Connecticut
Smarter Balanced Assessments
for
English Language Arts/Literacy
and
Mathematics

CSDE

CONNECTICUT STATE
DEPARTMENT OF EDUCATION

2017

REVISED July, 2017

Interpretive Guide

Connecticut State Board of Education

Contents

Purpose of the Interpretive Guide 1
General Principles of Test Interpretation and Use 1
Accessing Online Assessment Results 2
The Assessment Development Process 3
Overview 3
Key Components of the Assessment Development Process 4
The Smarter Balanced Assessment System 6
Overview 6
The Summative Assessments 6
Smarter Balanced Mathematics Claims 8
Smarter Balanced English Language Arts/Literacy Claims 8
Connecticut's English Language Arts/Literacy Reporting Categories 9
The Scores 9
Comparing ELA/Literacy Scores from 2014-15 to 2015-16 and 2016-17 10
Achievement Levels 11
Overview 11
Smarter Balanced Achievement Levels 11
Content-Specific Claim Scores 16
Mathematics 16
English Language Arts/Literacy 16
Aggregate Target-Level Results 17
Interpreting Scores in the Early Years of Implementation 18
Individual Student Reports 19
References 24
Smarter Balanced Scoring and Reporting FAQ – Appendix A 25

The Connecticut State Department of Education is committed to a policy of equal opportunity/affirmative action for all qualified persons. The Connecticut State Department of Education does not discriminate in any employment practice, education program, or educational activity on the basis of **race, color, religious creed, sex, age, national origin, ancestry, marital status, sexual orientation, gender identity or expression, disability (including, but not limited to, intellectual disability, past or present history of mental disorder, physical disability or learning disability), genetic information, or any other basis prohibited by Connecticut state and/or federal nondiscrimination laws. The Connecticut State Department of Education does not unlawfully discriminate in employment and licensing against qualified persons with a prior criminal conviction.** Inquiries regarding the Connecticut State Department of Education's nondiscrimination policies should be directed to: Levy Gillespie, Equal Employment Opportunity Director/Americans with Disabilities Act Coordinator, Connecticut State Department of Education, 450 Columbus Boulevard, Hartford, CT 06103 | 860-807-2071 | Levy.Gillespie@ct.gov

Purpose of the Interpretive Guide

The *Connecticut Smarter Balanced Assessments Interpretive Guide* is designed to help educators, parents, students, and members of both the public and the media understand and properly explain the results of the Smarter Balanced Summative Assessments. This guide provides interpretation rules to consider when analyzing Smarter Balanced assessment data to ensure proper interpretation and use of these data to inform decisions around both classroom instruction, and professional development.

The following general principles section was excerpted from L. Hammond, et al., (2015). [1]

General Principles of Test Interpretation and Use

Educational assessments can offer valuable information to students, parents, educators, and policymakers regarding what students know and are able to do. When used appropriately, they can provide an objective and efficient way to gauge some aspects of student learning and achievement and can inform the decision-making process about future instruction. All assessments have limitations; for example, a single assessment cannot measure all the aspects of an individual's knowledge, skills, and abilities, and no assessment can measure learning perfectly. The following general principles of test-score interpretation and use are generally accepted by measurement experts and are articulated in the newly revised *Standards for Educational and Psychological Testing.*

Tests are imprecise: Even a well-designed assessment may contain measurement error (AERA, APA, & NCME, 2014; NRC, 2007), which is the degree of imprecision or uncertainty in the assessment procedure. Measurement error occurs due to factors unrelated to student learning. For example, student performance on an assessment may be affected by mood, health, testing conditions, and motivation, as well as potential variability related to human scoring. Furthermore, the questions on a given test are only a sample of all the knowledge and skills that pertain to the subject being tested. If a different sample of questions had been chosen, or the questions had been posed in a different form, the student could have scored differently. Therefore, a test score is not an exact measure of a student's competencies since measurement error is inherent in all tests.

Tests provide only partial evidence about performance; thus, they should be combined with other sources of evidence for decision-making: In drawing any conclusion or making any decision, test scores should always be used in conjunction with multiple sources of evidence about performance (AERA, APA, & NCME, 2014; NRC, 2007). Consequential decisions about a student, educator, or a school should not be made only or primarily based on a single test score. Because a test score is not perfect and only tells part of the story, other relevant information (e.g., student work samples, course grades, course-taking record, teacher observations, and other measures) should be included to place test scores in context and allow for a broader view of performance.

The extent and nature of evidence needed may depend on characteristics of the learner (e.g., age, prior schooling, native language, learning differences), as well as the interpretation to be made (e.g., next steps for instruction, program placement, readiness for a specific experience, etc.). A range of appropriate measures about an individual's competencies will enhance the validity of the overall interpretation of the test score and the appropriateness of decisions that rely in part on test data.

The more consequential the test use, the stronger the evidence must be to support that

[1] L. Hammond, E. Haertel, J. Pellegrino. (2015). *Making Good Use of New Assessment: Interpreting and Using Scores from the Smarter Balanced Assessment Consortium.*

use (AERA, APA, & NCME, 2014; NRC, 2007). High stakes demand that a stronger body of additional supporting evidence is provided in order to "minimize errors of measurement or errors in classifying individuals into categories such as 'pass,' 'fail,' 'admit,' or 'reject'" (AERA, APA, & NCME, 2014, p. 188). When multiple sources of evidence agree, we can have greater confidence that the inferences we base on test scores are sound.

Validity depends on test design and use: An assessment is valid only when used with the intended population of test takers for the specific purposes and under the conditions (including prior preparation, motivation, and other administration conditions) for which it was designed and validated (AERA, APA, & NCME, 2014; NRC, 2007). Test validity refers to the extent to which inferences about individuals based on their scores on a particular test are defensible. When used as designed, test data can provide useful information. However, any test may function poorly or have unintended consequences if used outside the specific purposes and populations for which it was designed and validated.

Test score interpretations or judgments are validated for specific purposes and validity does not automatically transfer to new uses: each different purpose must be justified and validated in its own right. No assessment is valid for all possible purposes.

Opportunities to learn influence valid inferences as well as fairness: In educational contexts, valid inferences about student ability derived from tests depend on students having been provided opportunities to learn the tested material prior to the assessment being administered. The degree to which students are afforded high-quality instruction, and are supported to perform to their full potential, affects the degree to which test scores can appropriately support consequential decisions about their knowledge, skills, and abilities (NRC, 2007).

Accessing Online Assessment Results

Smarter Balanced non-confidential aggregate results are publicly reported through EdSight (http://edsight.ct.gov), an interactive Web site that integrates important school and district information collected by the Connecticut State Department of Education (CSDE) that serves as a single source for all data-driven analyses and reporting. Information can be sorted, filtered, and compared across schools and districts, and across race/ethnicity, and gender.

Smarter Balanced results for individual students are password protected and available to authorized school district personnel in the Score Reports feature of the Online Reporting System (ORS) located on the Connecticut Smarter Balanced Assessment Portal (CT.portal.airast.org).

The ORS is a web-based system that provides school district users access to individual student performance results. Users can compare score data between individual students and the school, or district. The ORS also provides information in the aggregate about performance on Smarter Balanced Assessment Claims and Targets. These data can be disaggregated by gender, special education, and English learner (EL) status. The ORS User Guide describes features of ORS, including an overview of the available score reports, and is available on the Connecticut Smarter Balanced Assessment Portal.

Additional information about the Smarter Balanced Assessment System is available through the Student Assessment link on the Connecticut State Department of Education Web site (www.ct.gov/sde) and on the Smarter Balanced Web site (www.smarterbalanced.org).

General questions about the Smarter Balanced Assessments should be directed to the Student Assessment Office at 860-713-6860 or ctstudentassessment@ct.gov.

Specific questions about individual student results should be directed to local school personnel.

The Assessment Development Process

Overview

In 2009, the Council of Chief State School Officers and the National Governors Association Center for Best Practices coordinated a state-led effort to develop the Common Core State Standards referred to in Connecticut as the Connecticut Core Standards (CCS). The goal of the collaboration was to establish clear and consistent education standards in mathematics and English language arts/literacy that would help prepare all students for success in college and careers. The CCS define what students should learn as described in learning progressions and grade-level expectations. The standards were adopted in Connecticut on July 7, 2010.

The adoption and implementation of the CCS required the development of next-generation assessments. The Smarter Balanced Assessment Consortium was one of two state-led consortia that developed systems of assessments aligned to the CCS under the Race To The Top (RTTT) Fund Assessment Program: Comprehensive Assessment Systems Grant.

In 2010, the consortium laid out its vision for an innovative assessment system intended to inform parents, students, teachers and policymakers about student achievement in relation to the CCS. The consortium's work is guided by the following principles:

1. Assessments are grounded in a thoughtful, standards-based curriculum, and are managed as part of an integrated system of standards, curriculum, assessment, instruction, and teacher development.
2. Assessments produce evidence of student performance on challenging tasks that evaluate CCS.
3. Teachers are integrally involved in the development of assessments and the scoring of the interim assessments.
4. The development and implementation of the assessment system is a state-led effort with a transparent and inclusive governance structure.
5. Assessments are structured to continuously improve teaching and learning.
6. Assessment, reporting, and accountability systems provide useful information on multiple measures that is instructive for all stakeholders.
7. Design and implementation strategies adhere to established professional standards.

Connecticut joined the Smarter Balanced Assessment Consortium as a governing state in June 2010. In January 2011, ten workgroups were established:

1. Item Development
2. Performance Tasks
3. Test Administration
4. Accessibility and Accommodations
5. Reporting
6. Technology Approach
7. Formative Assessment Practices and Professional Learning
8. Test Design
9. Test Administration
10. Validation and Psychometrics

The work groups were made up of 110 state-level staff, including CSDE assessment consultants, who were responsible for overseeing the work of the consortium in each area. Work group members participated in the vendor-selection process and provided ongoing feedback and guidance during the development of the assessment system. Thousands of K-12 educators and higher education faculty from across member states, including over 300 from Connecticut, also participated in various aspects of the assessment system development.

Since the end of the assessment grant in September 2014, Smarter Balanced has operated as a public agency. Smarter Balanced is housed at the University of California, Santa Cruz (UCSC) Silicon Valley Extension.

Key Components of the Assessment Development Process

Technical Advisory Committee (TAC): At the beginning of the grant period, the consortium assembled a TAC comprised of highly regarded national experts. The TAC met regularly over the grant period and continues to do so in order to provide technical advice and support on key decisions related to all components of the assessment system.

Evidence-Centered Design (ECD): As described in the Smarter Balanced Content Specifications for mathematics and English language arts (ELA), the consortium made a commitment to employ an ECD approach in the development of the assessment system. Central to ECD is the idea of collecting evidence through a student's response to an item or task that supports a claim about the extent to which a student has developed the knowledge, skill, and ability that is contained in a content standard or target of instruction.

Content Specifications Development: Initial drafts of the Smarter Balanced Content Specifications for mathematics and English language arts were completed during the summer of 2011. The consortium assembled a team of experts in the fields of mathematics, English language arts, and assessment, along with the lead authors of the CCS to write the content specifications. These documents established the assessment claims that are described below along with the evidence that the consortium would need to collect in order to support each claim by grade level. The documents specify assessment targets and lay out accessibility strategies for English learners and students with disabilities to be considered in addressing each target. Consortium staff, state work group members, and the consortium's TAC reviewed this initial draft. A revised version went through two rounds of public review during which more than 200 individuals and organizations provided feedback on the content specifications. Using the public's feedback, the documents were revised and then the governing states voted on the claims.

