Connecticut Next Generation Science Standards Assessment

2020–2021

Volume 6: Score Interpretation Guide

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. CONNECTICUT SCORE REPORTS

In spring 2021, the Connecticut Next Generation Science Standards (NGSS) Assessments were administered to Connecticut students in grades 5, 8, and 11 science. The purpose of the *Score Interpretation Guide* is to document the features of the Connecticut Centralized Reporting System (CRS), which is designed to assist stakeholders in reviewing and downloading the test results and understanding and appropriately using the results of the state assessments. Additionally, this volume describes the score types reported for the spring 2021 assessments, the appropriate uses of those score types, the inferences that can be drawn from them, and features of the score report.

1.1 OVERVIEW OF CONNECTICUT'S SCORE REPORTS

The Connecticut NGSS Assessments were first administered operationally statewide in spring 2019. Due to the Covid-19 pandemic, testing in spring 2020 was cancelled. Testing resumed in Spring 2021 and test scores from the spring 2021 assessment were provided to districts and schools through the CRS on June 25, 2021. The CRS provided information on student performance and aggregated summaries at several levels—district, school, and roster.

The CRS (https://ct.portal.cambiumast.com/) is a web-based application that provides Connecticut NGSS Assessment results at various levels. Test results are available for users based on their roles and the privileges they receive based on the authentication granted to them. There are four basic levels of user roles, including the state, district, school, and teacher levels. Each user is granted drill-down access to reports in the system based on his or her assigned role. This means that teachers can access data for their roster(s) of students only, schools can access data for the students in their school only, and districts can access data for all schools and students in their district only.

The following users have access to the system:

- **State Users.** Access to all data at the state, district, school, teacher, and student levels
- **District Administrator (DA) and District Test Coordinator (DC) Users.** Access to all data for their district and the schools and students in their district
- **School Coordinator (SC) Users.** Access to all data for their school and the students in their school
- **Teacher Examiner (TE) and Screener Administrator (SA) Users.** Access to all aggregate data for their rosters and the students within their rosters

Access to reports is password protected, and users can access data at and below their assigned level. For example, an SC user can access the school report of students for his or her school but not for another school.

1.2 OVERALL SCORES AND DISCIPLINE-LEVEL SCORES

Each student receives a single scale score for each subject tested if there is a valid score to report. A student's score is based only on the operational items on the assessment. A scale score is used to describe how well a student performed on a test and can be interpreted as an estimate of the

student's knowledge and skills measured. The scale score is transformed from a theta score, which is estimated based on mathematical models. Low scale scores can be interpreted as an indication that the student does not possess sufficient knowledge and skills measured by the test. Conversely, high scale scores can be interpreted as an indication that the student has proficient knowledge and skills measured by the test. Interpretation of scale scores is more meaningful when the scale scores are used along with performance levels and Performance-Level Descriptors (PLDs).

Based on the scale score, students will receive an overall performance level. Performance levels are proficiency categories on a test, which students fall into based on their scale scores. For the Connecticut NGSS Assessment, scale scores are mapped into four performance levels:

1. Does Not Meet Standard
2. Approaching Standard
3. Meets Standard
4. Exceeds Standard

For details on the standard-setting process, refer to Volume 3, Standard Setting.

PLDs are a description of content area, knowledge, and skills that students at each performance level are expected to possess. Thus, performance levels can be interpreted based on PLDs. Generally, students performing on the Connecticut NGSS Assessment at Levels 3 and 4 are considered on track to demonstrate progress toward mastery of the knowledge and skills necessary for college and career readiness.

In addition to an overall score, students will receive discipline-level scores. For the Connecticut NGSS Assessment, student performance on each discipline level is reported on three performance categories:

1. Below Standard
2. Approaching Standard
3. Above Standard

Unlike the performance levels for the overall test, student performance on each of the discipline levels is evaluated with respect to the *Meets Standard* performance standard. Student performance at either *Below Standard* or *Above Standard* can be interpreted as student performance clearly below or above the *Meets Standard* cut score for a specific discipline. Student performance at *Approaching Standard* can be interpreted as student performance that does not provide enough information to tell whether students reached the *Meets Standard* mark for the specific discipline.

