Delaware System of Student Assessments (DeSSA)

Q&A on Initial Achievement Levels

Q. What is achievement level setting?

Achievement level setting, also known as standard setting, is the process for establishing one or more threshold scores on an assessment, making it possible to create categories of performance. Smarter Balanced Governing States approved a three-phase design for achievement level setting:

- An Online Panel (scheduled for October 6–17) allowed thousands of K–12 educators, higher education faculty, parents, and other interested parties to participate virtually in recommending achievement levels.
- An In-Person Panel (October 13–19) with educators and other stakeholders working in grade-level teams deliberated and made recommendations for the thresholds of the four achievement levels.
- The Cross-Grade Review Committee (October 20), a subset of the In-Person Panel, examined recommendations across all grades and considered the reasonableness of the system of cut scores.

Q. Who participated in achievement level setting process?

All Smarter Balanced states were actively involved in setting the achievement levels. Teachers, parents, higher education faculty, business leaders, and other community members from all of the Smarter Balanced states took part in a highly inclusive, consensus-based process that asked participants to closely examine assessment content and determine threshold scores for each achievement level. Educators who work with English language learners and students with disabilities were included to help ensure the fairness and appropriateness of the achievement levels for all students. Delaware was actively involved in this process with 142 participants serving on the Online Panel, 25 participants on the In-Person Panel, and 1 participant on the Cross-Grade Review Committee.

Q. How will these achievement levels be used?

The achievement levels serve as a starting point for discussion about the performance of individual students and groups of students in English language arts (ELA) and mathematics. Scale scores, growth models, and portfolios of student work can also be used to help evaluate the academic progress of students and schools.

Q. When will these achievement levels go into effect?

Once adopted by the Delaware State Board of Education, these achievement levels will appear on score reports for assessments administered this school year.

Q. We hear that our state's results are going to drop dramatically. Does this mean my child or my school is failing or that our state is doing worse than before?

The new tests are designed to assess college and career readiness using the new content standards as a benchmark. Expectations for student achievement with these new standards are higher than they used to be. Based on projections from the Smarter Balanced field test conducted in 2014, it is likely that fewer students will score at the higher achievement levels on the assessments, especially for the first few years. Results should improve as students have more years of instruction based on the new standards.

It is important to keep in mind that the tests have changed and are measuring different things—such as whether students are developing critical-thinking and problem-solving skills. Because we have raised expectations, fewer students are likely to meet those expectations initially.

Lower test scores do not necessarily mean that schools are performing worse or that students are learning less. Rather, it means we now have a more accurate measure of where students are on the path to success, based on the higher standards we set to ensure that our students are challenged and prepared to compete nationally and globally. Think of it as a fresh start, a new baseline.

As students have more years of instruction aligned to new standards, results typically improve. We have seen this pattern in several states. For example, in Kentucky, the first state to begin using the Common Core standards, student test scores went down at first. But, over the next four years as teachers and students worked to meet higher standards, the percentage of high school graduates meeting the state's benchmark for college and career readiness increased from 34% to 62%. There also have been impressive gains in ACT sores in Tennessee since the adoption of more rigorous standards. In California, where students have taken an early assessment of college readiness and participated in 12th grade courses to improve their preparation since 2007, the proportion of students needing remediation at the California State University has dropped from 56% to 43%.

Q. Why are these tests measuring college preparedness when some students may say they are not planning to go to college?

To be ready for success after high school, all students need to master skills such as critical thinking, analytical writing, and problem solving. Smarter Balanced assessments have been specifically developed to measure these real-world skills that students will need when they graduate, whether they are headed into a traditional degree program or postsecondary career training.

Q. What does this mean for teacher evaluations and school/district accountability?

Each state determines individually how to use student assessment results. For school and district accountability, we will use the Smarter assessment as part of the accountability measure. For teacher evaluation, Delaware has determined that the Smarter assessments will not be used for the 2014–2015 school year. A stakeholder group is currently working with the Department of Education to determine how the Smarter assessments will be used for school and district accountability.

For teacher evaluation, Delaware will use Smarter assessments in calculations for teacher evaluation for the 2014–2015 school year "for information only" (i.e., not counting toward an educator's evaluation), and plan to fully incorporate a growth model using Smarter assessments

in 2015–2016. The Department is working to align teacher evaluation with the school accountability measures (see above). A communications and engagement process will be rolled out in early 2015. The teacher evaluation model will continue to be based on student growth and not proficiency levels.

Q. Why should we trust computer-based tests?

The Smarter assessments will provide a more accurate measure of a student's knowledge and skills than tests that just asked students to pick the right answer from a multiple-choice list. The Smarter assessment system capitalizes on the ability of well-established, computer-adaptive technology. This approach individually tailors the test for every student so that each test question improves the accuracy of the student's scores across the full range of the achievement continuum. Delaware students have also been using computers for statewide testing for four years now, and the students are familiar with this kind of testing.

Q. What is this test going to tell me that I don't already learn from others?

A lot. Smarter Assessments offer significant improvements over tests of the past, including essay writing at every tested grade level—grades 3–8 and 11, and new performance tasks that ask students to demonstrate an array of research, critical-thinking, and problem-solving skills in response to real-world problems. We are excited to have these state-of-the-art tools to help us do a better job of both measuring student achievement and identifying where students need help.

Q. How do the Smarter Balanced achievement levels compare to NAEP?

Smarter Balanced projections for student achievement closely align with how students have performed historically on NAEP. The achievement levels also generally align with the results of a comprehensive research study on college readiness conducted by the National Assessment Governing Board, the oversight body for the National Assessment of Educational Progress (see https://www.nagb.org/what-we-do/preparedness-research.html for more information).

Q. What are our state projections based on the field test data?

Valid state-level projections are not available. The Smarter Balanced field test was designed to measure the difficulty and quality of the assessment items and to project outcomes for all students across the entire Consortium. Students were not sampled to be representative of any single state, but instead to represent the demographic characteristics of students across the entire Consortium. Therefore, projections of student results from the field test are only valid for the Consortium as a whole and cannot be interpreted on a state-by-state basis. Publishing individual state projections from the field test would create incorrect conclusions and would constitute an irresponsible use of these data.