Achievement-Level Descriptors (ALDs) Development: In October 2012, 30 K-12 educators and 21 higher-education faculty members convened to write ALDs for ELA/literacy and mathematics. The K-12 educators were chosen to represent the diversity in schools across member states. Educators with Grades 3–8 experience made up the panels for these grades. Following the workshop, there were three rounds of review including Smarter Balanced staff, committees, and more than 350 members of the public representing K-12 and higher education. Collectively, these groups contributed to the wording of the final version. The Achievement Level Descriptors are located on the 2017 Communication Tools for Smarter Balanced page of the CSDE Web site.

Item and Task Development: The consortium developed item and task specifications to ensure that the assessment items and tasks measure the assessments' claims. The specifications delineate the types of evidence that should be elicited for each claim within a grade level. They also provide explicit guidance on how to write items in order to elicit the desired evidence. The consortium developed many different types of items beyond traditional multiple-choice items. This was done to measure the claims and assessment targets with varying degrees of complexity by allowing students to construct their responses rather than simply recognizing a correct response. All items were created using principles of universal design, which aim to create items that accurately measure the assessment target for a wide range of student abilities. Item writers have been trained to consider all students who may answer a question – including students from various demographic regions or socioeconomic status, students with disabilities, and English

learners – to ensure that the context of the item is familiar to the majority of students in a particular grade level. The various item types are illustrated on the Item Type Tutorials page of the CSDE Smarter Balanced Assessment Portal.

The Smarter Balanced Content Specifications and Item and Task Specifications are available on the Smarter Balanced Web site.

Small-Scale Trials, Pilot Testing, and Field Testing: A small set of items was developed and administered in the fall of 2012 during a small-scale trial. New item types were tested prior to large-scale development for later field testing. During the small-scale trials, the consortium conducted cognitive laboratories to better understand how students solve various types of items. A cognitive laboratory uses a think-aloud methodology in which students verbalize their thinking while answering a test question. The Item and Task Specifications were again revised based on the findings of the small-scale trials. These specifications were used to develop items for the 2013 pilot test and were again revised based on the pilot test results.

A large-scale field test was administered to approximately 4.2 million students in over 16,500 schools across the 21 governing states and the U.S. Virgin Islands in spring 2014. The field test was a practice run of the assessment system that helped ensure the accuracy and fairness of the test questions. It also gave teachers and schools a chance to gauge their readiness in advance of the first operational assessment in the spring of 2015. This field test allowed the consortium to evaluate the performance of the more than 19,000 items and performance tasks in the item pool. Field-test data were used to identify those items that performed well and those items that needed to be improved or even rejected for use on an operational assessment. This information was also used to inform future item-writing efforts. Both before and after the field test, panels of educators reviewed all items, performance tasks, and item stimuli for accessibility, bias/sensitivity, and content.

Accessibility Features: To provide every student with a positive and productive assessment experience, and to generate results that are a fair and accurate estimate of each student's achievement, member states worked together to create an accessibility framework that includes universal tools, designated supports, and accommodations. These tools and supports all yield reportable scores when used in the manner specified by the Smarter Balanced Usability, Accessibility, and Accommodations Guidelines.

Achievement-Level Setting: In November 2014, the consortium involved thousands of stakeholders in setting achievement levels, using a process known as the "bookmark method." Approximately 500 teachers, school leaders, higher education faculty, parents, business and community leaders met in person to review test questions and determine the threshold scores (i.e., cut scores) for four achievement levels for each grade and content area. Representatives of each member state and educators with experience teaching English learners, students with disabilities, and other traditionally underrepresented students participated to help ensure that the achievement levels were fair and appropriate for all students. In addition, an online panel was open to educators, parents, and other interested members of the community to provide input on the achievement levels. More than 9,500 people registered to participate in the online panel. A cross-grade review committee, composed of 72 members of the in-person panels, then took the results of the online and in-person panels into account to develop recommendations that coherently aligned across grades and that reflect student progress from year to year.

The Smarter Balanced Assessment System

Overview

Smarter Balanced is an online assessment system aligned to the Connecticut Core Standards (CCS). The components of the system are designed to work together to help ensure that every student meets the overarching goal that all students leave high school prepared for postsecondary success in college or a career through increased learning and improved teaching. The assessment system is comprised of three components:

- **a summative assessment** administered near the end of the school year;
- **optional interim assessments** administered at locally determined intervals; and
- **an optional digital library** that houses professional development and professional learning materials, resources, and tools aligned to the CCS, with a strong focus on formative assessment processes.

The Summative Assessments

Connecticut General Statues (Section 10-14n) mandates that all public school students enrolled in Grades 3 through 8, and 10 or 11 participate in a "mastery examination" approved by the State Board of Education that measures essential and grade-appropriate skills in reading, writing, mathematics, or science.

> Connecticut General Statute 10-14n as amended by Public Act No. 15-238:
>
> (b) (1) For the school year commencing July 1, 2015, and each school year thereafter, each student enrolled in grades three to eight, inclusive, and grade eleven in any public school shall, annually, take a mastery examination in reading, writing and mathematics during the regular school day.
>
> (2) For the school year commencing July 1, 2013, and each school year thereafter, each student enrolled in grades five, eight and ten in any public school shall, annually, in March or April, take a state-wide mastery examination in science during the regular school day.

The Smarter Balanced Summative Assessment is Connecticut's mastery examination for Grades 3-8 in English language arts and mathematics. It is the culminating evaluation of student performance relative to the Connecticut Core Standards. It provides an efficient and reliable estimate of a student's overall performance in a subject area relative to grade-appropriate standards that enable valid interpretations of student achievement and progress.

In Connecticut, the Smarter Balanced ELA and mathematics summative assessments include a computer adaptive test; the mathematics test also includes a performance task.

***Computer Adaptive Test (CAT)*:** A computer adaptive test adjusts the test to each student by basing the difficulty of future questions on previous answers. This provides a more accurate measurement of student achievement. The CAT represents a significant improvement over traditional paper-and-pencil assessments, providing more accurate scores for all students across the full range of the achievement continuum.

Performance Tasks (PT): Performance tasks are designed to provide students with an opportunity to demonstrate their ability to apply their knowledge and higher-order thinking skills to explore and analyze a complex, real-world scenario. They can best be described as collections of questions and activities that are coherently connected to a single theme or scenario. These activities are meant to measure capacities such as depth of understanding or complex analysis with relevant evidence, which cannot be adequately assessed with traditional assessment questions. While the performance tasks are also administered on a computer, they are not computer adaptive.

The Smarter Balanced Content Specifications: The Smarter Balanced Content Specifications in English language arts/literacy and mathematics were developed to ensure that the assessments cover the range of knowledge and skills in the CCS. The Content Specifications serve as the basis for the development of the Smarter Balanced Summative and Interim Assessments. They describe clear and prioritized assessment claims and targets that are used to translate the grade-level standards into content frameworks from which test blueprints and item and task specifications are established.

The Smarter Balanced Assessment Claims: The assessments were developed using an evidence-based design that identifies claims—one overall composite claim associated with each content area assessment, and additional specific content claims. Assessment claims are broad evidence-based statements about what students know and can do as demonstrated by their performance on subsets of the assessment. Students will receive an overall scale score for each content area (i.e., ELA or Mathematics) and sub-scores for each content-specific claim. These scores are derived from clusters of items in both the CAT and PT in mathematics and the CAT for ELA.

The Smarter Balanced Assessment Targets: Each content-specific claim is accompanied by a set of assessment targets that provide more detail about the range of content and Depth of Knowledge levels. The targets were drawn from the CCS and are intended to support the development of high-quality items and tasks that contribute evidence to the claims.

For mathematics Claim 1, the targets are drawn from the cluster-level headings of the Standards for Mathematical Content. Use of more fine-grained descriptions would risk a tendency to atomize the content and might lead to assessments that would not meet the intent of the standards. For Claims 2, 3 and 4, the targets are drawn from the language in the Standards for Mathematical Practice. These targets are the same across all grade levels.

For ELA, the statements drawn from the Connecticut Core Standards for English Language Arts and Literacy in History/Social Studies, Science, and Technical Studies were reorganized or combined into targets, thus changing the presentation of the standards without changing the content. This was done to develop efficient strategies for assessment and reporting. Rather than tapping only isolated skills within one strand, such as reading, standards-based instruction requires students to integrate skills and concepts across strands; subsequently, Smarter Balanced ELA/literacy Assessment Claims and Targets represent the ways in which students may be expected to learn and demonstrate their knowledge of ELA. The ELA assessment targets are focused on a subset of skills and aligned to a variety of standards. The demands within the assessment targets vary by grade and demonstrate the progression of learning as students advance from grade to grade.

The Mathematics Assessment: The CCS for Mathematics require that mathematical content and mathematical practices are connected. Students are expected to make connections between content and practice, model a mathematical situation, and explain their reasoning when solving problems. In addition, two of the major design principles of the standards are *focus* and *coherence*. *Coherence* implies that the standards are more than a mere checklist of disconnected topics, while attending to *focus* will allow the student the time necessary to learn and master

grade-level content in order to be able to build upon it the following year. Together, these features of the standards had an important influence on the design of the Smarter Balanced mathematics assessment. The mathematics claims are described below:

Smarter Balanced Mathematics Claims

Overall Claim for Grades 3-8	Students can demonstrate progress toward college and career readiness in mathematics.
Claim 1 Concepts and Procedures	Students can explain and apply mathematical concepts and interpret and carry out mathematical procedures with precision and fluency.
***Claim 2 Problem Solving**	Students can solve a range of complex well-posed problems in pure and applied mathematics, making productive use of knowledge and problem-solving strategies.
Claim 3 Communicating Reasoning	Students can clearly and precisely construct viable arguments to support their own reasoning and to critique the reasoning of others.
***Claim 4 Modeling and Data Analysis**	Students can analyze complex, real-world scenarios and can construct and use mathematical models to interpret and solve problems.

*NOTE: For reporting purposes in CT, Claims 2 and 4 are combined into one reporting category.

The English Language Arts Assessment: Aligned to the CCS for English Language Arts, Smarter Balanced assessments measure the success of students as they progress towards college and career readiness in reading, writing, listening, and research. The CCS reinforce the importance of students being able to learn to read, write, speak, listen, and use language effectively in a variety of content areas, as well as to think critically. The ELA/literacy Claims are described below:

Smarter Balanced English Language Arts/Literacy Claims

Overall Claim for Grades 3-8	Students can demonstrate progress toward college and career readiness in English language arts and literacy.
Claim 1 Reading	Students can read closely and analytically to comprehend a range of increasingly complex literary and informational texts.
***Claim 2 Writing**	Students can produce effective writing for a range of purposes and audiences.
Claim 3 Listening	Students can employ effective speaking and listening skills for a range of purposes and audiences.
***Claim 4 Research**	Students can engage in research/inquiry to investigate topics, and to analyze, integrate, and present information.

*NOTE: For reporting purposes in CT, Claims 2 and 4 are combined into one reporting category.

