

2014-2015 Smarter Balanced Scoring and Reporting Factsheet

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

Below are the Smarter Balanced rules for calculating participation and performance. States may elect to implement the Consortium rules regarding calculating participation and achievement rates, but are not required to do so:

- **Participation Reports:** Students are reported as having “participated” in the test if they logged in to both the Performance Task and Non-Performance Task parts of the test, even if they did not answer any items. For computer-based Summative tests, the Non-Performance Task is a Computer Adaptive Test (CAT).
- **Score Reports:** For a student’s score to be reported, the student must have answered at least one Non-Performance Task (i.e. CAT) item **and** one Performance Task item.

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 in a relatively fine-grained fashion.
- 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 purpose of the Standard Error of Measurement (SEM) is to show the score range that a student would likely fall within if they took the same Smarter Balanced ELA or math 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, they would likely score between 2513 and 2557. Scale scores will vary based on the test and on the student.

Figure 1: Example of a Student Listing in the Online Reporting System that displays both scale scores and SEM.

Scale Score and Achievement Level			
Smarter Summative ELA/Literacy Grade 3 Test for Students in demo roster			
Name	EDUID	Scale Score	Achievement Level
Ia, Fi M.	999990001	2535 ± 22	Level 4
La, Fi M.	999990003	2480 ± 9	Level 4

4. Why is Standard Error of Measurement important?

The Standard Error of Measurement (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 and not just a precise number. All test results, including scores on tests and quizzes designed by classroom teachers, are subject to Standard Error of Measurement.

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

- Achievement levels are categories that describe what a student knows and is able to do, based on the student's scale score on the test
- A high score will place a student in a high achievement level. A higher score on the test reflects a greater accumulation of knowledge, skills, and processes.
- The achievement levels on the Smarter Balanced Summative and ICA tests are Level 1, Level 2, Level 3, and Level 4. To meet standard, a student must be in either the Level 3 or Level 4 achievement level.

6. What are achievement level descriptors?

- Achievement Level Descriptors, or ALDs, describe a student's overall content readiness in the core areas of ELA/literacy 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 students 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 (ELA) test. That student's score is higher than the threshold for Level 4, which places him or her in the Level 4 achievement level. 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.
- 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 be able 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 are developed by thousands of K–12 educators, higher education faculty, experts, parents, and other stakeholders through a process called standard setting.

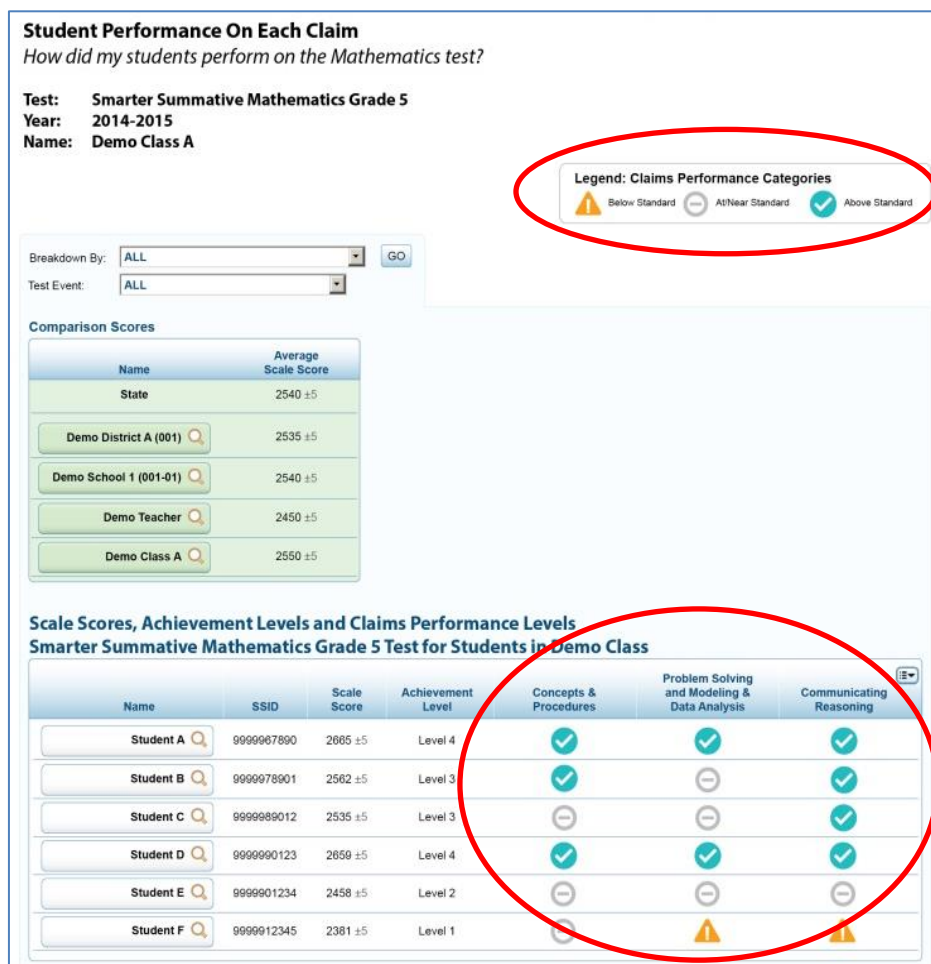
8. What are assessment claims?