Table 1 displays the disciplines and discipline level claims for science, by grade.

Table 1. Disciplines and Discipline Level Claims for Science

Grade	Discipline	Claim
5, 8, 11	Practices and Concepts in Life Sciences	The student is able to use the science and engineering practices to demonstrate understanding of the disciplinary core ideas and crosscutting concepts in Life Science.
	Practices and Concepts in Physical Sciences	The student is able to use the science and engineering practices to demonstrate understanding of the disciplinary core ideas and crosscutting concepts in Physical Science.
	Practices and Concepts in Earth and Space Sciences	The student is able to use the science and engineering practices to demonstrate understanding of the disciplinary core ideas and crosscutting concepts in Earth and Space Science.

1.3 CENTRALIZED REPORTING SYSTEM

The Centralized Reporting System (CRS) generates a set of online score reports that describe student performance for students, families, educators, and other stakeholders. The online score reports are produced after the tests are submitted by the students, handscored and machine-scored, and finally processed into the CRS. In addition to each individual student's score report, the CRS produces aggregate score reports for teachers, schools, districts, and states.

Furthermore, to facilitate comparisons, each aggregate report contains the summary results for the selected aggregate unit, as well as all aggregate units above the selected aggregate. For example, if a school is selected, the summary results of the district to which the school belongs are provided so that the school performance can be compared with the district performance. If a teacher is selected, the summary results for the school and the district above the teacher are also provided for comparison purposes.

1.4 AVAILABLE REPORTS ON THE CONNECTICUT CENTRALIZED REPORTING SYSTEM

The Connecticut CRS is hierarchically structured. An authorized user is able to view reports at his or her own aggregated unit and any lower level of aggregation. For example, a school user can view only the reports and data at the school and student levels of his or her school. DA users can view the reports and data for their districts and also the student-level results for all of their schools.

Table 2 summarizes the types of score reports that are available in the CRS and the levels at which the reports can be viewed. A description of each report is also provided. Data files are also accessible for district to download.

For detailed information on available reports and available features, educators can refer to the CRS user guide. The 2020–2021 *Centralized Reporting System User Guide* is included in Appendix A, Centralized Reporting System User Guide.

Table 2. Connecticut Score Reports Summary

Report	Description	Unit of Aggregation				
		State	District	School	Roster	Student
Summary Performance	Summary of performance (to date) across grades and subjects or courses for the current administration	✓	✓	✓	✓	
Aggregate-Level Subject Report	Summary of overall performance for a subject and a grade for all students in the defined level of aggregation	✓	✓	✓	✓	
Aggregate-Level Discipline Level Score Report	Summary of overall performance on each discipline level for each grade across all students within the selected level of aggregation	✓	✓	✓	✓	
Aggregate-Level Disciplinary Core Ideas Report	Summary of overall performance on each disciplinary core idea for a given subject and grade across all students within the selected level of aggregation	✓	✓	✓	✓	
Student-Level Subject Report	List of all students who belong to a school, teacher, or roster with their associated subject or course scores for the current administration			✓	✓	✓
Student-Level Discipline-Level Score Report	List of all students who belong to a school, teacher, or roster with their associated discipline level performance for the current administration			✓	✓	✓
Individual Student Report (ISR)	Detailed information about a selected student's performance in a specified subject or course, includes overall subject and discipline level results					✓
Data Files	Text/CSV file containing overall and discipline-level scale scores and performance levels along with demographic information		✓	✓	✓	✓

1.4.1 Reporting by Subgroup

The aggregate score reports provide overall student results by default but can at any time be analyzed by subgroups based on demographic data. When used on aggregate-level reports, an additional level of analysis will be provided by aggregating students based on subgroup. For example, when the Gender subgroup is selected, the CRS will display aggregate results for all students, male and female. When used on student-level reports, subgroups can instead be used to filter individual results. For example, a user will have the option to select Male or Female after the Gender subgroup is selected.

Users can see student assessment results by any subgroup at any time by selecting the desired subgroup from the *Breakdown By* drop-down menu available. Table 3 presents the types of subgroups and subgroup categories provided in the CRS.