Because the ELA Performance Task is no longer administered as part of the Connecticut Smarter Balanced Summative Assessment, Claims 2 and 4 can no longer be reported as standalone categories. A reporting category, unique to Connecticut, encompasses components of Writing (Claim 2) with Research (Claim 4). This reporting category, Writing and Research/Inquiry, cannot be compared to the Smarter Balanced Claim 2 and Claim 4 data from the spring 2015 administration. Only Reading (Claim 1) and Listening (Claim 3) are comparable from the 2014-15 to 2015-16 and 2016-17 school years.

Connecticut's English Language Arts/Literacy Reporting Categories

Reading	Students can read closely and analytically to comprehend a range of increasingly complex literary and informational texts.
Listening	Students can employ effective speaking and listening skills for a range of purposes and audiences.
Writing and Research/Inquiry	Students can strengthen writing by revising, editing, and rewriting a range of text. Students can also analyze and integrate evidence-based information to support analysis and research.

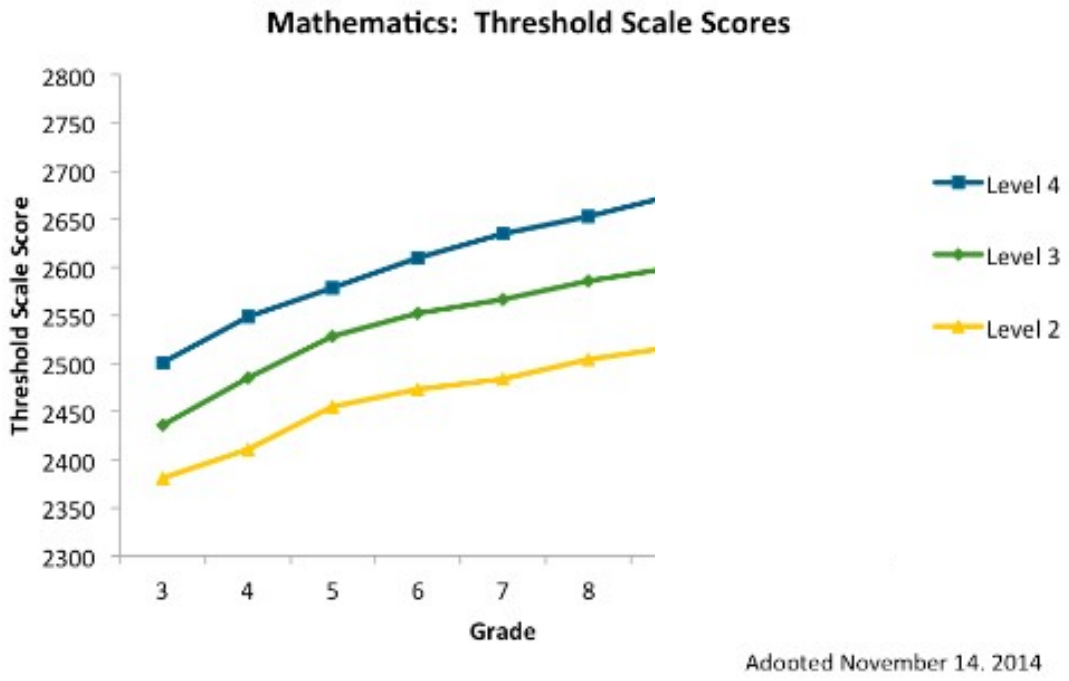
The Scores

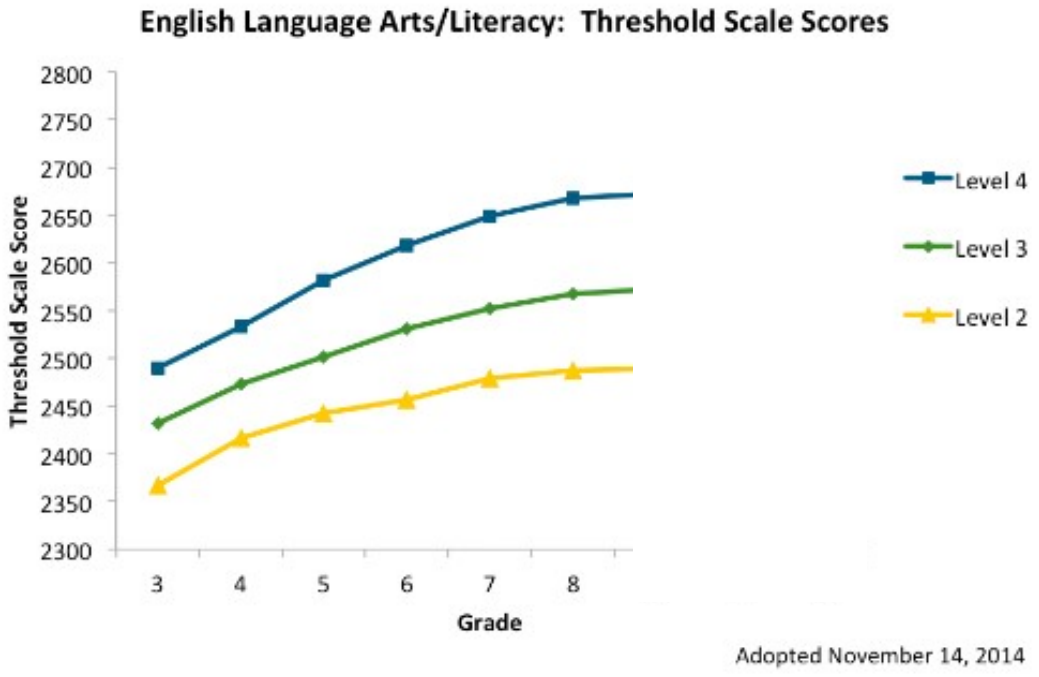
Each student who completes the Smarter Balanced Summative Assessment receives a total scale score and associated achievement level for each content area. Scale scores are the basic unit of reporting. A scale score is derived from a total number of obtained score points that is statistically adjusted and converted into a consistent, standardized scale that permits direct and fair comparisons of scores from different forms of a test, either within the same administration year or across years (Tan & Michel, 2011). Established psychometric procedures are used to ensure that a given scale score represents the same level of performance regardless of the test form. This allows for the fair comparison of scale scores from a computer adaptive test where different students are presented with different test questions. While scale scores are comparable across tests in a given content area, they are not comparable across content areas; a scale score on the mathematics test should not be compared to a scale score on the ELA test.

Each overall scale score is indicated by a single number. An error band is described on the Individual Student Report (ISR) for each scale score. The error band indicates the range of scores that the student would be likely to achieve if he or she were to take the test multiple times.

The Smarter Balanced overall scale scores fall along a continuous vertical scale (from approximately 2000 to 3000) that increases across grade levels. These scores can be used to illustrate students' current level of achievement and their growth over time. When aggregated, these scores can also describe school- or district-level changes in performance on the tests and can be used to measure gaps in achievement among different groups of students.

The mathematics and ELA threshold scores (scores that separate achievement levels from one another) are provided in the graphs below.





Comparing ELA/Literacy Scores from 2014-15 to 2015-16 and 2016-17

Connecticut discontinued the ELA Performance Task for 2015-16; however, the 2014-15 ELA results include the Performance Task. Therefore, to enable the most valid comparison of aggregate results from 2014-15 to 2015-16 and 2016-17, the CSDE scored the 2014-15 test

unofficially, based solely on the computer-adaptive test (CAT) portion of the ELA test. Aggregate district, school, and grade-level results are provided for district/schools to enable valid comparisons of the "CAT-only" ELA results from 2014-15 to 2015-16 and 2016-17. The discontinuance of the ELA Performance Task affected the Writing and Research/Inquiry claims, but did not affect the Reading and Listening claims; therefore, claim performances in Reading and Listening can be compared between 2014-15, 2015-16, and 2016-17.

Achievement Levels

Overview

The Consortium developed a set of initial, policy Achievement-Level Descriptors (ALDs) for ELA/literacy and mathematics that are aligned with the CCS and the Smarter Balanced Assessment Claims. The purpose of these descriptors is to specify, in content terms, the knowledge and skills that students display at four levels of achievement.

Defining these achievement levels is a reporting feature that has become familiar to many educators. However, characterizing a student's achievement solely in terms of falling in one of four categories is an oversimplification. Achievement levels should serve only as a starting point for discussion about the performance of students and of groups of students. That is, the achievement levels should never be interpreted as infallible predictors of a student's future. They must continuously be validated, and should only be used in the context of the multiple sources of information that we have about students and schools. Achievement-Level Descriptors do not equate directly to expectations for "on-grade" performance; rather, they represent differing levels of performance for students within a grade level. Additionally, the achievement levels do not preclude or replace other methods of evaluating assessment results, including measures of year-to-year growth that use the vertical scale scores.

Although the ALDs are intended to aid interpretation of achievement levels, they will be less precise than scale scores for describing student gains over time or changes in achievement gaps among groups, since they do not reveal changes of student scores within the bands defined by the achievement levels. Furthermore, there is not a critical shift in student knowledge or understanding that occurs at a single cut-score point. Thus, the achievement levels should be understood as representing approximations of levels at which students demonstrate mastery of a set of concepts and skills, and the scale scores just above and below an achievement level as within a general band of performance.

Smarter Balanced Achievement Levels

Content Area Mathematics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Level 4	2501-2621	2549-2659	2579-2700	2610-2748	2635-2778	2653-2802
Level 3	2436-2500	2485-2548	2528-2578	2552-2609	2567-2634	2586-2652
Level 2	2381-2435	2411-2484	2455-2527	2473-2551	2484-2566	2504-2585
Level 1	2189-2380	2204-2410	2219-2454	2235-2472	2250-2483	2265-2503
ELA/Literacy						
Level 4	2490-2623	2533-2663	2582-2701	2618-2724	2649-2745	2668-2769
Level 3	2432-2489	2473-2532	2502-2581	2531-2617	2552-2648	2567-2667
Level 2	2367-2431	2416-2472	2442-2501	2457-2530	2479-2551	2487-2566
Level 1	2114-2366	2131-2415	2201-2441	2210-2456	2258-2478	2288-2486

Achievement-Level Descriptors

An ALD is included for each content area on the Individual Student Report. The ALDs are intended to help parents and educators understand the general characteristics of students who score at a particular achievement level.

The Connecticut ALDs for the Grades 3-5 mathematics test are shown in the table below.

Achievement Level	Grades 3-5 Achievement-Level Descriptors for Mathematics
Level 4	***Exceeds the Achievement Standard:*** The student has ***exceeded the achievement standard*** for mathematics expected for this grade. Students performing at this standard are ***demonstrating advanced progress toward mastery*** of mathematics knowledge and skills. Students performing at this standard are on track for likely success in the next grade.
Level 3	***Meets the Achievement Standard:*** The student has ***met the achievement standard*** for mathematics expected for this grade. Students performing at this standard are ***demonstrating progress toward mastery*** of mathematics knowledge and skills. Students performing at this standard are on track for likely success in the next grade.
Level 2	***Approaching the Achievement Standard:*** The student has ***nearly met the achievement standard*** for mathematics expected for this grade. Students performing at this standard r***equire further development toward mastery*** of mathematics knowledge and skills. Students performing at this standard will likely need support to get on track for success in the next grade.
Level 1	***Does Not Meet the Achievement Standard:*** The student has ***not yet met the achievement standard*** for mathematics expected for this grade. Students performing at this standard ***require substantial improvement toward mastery*** of mathematics knowledge and skills. Students performing at this standard will likely need substantial support to get on track for success in the next grade.