- Assessment Claims are broad evidence-based statements about what students know and can do as demonstrated by their performance on the assessments.
- For example, there are three assessment claims on the Smarter Balanced Summative Grade 5 Math test: Concepts and Procedures, Problem Solving and Modeling & Data Analysis, and Communicating Reasoning.
- In addition to receiving a scale score and achievement level for the whole test, students are also placed into performance categories for specific claims.

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

- For each claim, students are assigned to one of three claim performance categories: Below Standard, At/Near Standard, or Above Standard.
- In ELA, claim performance categories will be reported for each of the four claims. In math, claim performance categories will be reported for Claim 1, Claims 2 and 4 combined, and Claim 3 (see Figure 3 for an example of a Claims Detail page).
- A student's performance category for each claim is derived based on their performance on the items contained in that claim.
- For example, a student may receive a scale score on the Smarter Balanced Summative Grade 3 ELA test that places him or her in the Level 3 performance level (meeting standard). The student may have performed "Above Standard" on the Reading and Writing claims, "At/Near Standard" on the Listening claim, and "Below Standard" on the Research/Inquiry claim. These claim performance categories contain useful information about math and ELA strengths and weaknesses for the student's teachers and parents.

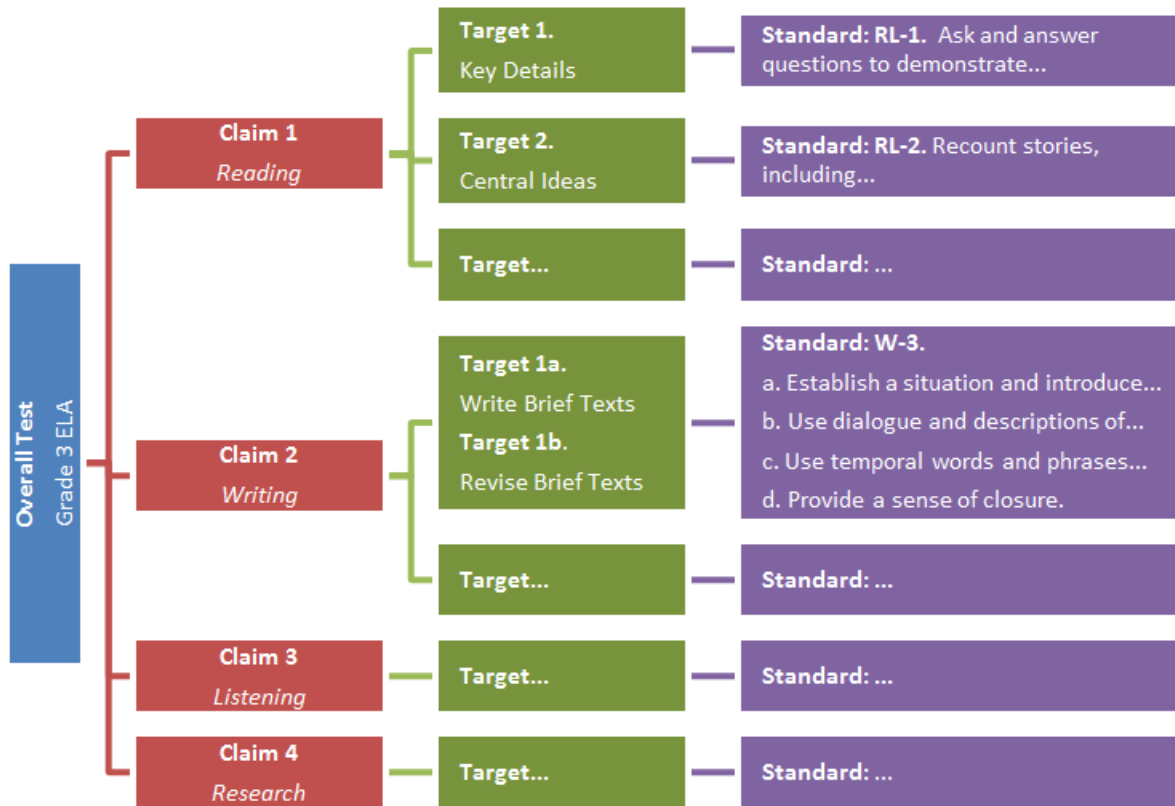
Figure 3: Claims Detail page with Claim Performance Categories (Online Reporting System)



