this one.

For the full legal code of this Creative Commons license, please visit www.creativecommons.org. If you have any questions about citing or reusing Bellwether content, please contact us.