The Connecticut ALDs for the Grades 6-8 mathematics test are shown in the table below.

Achievement Level	Grades 6-8 Achievement-Level Descriptors for Mathematics
Level 4	***Exceeds the Achievement Standard:*** The student has ***exceeded the achievement standard*** for mathematics expected for this grade. Students performing at this standard are ***demonstrating advanced progress toward mastery*** of mathematics knowledge and skills. Students performing at this standard are on track for likely success in high school and college coursework or career training.
Level 3	***Meets the Achievement Standard:*** The student has ***met the achievement standard*** for mathematics expected for this grade. Students performing at this standard are ***demonstrating progress toward mastery*** of mathematics knowledge and skills. Students performing at this standard are on track for likely success in high school and college coursework or career training.
Level 2	***Approaching the Achievement Standard:*** The student has ***nearly met the achievement standard*** for mathematics expected for this grade. Students performing at this standard ***require further development toward mastery*** of mathematics knowledge and skills. Students performing at this standard will likely need support to get on track for success in high school and college coursework or career training.
Level 1	***Does Not Meet the Achievement Standard:*** The student has ***not yet met the achievement standard*** for mathematics expected for this grade. Students performing at this standard ***require substantial improvement toward mastery*** of mathematics knowledge and skills. Students performing at this standard will likely need substantial support to get on track for success in high school and college coursework or career training.

The Connecticut ALDs for the Grades 3-5 ELA/literacy test are shown in the table below.

Achievement Level	Grades 3-5 Achievement-Level Descriptors for ELA and Literacy
Level 4	***Exceeds the Achievement Standard:*** The student has ***exceeded the achievement standard*** for English language arts and literacy expected for this grade. Students performing at this standard are ***demonstrating advanced progress toward mastery*** of English language arts and literacy knowledge and skills. Students performing at this standard are on track for likely success in the next grade.
Level 3	***Meets the Achievement Standard:*** The student has ***met the achievement standard*** for English language arts and literacy expected for this grade. Students performing at this standard are ***demonstrating progress toward mastery*** of English language arts and literacy knowledge and skills. Students performing at this standard are on track for likely success in the next grade.
Level 2	***Approaching the Achievement Standard:*** The student has ***nearly met the achievement standard*** for English language arts and literacy expected for this grade. Students performing at this standard ***require further development toward mastery*** of English language arts and literacy knowledge and skills. Students performing at this standard will likely need support to get on track for success in the next grade.
Level 1	***Does Not Meet the Achievement Standard:*** The student has ***not yet met the achievement standard*** for English language arts and literacy expected for this grade. Students performing at this standard ***require substantial improvement toward mastery*** of English language arts and literacy knowledge and skills. Students performing at this standard will likely need substantial support to get on track for success in the next grade.

The Connecticut ALDs for the Grades 6-8 ELA/literacy test are shown in the table below.

Achievement Level	Grades 6-8 Achievement-Level Descriptors for ELA and Literacy
Level 4	***Exceeds the Achievement Standard:*** The student has ***exceeded the achievement standard*** for English language arts and literacy expected for this grade. Students performing at this standard are ***demonstrating advanced progress toward mastery*** of English language arts and literacy knowledge and skills. Students performing at this standard are on track for likely success in high school and college coursework or career training.
Level 3	***Meets the Achievement Standard:*** The student has ***met the achievement standard*** for English language arts and literacy expected for this grade. Students performing at this standard are ***demonstrating progress toward mastery*** of English language arts and literacy knowledge and skills. Students performing at this standard are on track for likely success in high school and college coursework or career training.
Level 2	***Approaching the Achievement Standard:*** The student has ***nearly met the achievement standard*** for English language arts and literacy expected for this grade. Students performing at this standard ***require further development toward mastery*** of English language arts and literacy knowledge and skills. Students performing at this standard will likely need support to get on track for success in high school and college coursework or career training.
Level 1	***Does Not Meet the Achievement Standard:*** The student has ***not yet met the achievement standard*** for English language arts and literacy expected for this grade. Students performing at this standard ***require substantial improvement toward mastery*** of English language arts and literacy knowledge and skills. Students performing at this standard will likely need substantial support to get on track for success in high school and college coursework or career training.

Content-Specific Claim Scores

Content-specific claim scores are useful when analyzing data about the knowledge and skills students are expected to demonstrate on the assessment related to a particular aspect of the Connecticut Core Standards. For example, mathematics Claim 2, Problem Solving and Modeling & Data Analysis, provides evidence of each student's ability to solve a range of well-posed problems in pure and applied mathematics, making constructive use of prior knowledge and problem solving strategies. This claim addresses the core of mathematical expertise—the set of competencies that students can use when they are confronted with challenging tasks. English language arts/literacy Claim 1, Reading, provides evidence of each student's ability to read closely and analytically to comprehend a range of increasingly complex literary and informational texts. Being able to read and analyze a variety of complex texts helps students make sense of information, understand diverse viewpoints, and become active, productive, and informed citizens.

The content-specific claim scores are referred to as Areas of Knowledge and Skills on the paper version of the Individual Student Report. The content-specific claim scores are reported as Above Standard, Approaching Standard, or Below Standard.

Mathematics

For mathematics, content-specific claim scores are reported for Claim 1, and Claim 3. Claims 2 and 4 have been combined by the Connecticut State Department of Education into a single reporting category.

English Language Arts/Literacy

For ELA/literacy, content-specific claim scores are reported for Claim 1 and Claim 3. Claims 2 and 4 have been combined by the Connecticut State Department of Education into a single reporting category, Writing and Research/Inquiry.

Aggregate Target-Level Results

Unlike an overall ELA or mathematics score, the Assessment Target Report does not present absolute performance; instead it presents relative performance. The report provides an indicator of relative strength/weakness. Target-level results are provided for a group of students, but not for an individual student.

To determine relative strength/weakness, the actual performance of the group of students on the items in a particular target is compared to their expected performance on those items. If actual performance is significantly better than expected performance, then the group receives a "+." If actual performance is significantly worse than expected performance, then the group receives a "-." If actual performance is statistically no different than expected performance, then the group receives an "=" for that target.

Two statistical approaches are used to establish expected student performance.

Relative to Overall Performance
The expected performance is determined based on the students' overall performance on the entire test. For example, if the students in the group are extremely high performing overall, those students will likely be expected to do well on items in each target. If however, they do significantly worse than expected, then a minus sign may be displayed. This may not mean that the students are really low performing on that target; it may simply mean that their performance on that target was significantly lower than expected.

Icon	Target Level	Description
+	Better than performance on the test as a whole	This target is a relative strength. The group of students performed better on items from this target than they did on the test as a whole.
=	Similar to performance on the test as a whole	This target is neither a relative strength nor a relative weakness. The group of students performed about as well on items from this target as they did on the test as a whole.
−	Worse than performance on the test as a whole	This target is a relative weakness. The group of students did not perform as well on items from this target as they did on the test as a whole.
*	Insufficient Information	Not enough information is available to determine whether this target is a relative strength or weakness.

Relative to (Minimum Overall) Proficiency
The expected performance is determined based on a hypothetical student with minimum overall proficiency – one who is performing at the cut score separating levels 2 and 3 (i.e., the lowest score in level 3). Continuing the above example, the extremely high performing group may have done worse than expected on a target with somewhat challenging items but still better than the minimum overall proficiency would have done on those items. These students may earn a "+" to mean that their "Performance is above the Proficiency Standard."

Icon	Target Level	Description
+	Performance is above the Proficiency Standard	The target performance is above the proficiency standard. The group of students performed above the proficiency standard on this target.

=	Performance is near the Proficiency Standard	The target performance is near the proficiency standard. The group of students performed near the proficiency standard on this target.
-	Performance is below the Proficiency Standard	The target performance is below the proficiency standard. The group of students performed below the proficiency standard on this target.
*	Insufficient Information	Not enough information is available to determine whether the performance on this target is above, near, or below the proficiency standard.

When used together, the two methods can provide greater insight into a group of students' strengths and weaknesses. The table below illustrates how to interpret the results for a target, based on the results from the two approaches.

		Relative to Overall Performance		
		-	=	+
Relative to (Minimum Overall) Proficiency	-	Worse than expected and below the proficiency standard	As expected but below the proficiency standard	Better than expected but below the proficiency standard
	=	Worse than expected but near the proficiency standard	As expected and near the proficiency standard	Better than expected but near the proficiency standard
	+	Worse than expected but above the proficiency standard	As expected but above the proficiency standard	Better than expected and above the proficiency standard

Target reports are not available for an individual student because the number of items administered to individual students by target is not large enough to provide a score from which valid inferences can be made.

Interpreting Scores in the Early Years of Implementation

Because the CCS for each grade level build on learning at prior grade levels, students' instructional experience with CCS-aligned curriculum and pedagogical strategies should also be considered when interpreting test results. In the early years of implementation, this may be an important consideration for students at higher grade levels. One must keep in mind that when new content standards are assessed, the summative assessment scores will reflect both the degree to which the content standards are well implemented in a school and the degree to which students have learned them. Summative assessment results should be viewed as one indicator among multiple sources of evidence such as classroom-based tests, course grades, and samples of student work when making decisions about student performance.

IMPORTANT NOTE: Connecticut discontinued the ELA Performance Task for 2015-16; however, the 2014-15 ELA results include the Performance Task. Therefore, to enable the most valid comparison of aggregate results from 2014-15 to 2015-16 and 2016-17, the CSDE scored the 2014-15 test unofficially based solely on the computer-adaptive test (CAT) portion of the ELA

test. Aggregate district-, school-, and grade-level results are provided for district/schools to enable valid comparisons of the “CAT-only” ELA results from 2014-15 to 2015-16 and 2016-17. The discontinuance of the ELA Performance Task affected the Writing and Research/Inquiry claims, but did not affect the Reading and Listening claims; therefore, claim performances in Reading and Listening can be compared between 2014-15, 2015-16, and 2016-17.

Individual Student Reports

Sample Individual Student Reports (ISRs) for Grades 5 and 8 are provided on the pages that follow.

Two paper copies of Individual Student Reports are shipped to the districts. One copy is to be provided to parents or guardians by the school district and the other is to be retained by the district for the student’s cumulative record. The Individual Student Report provides a summary of the student’s performance on the mathematics and ELA tests.

On page 2, a customized message indicates the student’s overall performance for each content area. Below the message is a chart that indicates student achievement over the past two test years. This includes the total scale score and an achievement level. If a student only took one test over the previous two years, only one set of test scores will be displayed. In addition to the student score, the district and school average is presented for comparison purposes. A measurement error band is described, indicating the range of scores the student would likely receive if the test were taken several times.

Information is provided about the student’s performance on the Areas of Knowledge and Skill for each content area. These scores are reported as Above Standard, Approaching Standard, and Below Standard.