11. How do claims scores relate to overall scores?

A student’s scale score and achievement level are based on the student’s performance on the test as a whole, while a student’s claim performance category is based only on the student’s performance on the test items contained in that particular claim (see Figure 4 for a chart showing the relationship between the test subject, claims, and targets).

Figure 4: Illustration of the relationship between test subject, claims, targets, and standards



12. What are assessment targets?

- Targets are narrowly defined skills that are tied directly to the Common Core standards. Because of the breadth in coverage of the individual 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 (Informational Text): Use explicit details and implicit information from the text to support answers or inferences about information presented.” Targets are reported only at the aggregate level (district, school, roster) and not for individual students.

13. What are strength and weakness indicators?

- Targets are assigned “strength and weakness indicators” instead of achievement levels or performance levels.
- Unlike achievement levels and claim performance categories, which are assigned to individual students, strength and weakness indicators display a group of students’ performance on a target **relative to the group’s performance on the test as a whole** (see Figure 5 for a description of the indicators).

- Unlike achievement levels based on overall test performance, these strengths and weaknesses do not imply proficiency. Instead, they show how a group of students' performance is distributed across the target content relative to their overall subject performance on a test.

Figure 5: Target Level Indicator Descriptions

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 rest of 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 rest of 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 rest of the test as a whole.
	Insufficient Information	Not enough information is available to determine whether this target is a relative strength or weakness.

14. How are strengths and weaknesses for targets determined?

- In math, target scores will be computed for the targets in Claim 1 only. In ELA, target scores will be computed for the targets within each Claim.
- Strengths and weaknesses are reported for groups of students based on whether there is a statistically significant difference between that group's performance on each target compared to their performance on the test as a whole (see Figure 6 for an example of a Target Detail page).
- For example, a group of students may have performed very well in a subject, but performed slightly lower in several targets. Thus, the minus sign for a target does not imply a lack of proficiency. Instead, it simply communicates that these students' performance on that target was statistically lower than their performance across all other targets put together. Although the students are doing well, an educator may want to focus instruction on these areas.

Figure 6. Target Detail page in the Online Reporting System showing strength and weakness indicators.

School Performance on Each Target for the ELA/Literacy Test
What are my school's relative strengths and weaknesses in the ELA/Literacy Targets?