Sample Grade 5 Individual Student Report

CONNECTICUT STATE DEPARTMENT OF EDUCATION CSDE

Student Name:	**Jonathan Doe**		
Grade:	**05**	School:	**Demo Elementary School**
Date of Birth:	**05/20/2005**	District:	**Demo District**
SASID:	**1234567891**	Test Year:	**2017**

Connecticut Smarter Balanced Summative Assessment Results

The 2017 Connecticut Smarter Balanced Summative Assessments are administered to students in grades 3 – 8 in English Language Arts (ELA)/Literacy and Mathematics. This report shows Jonathan's achievement on assessments based on the Connecticut Core Standards, which define learning expectations for what students should know and be able to do at each grade level.

Connecticut has a comprehensive plan for college and career readiness that includes challenging academic standards and assessments to measure student progress. Results from the Connecticut Smarter Balanced Summative Assessments are only one indicator of a student's performance. These results should be used along with other information, such as classwork and other tests, when making educational decisions. Specific questions about individual student results should be directed to local school personnel.

Scale Scores and Performance Levels

Overall scores from ELA/Literacy and Mathematics tests are reported in scale-score units. Within the scale-score range, four performance levels have been established for each content area. Scoring in the Level 3 or Level 4 range is a challenging, yet reasonable, expectation for Connecticut students.

English Language Arts/Literacy

Jonathan's overall ELA/Literacy scale score is reported as well as the associated performance level. The school and district average scale scores are also reported for comparative purposes. The ELA/Literacy test assesses mastery of grade-level English language arts and Literacy in three areas of knowledge and skills aligned to the Connecticut Core Standards.

Three Areas of Knowledge and Skills

Reading
Students can read closely and analytically to comprehend a range of increasingly complex literary and informational texts.

Listening
Students can employ effective speaking and listening skills for a range of purposes and audiences.

Writing and Research/Inquiry
Students can strengthen writing by revising, editing, and rewriting a range of text. Students can also analyze and integrate evidence-based information to support analysis and research.

Mathematics

Jonathan's overall Mathematics scale score is reported as well as the associated performance level. The school and district average scale scores are also reported for comparative purposes. The Mathematics test assesses mastery of grade-level mathematics in three areas of knowledge and skills aligned to the Connecticut Core Standards.

Three Areas of Knowledge and Skills

Concepts and Procedures
Students can explain and apply mathematical concepts and interpret and carry out mathematical procedures with precision and fluency.

Problem Solving and Modeling & Data Analysis
Students can solve a range of complex, well-posed problems in pure and applied mathematics, making productive use of knowledge and problem solving strategies. Students can analyze complex, real-world scenarios and can construct and use mathematical models to interpret and solve problems.

Communicating Reasoning
Students can clearly and precisely construct viable arguments to support their own reasoning and to critique the reasoning of others.

Frequently Asked Questions

Where can I find more information about Connecticut academic standards and state assessments?

Information about Connecticut academic standards and state assessments is available on the Connecticut State Department of Education Web site (www.ct.gov/sde).

What are the ELA/Literacy and Mathematics Areas of Knowledge and Skills?

Each area is based on a summary statement about the knowledge and skills students are expected to demonstrate on the assessment. Each relates to a particular aspect of the Connecticut Core Standards. These indicators provide additional information about a student's overall score.

Where can I find more information about the Smarter Balanced Assessment System?

Information about the Smarter Balanced Assessment System is available at www.smarterbalanced.org.

Sample Grade 5 Individual Student Report – Page 2

CONNECTICUT STATE DEPARTMENT OF EDUCATION CSDE

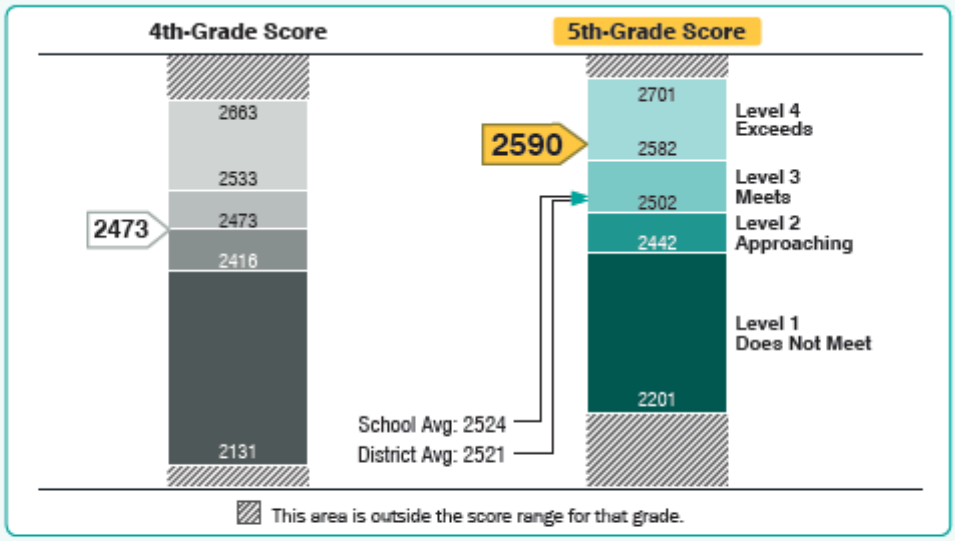
Student Name:	**Jonathan Doe**		
Grade:	**05**	School:	**Demo Elementary School**
Date of Birth:	**05/20/2005**	District:	**Demo District**
SASID:	**1234567891**	Test Year:	**2017**

Jonathan's ELA/Literacy Score for 2017

2590
Level 4
Exceeds

Jonathan has **exceeded the achievement standard** for English language arts and literacy expected for this grade. Students performing at this standard are **demonstrating advanced progress toward mastery** of English language arts and literacy knowledge and skills. Students performing at this standard are on track for likely success in the next grade.

Areas of Knowledge and Skill	Performance
Reading	Above Standard
Listening	Approaching Standard
Writing and Research/Inquiry	Above Standard



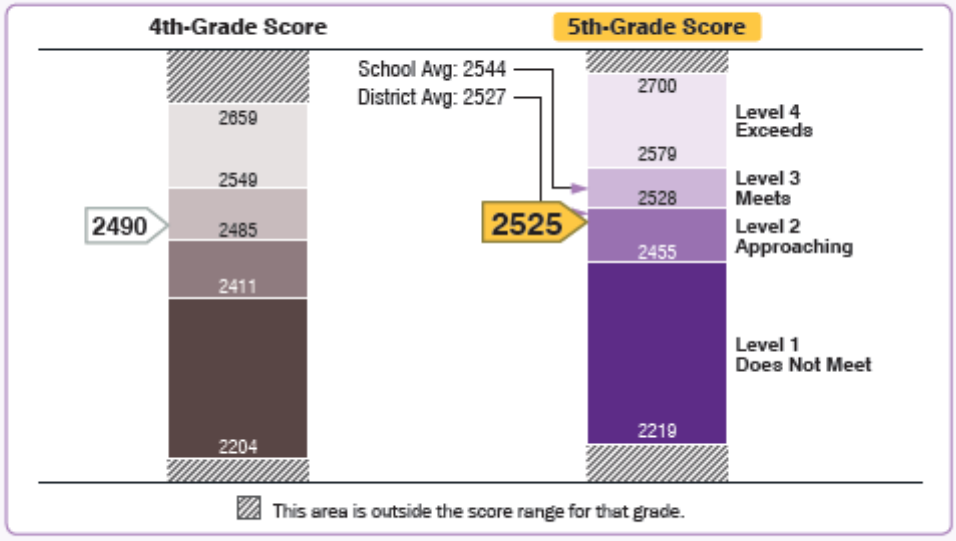
A student's test scores can vary if tests are taken several times. If Jonathan were tested again on ELA/Literacy, the new scale-score would probably fall between 2580 and 2600.

Jonathan's Mathematics Score for 2017

2525
Level 2
Approaching

Jonathan has **nearly met the achievement standard** for Mathematics expected for this grade. Students performing at this standard **require further development toward mastery** of Mathematics knowledge and skills. Students performing at this standard will likely need support to get on track for success in the next grade.

Areas of Knowledge and Skill	Performance
Concepts and Procedures	Below Standard
Problem Solving and Modeling & Data Analysis	Approaching Standard
Communicating Reasoning	Approaching Standard



A student's test scores can vary if tests are taken several times. If Jonathan were tested again on Mathematics, the new scale-score would probably fall between 2513 and 2537.

Sample Grade 8 Individual Student Report

Student Name:	**Jacqueline Doe**		
Grade:	**08**	School:	**Demo Middle School**
Date of Birth:	**05/20/2002**	District:	**Demo District**
SASID:	**1234567892**	Test Year:	**2017**

Connecticut Smarter Balanced Summative Assessment Results

The 2017 Connecticut Smarter Balanced Summative Assessments are administered to students in grades 3 – 8 in English Language Arts (ELA)/Literacy and Mathematics. This report shows Jacqueline's achievement on assessments based on the Connecticut Core Standards, which define learning expectations for what students should know and be able to do at each grade level.

Connecticut has a comprehensive plan for college and career readiness that includes challenging academic standards and assessments to measure student progress. Results from the Connecticut Smarter Balanced Summative Assessments are only one indicator of a student's performance. These results should be used along with other information, such as classwork and other tests, when making educational decisions. Specific questions about individual student results should be directed to local school personnel.

Scale Scores and Performance Levels

Overall scores from ELA/Literacy and Mathematics tests are reported in scale-score units. Within the scale-score range, four performance levels have been established for each content area. Scoring in the Level 3 or Level 4 range is a challenging, yet reasonable, expectation for Connecticut students.

English Language Arts/Literacy

Jacqueline's overall ELA/Literacy scale score is reported as well as the associated performance level. The school and district average scale scores are also reported for comparative purposes. The ELA/Literacy test assesses mastery of grade-level English language arts and Literacy in three areas of knowledge and skills aligned to the Connecticut Core Standards.

Three Areas of Knowledge and Skills

Reading
Students can read closely and analytically to comprehend a range of increasingly complex literary and informational texts.

Listening
Students can employ effective speaking and listening skills for a range of purposes and audiences.

Writing and Research/Inquiry
Students can strengthen writing by revising, editing, and rewriting a range of text. Students can also analyze and integrate evidence-based information to support analysis and research.

Mathematics

Jacqueline's overall Mathematics scale score is reported as well as the associated performance level. The school and district average scale scores are also reported for comparative purposes. The Mathematics test assesses mastery of grade-level mathematics in three areas of knowledge and skills aligned to the Connecticut Core Standards.

Three Areas of Knowledge and Skills

Concepts and Procedures
Students can explain and apply mathematical concepts and interpret and carry out mathematical procedures with precision and fluency.

Problem Solving and Modeling & Data Analysis
Students can solve a range of complex, well-posed problems in pure and applied mathematics, making productive use of knowledge and problem solving strategies. Students can analyze complex, real-world scenarios and can construct and use mathematical models to interpret and solve problems.

Communicating Reasoning
Students can clearly and precisely construct viable arguments to support their own reasoning and to critique the reasoning of others.

Frequently Asked Questions

Where can I find more information about Connecticut academic standards and state assessments?

Information about Connecticut academic standards and state assessments is available on the Connecticut State Department of Education Web site (www.ct.gov/sde).

What are the ELA/Literacy and Mathematics Areas of Knowledge and Skills?

Each area is based on a summary statement about the knowledge and skills students are expected to demonstrate on the assessment. Each relates to a particular aspect of the Connecticut Core Standards. These indicators provide additional information about a student's overall score.

Where can I find more information about the Smarter Balanced Assessment System?

Information about the Smarter Balanced Assessment System is available at www.smarterbalanced.org.

Sample Grade 8 Individual Student Report – Page 2

CONNECTICUT STATE DEPARTMENT OF EDUCATION CSDE

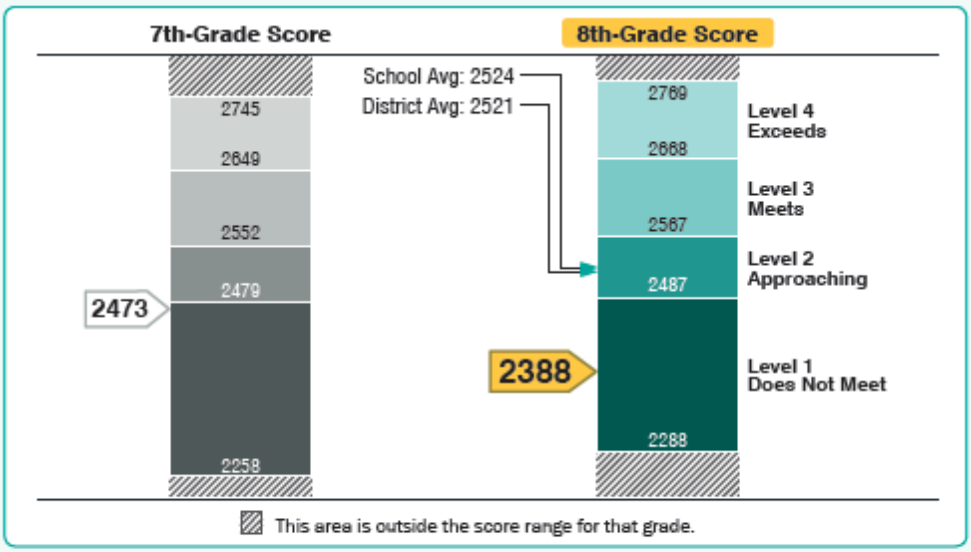
Student Name:	**Jacqueline Doe**		
Grade:	**08**	School:	**Demo Middle School**
Date of Birth:	**05/20/2002**	District:	**Demo District**
SASID:	**1234567892**	Test Year:	**2017**

Jacqueline's ELA/Literacy Score for 2017

2388
Level 1
Does Not Meet

Jacqueline has **not yet met the achievement standard** for English language arts and literacy expected for this grade. Students performing at this standard **require substantial improvement toward mastery** of English language arts and literacy knowledge and skills. Students performing at this standard will likely need substantial support to get on track for success in high school and college coursework or career training.

Areas of Knowledge and Skill	Performance
Reading	Below Standard
Listening	Approaching Standard
Writing and Research/Inquiry	Below Standard



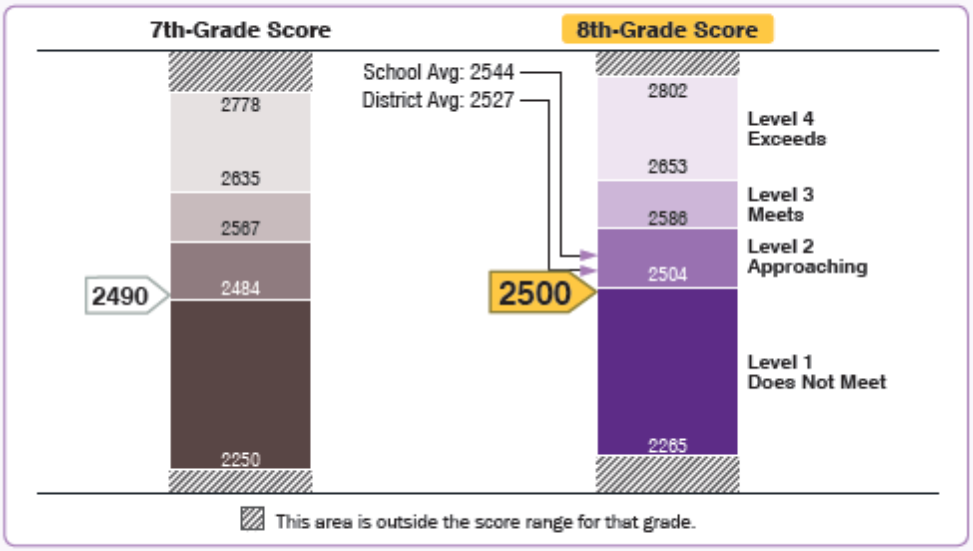
A student's test scores can vary if tests are taken several times. If Jacqueline were tested again on ELA/Literacy, the new scale-score would probably fall between 2378 and 2398.

Jacqueline's Mathematics Score for 2017

2500
Level 1
Does Not Meet

Jacqueline has **not yet met the achievement standard** for Mathematics expected for this grade. Students performing at this standard **require substantial improvement toward mastery** of Mathematics knowledge and skills. Students performing at this standard will likely need substantial support to get on track for success in high school and college coursework or career training.

Areas of Knowledge and Skill	Performance
Concepts and Procedures	Below Standard
Problem Solving and Modeling & Data Analysis	Below Standard
Communicating Reasoning	Approaching Standard



A student's test scores can vary if tests are taken several times. If Jacqueline were tested again on Mathematics, the new scale-score would probably fall between 2490 and 2510.

References

Core Standards. “About the Standards.” In Common Core State Standards Initiative. Accessed August 4, 2016. http://www.corestandards.org/about-the-standards/.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, *Standards for Educational and Psychological Testing.* Washington, D.C.: AERA, 2014.

American Institutes for Research. “Online Reporting System User Guide, 2016-2017.” In Connecticut State Department of Education Comprehensive Assessment Program Portal. Accessed June 29, 2017. http://ct.portal.airast.org/wp-content/uploads/CT_ORS_Guide.pdf.

American Institutes for Research. (2015). *Smarter Balanced Scoring Specifications Summative and Interim Assessments: ELA, Grades 3-8, 11 and Mathematics, Grades 3-8, 11. http://www.sde.ct.gov/sde/lib/sde/pdf/student_assessment/smarter_balanced/reporting/sb_testscoringspecs2014-2015.pdf*.

Smarter Balanced. “Content Specifications for the Summative Assessment of the Common Core State Standards for Mathematics.” Accessed June 29, 2017. https://portal.smarterbalanced.org/library/en/mathematics-content-specifications.pdf.

Smarter Balanced. “Content Specifications for the Summative Assessment of the Common Core State Standards for English Language Arts/Literacy in History, Social Studies, Science, and Technical Subjects.” Accessed June 29, 2017 https://portal.smarterbalanced.org/library/en/appendix-b-grade-level-tables.pdf.

L. Hammond, E. Haertel, J. Pellegrino. (2015). “Making Good Use of New Assessment: Interpreting and Using Scores from the Smarter Balanced Assessment Consortium.” (Working paper). Olympia, WA: Smarter Balanced Assessment Consortium.

The National Academies (2007). *Lessons learned about testing: Ten years of work at the National Research Council*. Washington, D.C.

Smarter Balanced. “Smarter Balanced Assessment Consortium: Theory of Action, an excerpt from the Smarter Balanced Race to the top application.” (2010). Accessed July 14, 2015. http://files.eric.ed.gov/fulltext/ED536956.pdf.

Core Standards. “Standards in Your State.” (2015). In Common Core State Standards Initiative. Accessed June 29, 2017. http://www.corestandards.org/standards-in-your-state/.

Tan, X., & Michel, R. “Why do standardized testing programs report scaled scores?” *R & D Connections,* 16 (2011): 1-6.

Smarter Balanced Scoring and Reporting FAQ – Appendix A

1. How are partially completed tests handled in participation reports and score reports?

Below are the Connecticut rules for calculating participation and performance:

- **Participation Reports**: Students are reported as having "participated" in the test if they logged in to both the Performance Task and CAT parts of the mathematics test, even if they did not answer any items. Students are reported as having "participated" in the test if they logged in to the CAT part of the ELA test, even if they did not answer any items.
- **Score Reports:** For a student's score to be reported, the student must have answered at least one CAT item **and** one Performance Task item in mathematics, and one item in the CAT for ELA.

2. What is a scale score?

- Scale scores are the basic units of reporting. These scores, which fall along a continuous vertical scale (from approximately 2000 to 3000) that increases across grade levels, can be used to illustrate students' current level of achievement and their growth over time.
- When aggregated, scale scores, unlike raw scores, can also describe school- or district-level changes in performance on the tests and can measure gaps in achievement among different groups of students.

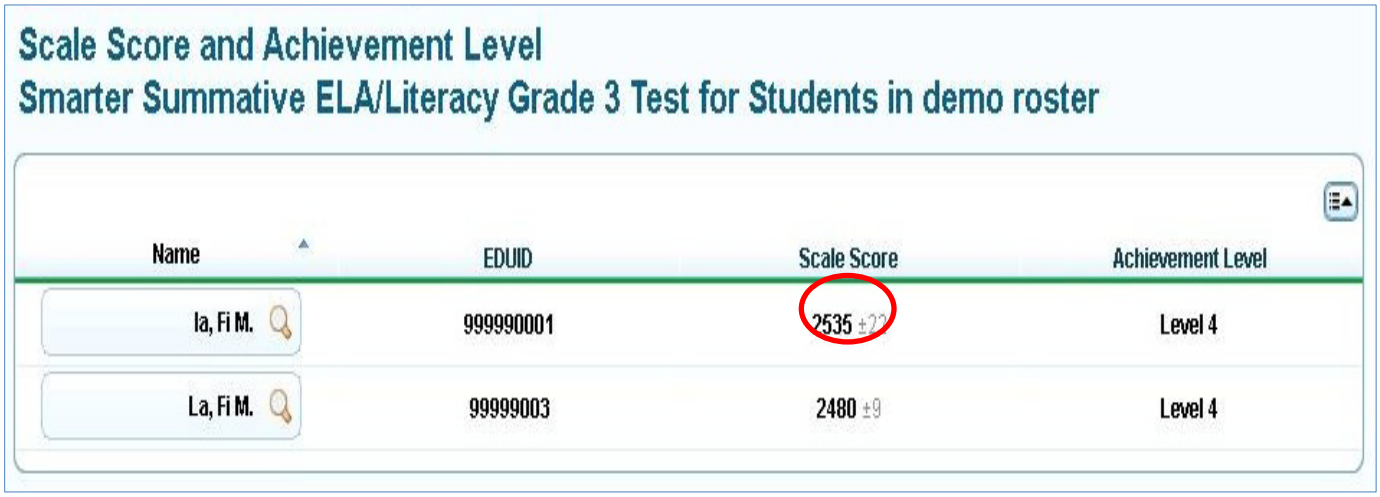
3. What is the standard error of measurement (SEM)?

The standard error of measurement (SEM) allows users to estimate the score range that a student would likely fall within if the student took the same Smarter Balanced English language arts or mathematics test multiple times with exactly the same level of knowledge and preparation. For example, as seen in Figure 1, a scale score of 2535 ±22 (circled in red) indicates that if the student could take the same test multiple times, the score would likely fall between 2513 and 2557. Scale scores will vary based on the test and on the student.