Test: Smarter Summative ELA/Literacy Grade 3
 Year: 2014-2015
 Name: Demo inst 9997

Legend: Strength And Weakness Indicator

- Better than performance on the test as a whole
- Similar to performance on the test as a whole
- Worse than performance on the test as a whole
- Insufficient Information

Performance on Each Target
Smarter Summative ELA/Literacy Grade 3 Test for Students in Demo inst 9997

Target	Performance Level
Reading	
(Informational Text) KEY DETAILS: Use explicit details and implicit information from the text to support answers or inferences about information presented.	=
(Informational Text) CENTRAL IDEAS: Identify or summarize central ideas/ key events, or procedures and details that support them.	=
(Informational Text) WORD MEANINGS: Determine intended meanings of words, including domain-specific (tier 3) words and academic (tier 2) words with multiple meanings, based on context, word relationships, word structure (e.g., common roots, affixes), or use of resources (e.g., beginning dictionary, glossary)	=
(Informational Text) REASONING & EVIDENCE: Use supporting evidence to interpret and explain how information is presented or connected within or across texts (author's point of view, ideas and supporting details, relationships)	=
(Informational Text) ANALYSIS WITHIN OR ACROSS TEXTS: Specify, integrate, or compare information within or across texts (e.g., cause effect, integrate information)	=
(Informational Text) TEXT STRUCTURES/ FEATURES: Relate knowledge of text structures or text features (e.g., graphics, bold text, headings) to obtain, interpret, or explain information	=

Comparison Scores

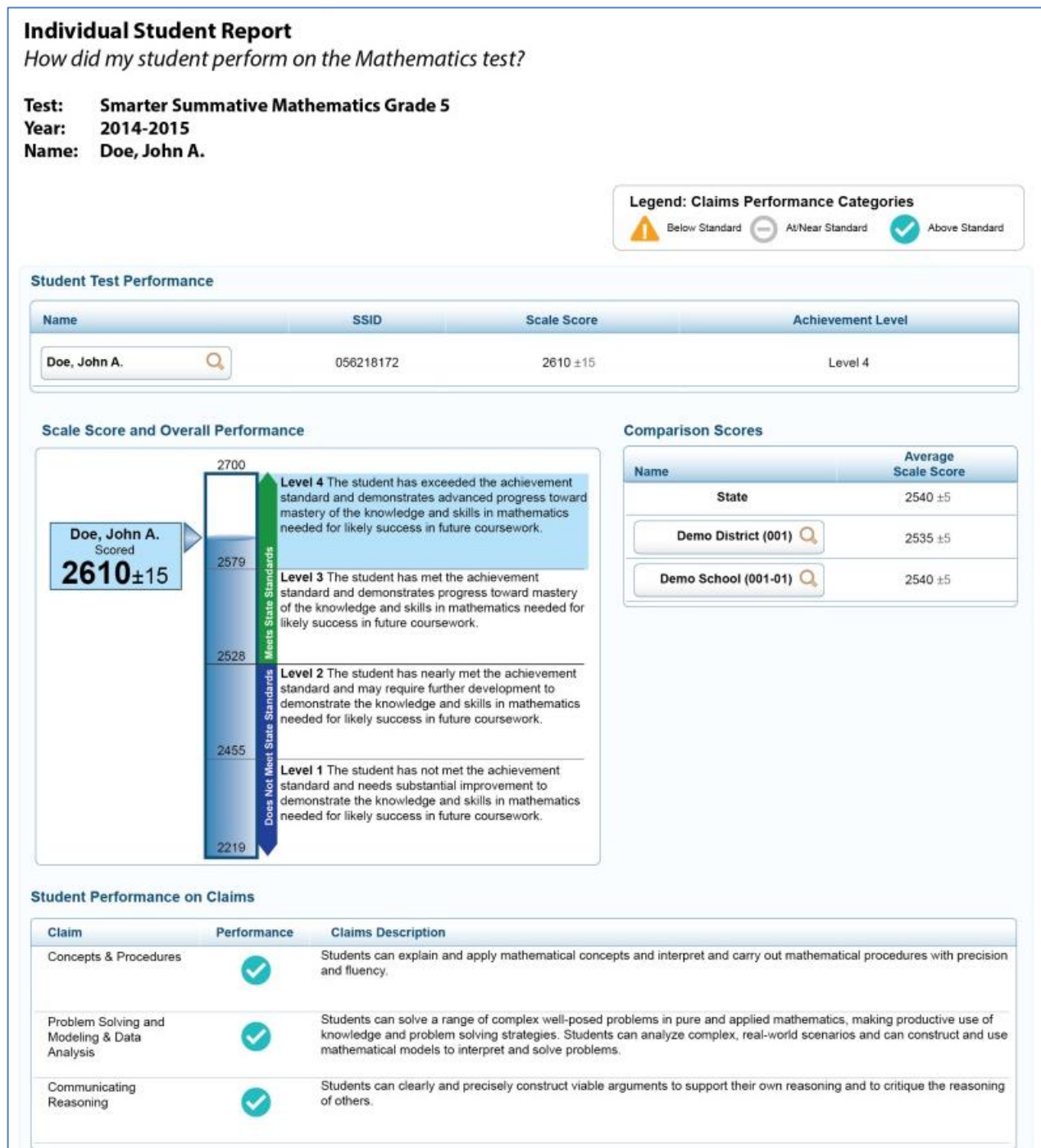
Name	Average Scale Score
Demo district 9999 (9999)	2402 ±12
Demo inst 9997 (9999_9997)	2411 ±13