4. Why is it important to report the standard error of measurement?

Reporting the SEM is important because a student's score is best interpreted when recognizing that the student's knowledge and skills fall within a score range. All test results, including scores on tests and quizzes designed by classroom teachers, are subject to measurement error.

Figure 1: Example of a Student Listing in the Online Reporting System That Displays Both Scale Scores and SEM.



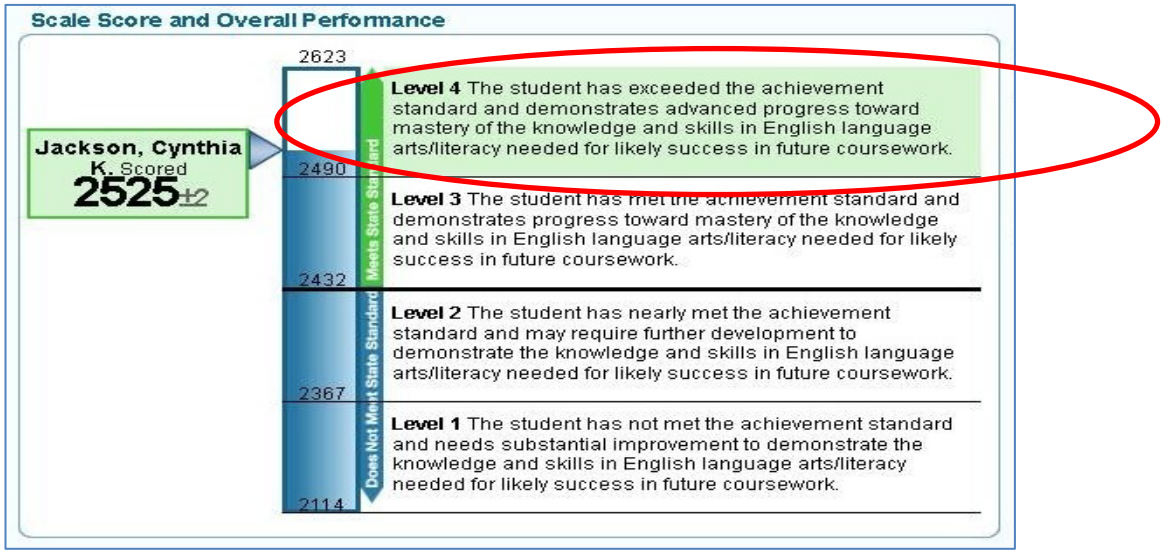
5. What do achievement levels represent and why are they useful?

- Achievement levels are categories used to describe student performance based on scale scores.
- A high score will place a student in a high achievement level. Generally, a higher score on the test reflects a greater accumulation of knowledge, skills, and processes when compared to students earning scores in lower achievement levels.
- The achievement levels on the Smarter Balanced Summative Assessment is Level 1, Level 2, Level 3, and Level 4. Level 4 is the highest performance level.

6. What are Achievement-Level Descriptors?

- Achievement-Level Descriptors (ALDs) describe a student's overall content readiness in the core areas of ELA and mathematics for a specific grade level.
- ALDs communicate the meaning of test scores by specifying, in content terms, the knowledge, skills, and processes that student's display at four levels of achievement.
  For example, Figure 2 shows a student scale score of 2525 on the Grade 3 English language arts/literacy test. That student's score is higher than the threshold for Level 4, which places him or her in Level 4. Level 4 is the highest achievement level of the Smarter Balanced Summative and Interim Comprehensive tests, indicating that the student has "exceeded the achievement standard and demonstrates advanced progress." The full achievement-level text is circled in the example image for Grade 3 ELA/literacy.
- Achievement-Level Descriptors are cumulative, where the knowledge, skills, and processes of lower level ALDs are assumed by the higher level ALDs. For instance, the Level 4 student in the example in Figure 2 is assumed to possess the knowledge, skills, and processes described in Levels 1, 2, and 3.
  Note from this example that the achievement-level scale scores are not always spaced evenly apart.

Figure 2: Example of an Individual Student Report Showing Achievement Levels.



7. Who determines where one achievement level ends and the next begins?

The scores that separate achievement levels from one another are called threshold scores. Threshold scores and achievement levels were developed by thousands of K–12 educators, higher education faculty, experts, parents, and other stakeholders through a process called standard setting.

8. What are assessment claims?

- Assessment claims are broad evidence-based statements about what students know and can do as demonstrated by their performance on the assessments.
- In addition to receiving scale scores and achievement levels for the mathematics and English language arts tests, students are also placed into performance categories by assessment claim within each subject.

9. What are the assessment claim performance categories and how are they derived?

- For each subject area assessment claim, students are assigned to one of three performance categories: Below Standard, Approaching Standard, or Above Standard.
- In ELA/literacy, performance categories are reported for three assessment claims (Reading; Listening; and Writing and Research/Inquiry).
- In mathematics, performance categories are reported for three assessment claims (Concepts and Procedures; Communicating Reasoning; and Problem Solving and Modeling and Data Analysis).
- A student's performance category for an assessment claim is derived from the student's performance on the items linked to that claim.
- For example, a student may receive a scale score on the Smarter Balanced Summative Grade 3 ELA/literacy test that places him or her in achievement level 3 (meeting standard). The student may have performed "Above Standard" on Reading and Listening, and "Approaching Standard" in Writing and Research/Inquiry. These performance categories contain information about a student's strengths and weaknesses that may be useful to parents and teachers.

Figure 3a: Student Performance on Each Claim

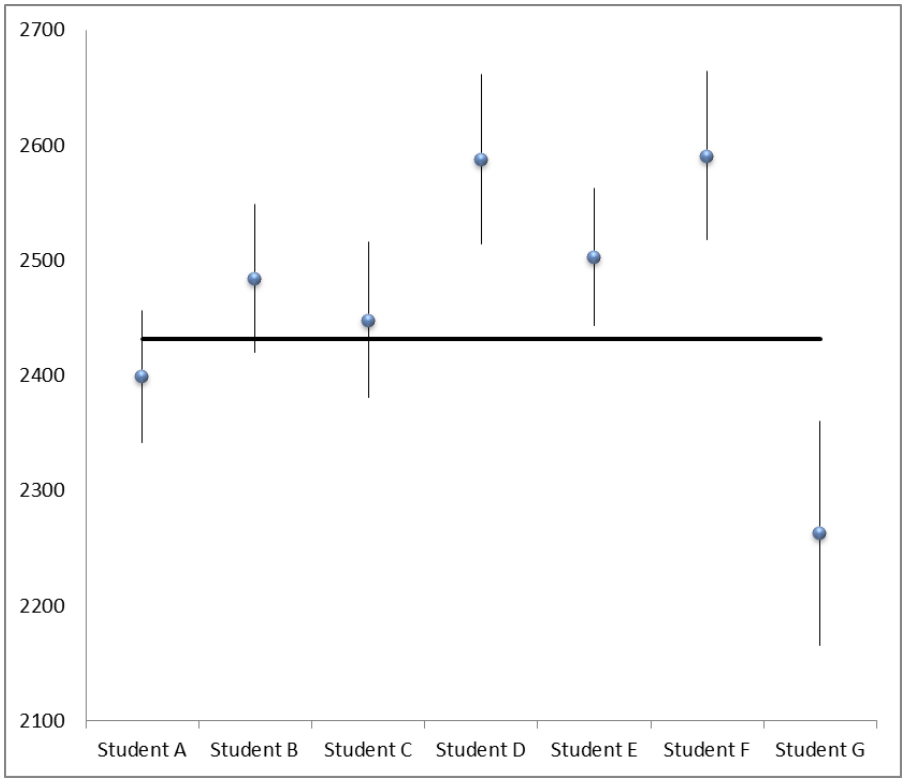
Name	SSID	Opportunities Taken	Scale Score	Achievement Level	Reading Achievement Category	Writing Achievement Category	Listening Achievement Category	Research/Inqu Achievement Category
Lastname, Firstname M.	9999990027	1	2482 ±23	3				
Lastname, Firstname M.	9999990039	1	2398 ±22	2				
Lastname, Firstname M.	9999990100	1	2385 ±22	2				
Lastname, Firstname M.	9999990026	1	2288 ±29	1				
Lastname, Firstname M.	9999990004	1	2623 ±29	4				
Lastname, Firstname M.	9999990008	1	2460 ±22	3				
Lastname, Firstname M.	9999990012	1	2420 ±22	2				
Testnew, test	567656765	1	2425 ±22	2				

10. How is a student's assessment claim performance category determined?

Assessment claim performance categories are assigned based on sub-scores. These sub-scores are derived from clusters of items in both the Computer Adaptive Test (CAT) and Performance Task (PT) portions for mathematics, and the CAT for ELA. Like the overall vertical scale scores, these sub-scores range from approximately 2000 to 3000.

See Figure 3b below for examples of Grade 3 ELA/literacy student scores. The horizontal line represents the *overall* minimum assessment scale score needed for a student to be performing in the highest performance category. In this example, the minimum score is 2432. The blue dots represent the Claim 1 scores for seven students. The lines above and below the blue dot represent the confidence interval for the Claim 1 score for each student.

Figure 3b: Grade 3 ELA/Literacy Claim Performance Category Determination



Examples:

- If a student's Claim 1 score ***and*** confidence interval are ***entirely below*** the horizontal line, then that student is said to be performing Below Standard on that Claim (Student G).
- If a student's Claim 1 score ***and*** confidence interval touch the horizontal line, then that student is said to be performing At/Near Standard on that Claim (Students A, B, and C).
- If a student's Claim 1 score ***and*** confidence interval are ***entirely above*** the horizontal line, then that student is said to be performing Above Standard on that Claim (Students D, E, and F).

11. How are items weighted?

The answer to this question is a bit more complex than a simple "weighting" system. Smarter Balanced is a computer adaptive assessment. In general, two students will see different sets of

items. The items that are presented to a student are selected based on two primary factors:

1. The test blueprints (ELA and Math) which define the claims and assessment targets that must be covered; and
2. Student performance on the test (i.e., students who do well on the test will tend to see more difficult items and vice versa).

A student's overall score is determined dynamically based on two parameters of each test item that are taken by the student. The two parameters are item difficulty and item discrimination. Item difficulty tells us whether the item is easy or hard. Item discrimination tells us how well the item can discriminate between low and high performers. It is possible that two students may get the same number of items correct but end up with different scale scores based on the difficulty and discrimination parameters of the items they took.

12. What are assessment targets?

Targets are narrowly defined skills that are tied directly to the Connecticut Core Standards. Because of the breadth in coverage of the individual assessment claims, the targets within them are needed to define more specific performance expectations within claim statements.

- For example, the "Reading" claim on the Smarter Balanced Summative ELA test includes targets such as "Target 1: KEY DETAILS (Literary Text): Given an inference or conclusion, use explicit details and implicit information from the text to support the inference or conclusion provided." Targets are reported only at the aggregate level (district, school, and roster) and not for individual students.

13. What are strength and weakness indicators?

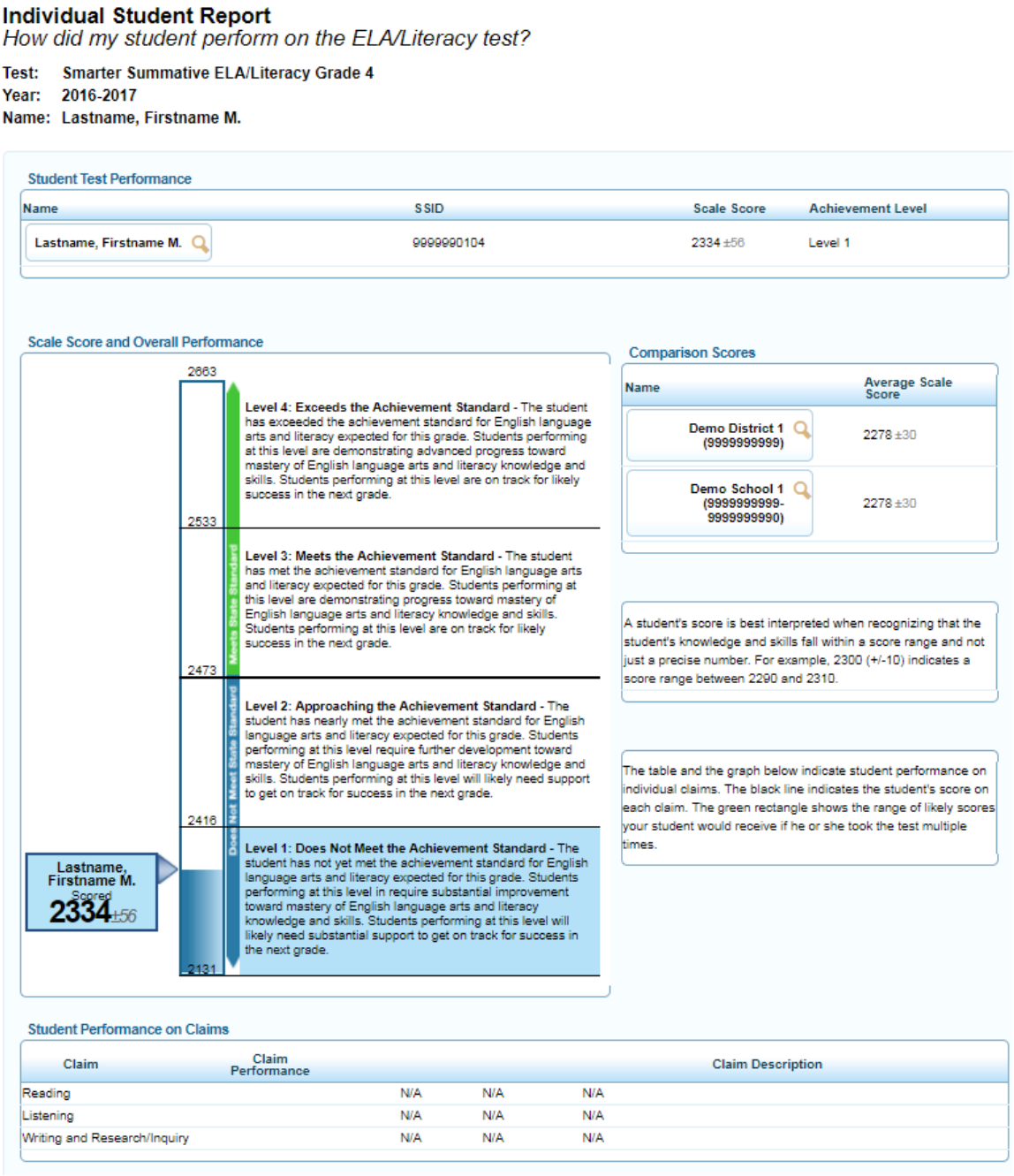
See page 17 and 18 of the Interpretive Guide.

14. Are summative and interim tests reported any differently?

Yes, there are differences in how interim and summative tests are reported. There are two types of Smarter Balanced Interim Assessments: Interim Assessment Blocks (IAB) and Interim Comprehensive Assessments (ICA). Below are the reporting features that are unique to the Smarter Balanced Summative, IAB, and ICA tests:

- **Summative tests** allow <u>one opportunity</u> per student. Summative reports, available through the Online Reporting System, include scale scores, achievement levels, claim-performance categories, and target strength and weakness indicators (see Figure 4a for an example of a Smarter Balanced Summative Individual Student Report).

Figure 4a: Individual Student Report in the Online Reporting System for a Summative Test.



Individual Student Report

How did my student perform on the ELA/Literacy test?

Test: Smarter Summative ELA/Literacy Grade 4
Year: 2016-2017
Name: Lastname, Firstname M.

Student Test Performance

Name	SSID	Scale Score	Achievement Level
Lastname, Firstname M.	9999990104	2334 ±56	Level 1

Scale Score and Overall Performance

2663

Level 4: Exceeds the Achievement Standard - The student has exceeded the achievement standard for English language arts and literacy expected for this grade. Students performing at this level are demonstrating advanced progress toward mastery of English language arts and literacy knowledge and skills. Students performing at this level are on track for likely success in the next grade.

2533

Meets State Standard

Level 3: Meets the Achievement Standard - The student has met the achievement standard for English language arts and literacy expected for this grade. Students performing at this level are demonstrating progress toward mastery of English language arts and literacy knowledge and skills. Students performing at this level are on track for likely success in the next grade.

2473

Does Not Meet State Standard

Level 2: Approaching the Achievement Standard - The student has nearly met the achievement standard for English language arts and literacy expected for this grade. Students performing at this level require further development toward mastery of English language arts and literacy knowledge and skills. Students performing at this level will likely need support to get on track for success in the next grade.

2416

Level 1: Does Not Meet the Achievement Standard - The student has not yet met the achievement standard for English language arts and literacy expected for this grade. Students performing at this level in require substantial improvement toward mastery of English language arts and literacy knowledge and skills. Students performing at this level will likely need substantial support to get on track for success in the next grade.

2131

Lastname, Firstname M.
Scored
2334 ±56

Comparison Scores

Name	Average Scale Score
Demo District 1 (9999999999)	2278 ±30
Demo School 1 (9999999999-9999999990)	2278 ±30

A student's score is best interpreted when recognizing that the student's knowledge and skills fall within a score range and not just a precise number. For example, 2300 (+/-10) indicates a score range between 2290 and 2310.

The table and the graph below indicate student performance on individual claims. The black line indicates the student's score on each claim. The green rectangle shows the range of likely scores your student would receive if he or she took the test multiple times.

Student Performance on Claims

Claim	Claim Performance			Claim Description
Reading	N/A	N/A	N/A	
Listening	N/A	N/A	N/A	
Writing and Research/Inquiry	N/A	N/A	N/A	

- **Interim Assessment Blocks (IABs)** consist of individually scored blocks. Students may be administered as many or as few blocks as necessary, and they may have multiple opportunities. The IAB reports look different from summative and ICA reports, because instead of a scale score and achievement level, students receive a Block Achievement Category for each block tested (see Figure 4b for an example of a Smarter Balanced IAB Individual Student Report). Scores and performance-level distributions are available by IAB, by class, and by individual student for eligible users through the Air Ways Reporting System (see Figure 4c).

- **Interim Comprehensive Assessments (ICAs)** use the same blueprint as summative tests and are reported in nearly the same way, except that target-level information is not

reported. Also note, that in the rare instance a student completes an ICA more than once, scores will be provided for each test-taking opportunity. Like the IABs, ICA data is available to eligible users through the AIR Ways Reporting System.

Figure 4b: Individual Student Report in the Online Reporting System for an IAB test.

Individual Student Report

How did my student perform on the test?

Test: Smarter Interim Assessment Blocks ELA/Literacy Grade 3
Year: 2016-2017
Name: Lastname, Firstname M.

Legend: Block Achievement Category

Below Standard | Approaching Standard | Above Standard

Student Information

Name	SSID	Opportunity
Lastname, Firstname M.	9999990026	Opportunity #1 11/13/2016

Student Test Performance

Block	
Read Literary Texts	Below Standard
Read Informational Texts	Approaching Standard
Editing	Below Standard
Language and Vocabulary Use	Below Standard
Revision	Below Standard
Brief Writes	Approaching Standard
Listen/Interpret	Below Standard
Research	Below Standard
Opinion Performance Task	Approaching Standard

Student Performance on Each Test Item - IAB Test

Item #/Target	Points Earned	Points Possible
Read Literary Texts		
1. CENTRAL IDEAS: Identify or summarize central ideas, key events, or the sequence of events presented in a text - Point(s) Possible: 1 Point	0	1
2. TEXT STRUCTURES & FEATURES: Relate knowledge of text structures or text features (e.g., illustrations) to gain, interpret, explain, or connect information - Point(s) Possible: 1 Point	0	1
3. REASONING & EVIDENCE: Use supporting evidence to interpret and explain inferences about character traits, motivations, feelings; point of view, author's lesson or message - Point(s) Possible: 1 Point	0	1
4. LANGUAGE USE: Interpret use of language by distinguishing literal from non-literal meanings of words and phrases used in context - Point(s) Possible: 1 Point	1	1
5. TEXT STRUCTURES & FEATURES: Relate knowledge of text structures or text features (e.g., illustrations) to gain, interpret, explain, or connect information - Point(s) Possible: 1 Point	0	1
6. ANALYSIS WITHIN OR ACROSS TEXTS: Specify or compare relationships across texts (e.g., literary elements, problem solution, theme) -		

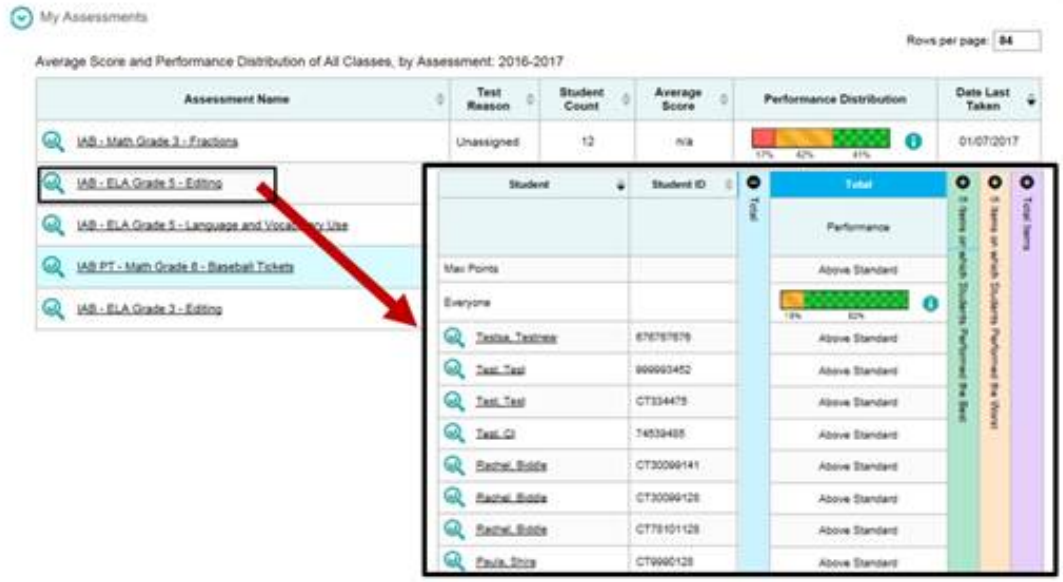



Figure 4c: Average Score and Performance for all Classes and a Class Roster