15. 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 Comprehensive Assessments (ICA) and Interim Assessment Blocks (IAB). Below are the reporting features that are unique to the Smarter Balanced Summative, ICA, and IAB tests:

- **Summative tests** allow one opportunity per student. Summative reports include scale scores, achievement levels, claim performance categories, and target strength and weakness indicators (see Figure 7a for an example of a Smarter Balanced Summative Individual Student Report).

Figure 7a: Individual Student Report in the Online Reporting System for a Summative test.






- **ICA tests** use the same blueprint as summative tests and are reported in nearly the same way, except that scores are shown for multiple student opportunities if the student has taken the test multiple times, and target-level information is not reported (see Figure 7b for an example of a Smarter Balanced ICA Individual Student Report).

Figure 7b: Individual Student Report in the Online Reporting System for an ICA test.

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

Test: Smarter Interim Comprehensive Assessment ELA/Literacy Grade 3
 Year: 2015-2016
 Name: Umpt, Stzn L.

Legend: Claims Performance Levels

 Below Standard  At/Near Standard  Above Standard

Student Test Performance

Name	EDUID	Opportunity	Scale Score	Achievement Level
Umpt, Stzn L.	8998577758	Opportunity #1 12/19/2014	2435 ±22	Level 3

Scale Score and Overall Performance

Umpt, Stzn L. Scored 2435 ±22

Level 4 The student has exceeded the achievement standard and demonstrates advanced progress toward mastery of the knowledge and skills in English language arts/literacy needed for likely success in future coursework.

Level 3 The student has met the achievement standard and demonstrates progress toward mastery of the knowledge and skills in English language arts/literacy needed for likely success in future coursework.





Level 2 The student has nearly met the achievement standard and may require further development to demonstrate the knowledge and skills in English language arts/literacy needed for likely success in future coursework.

Level 1 The student has not met the achievement standard and needs substantial improvement to demonstrate the knowledge and skills in English language arts/literacy needed for likely success in future coursework.

Comparison Scores

Name	Average Scale Score
Demo district 9999 (9999)	2435 ±22
Demo inst 9997 (9999_9997)	2435 ±22

Student Performance on Claims

Claim	Performance	Claim Description
Reading		Student has difficulty reading closely and analytically to comprehend a range of increasingly complex literary and informational texts.
Listening		Student may be able to employ effective listening skills for a range of purposes and audiences.
Writing		Student can produce effective and well-grounded writing for a range of purposes and audiences.
Research/Inquiry		Student can engage in research and inquiry to investigate topics, and to analyze, integrate, and present information.

- IAB tests** consist of 4–9 individually scored blocks per subject per grade. Students may be administered as many or as few blocks as necessary, and they may have multiple opportunities. IAB reports look different from Summative/ICA reports, because instead of a scale score and achievement level, students receive a Block Performance Level for each block tested (see Figure 7c below for an example of a Smarter Balanced IAB Individual Student Report).




Figure 7c: 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 6
 Year: 2014-2015
 Name: Dwoo, Rtjv







Legend: Blocks Performance Categories

 Below Standard
  At/Near Standard
  Above Standard

Student Information

Name	SSID	Opportunity
Dwoo, Rtjv 	9999886404	Opportunity #1 1/24/2015

Student Test Performance

Block	Performance Level
Read Literary Texts	
Read Informational Texts	
Edit/Revise	
Brief Writes	
Listen/Interpret	
Research	
Argument Performance Task	