Table 3. Connecticut List of Subgroups

Breakdown by Category	Displayed Category
Ethnicity	Two or More Races
	American Indian or Alaskan Native
	Asian
	Hispanic or Latino
	Black or African American
	White
	Pacific Islander
Gender	Male
	Female
IDEA Indicator	Yes
	No
Limited English Proficiency Status	Yes
	No
Enrolled Grade	Grade 5
	Grade 8
	Grade 11

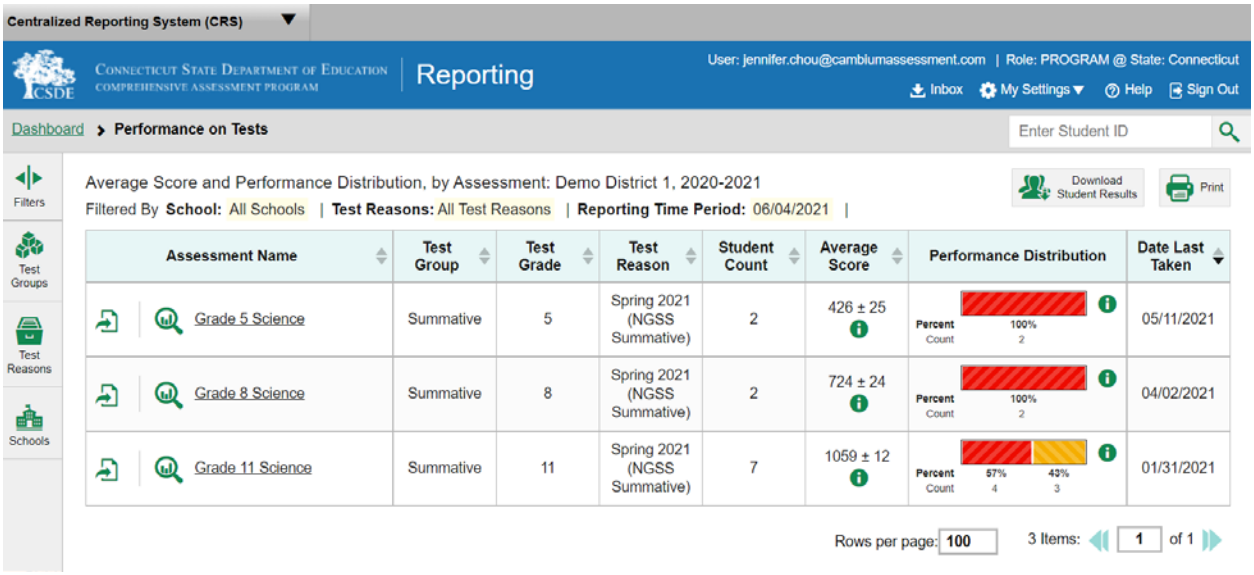
1.4.2 Summary Performance Report

Homepage-authorized users can log on to the CRS dashboard to view summaries of students' performance across grades and subjects. State personnel and district personnel can access district summaries, school personnel can access school summaries, and teachers can access summaries of their students through the dashboard. The Summary Performance Distribution has the following features:

- Displays summary data separated by grade and subject
- Level of aggregation based on a user's role
- Reports number of students tested and percentage meeting standard

Figure 1 presents sample Summary Performance Reports at the district level.

Figure 1. District-Level Summary Performance Report



1.4.3 Aggregate-Level Subject Report

Detailed summaries of student performance within a grade and subject area are made available in the Aggregate-Level Subject Report. The Aggregate-Level Subject Report presents results for the aggregate unit, as well as the results for any higher-level aggregate units. For example, a school's Aggregate-Level Subject Report will contain the summary results of the school's district so that school performance can be compared with district performance.

The Aggregate-Level Subject Report provides the aggregate summaries on a specific subject area, including

- number of students;
- average scale score and standard error of the average scale score;
- percentage of students meeting standard; and
- percentage of students in each performance level.

The summaries are also presented for overall students and by subgroups. Figure 2 presents an example of Aggregate-Level Subject Reports for grade 11 science at the district level without subgroups. Figure 3 highlights grade 11 science at the district level when a user selects a subgroup of gender.

Figure 2. District-Aggregate-Level Subject Report for Grade 11 Science

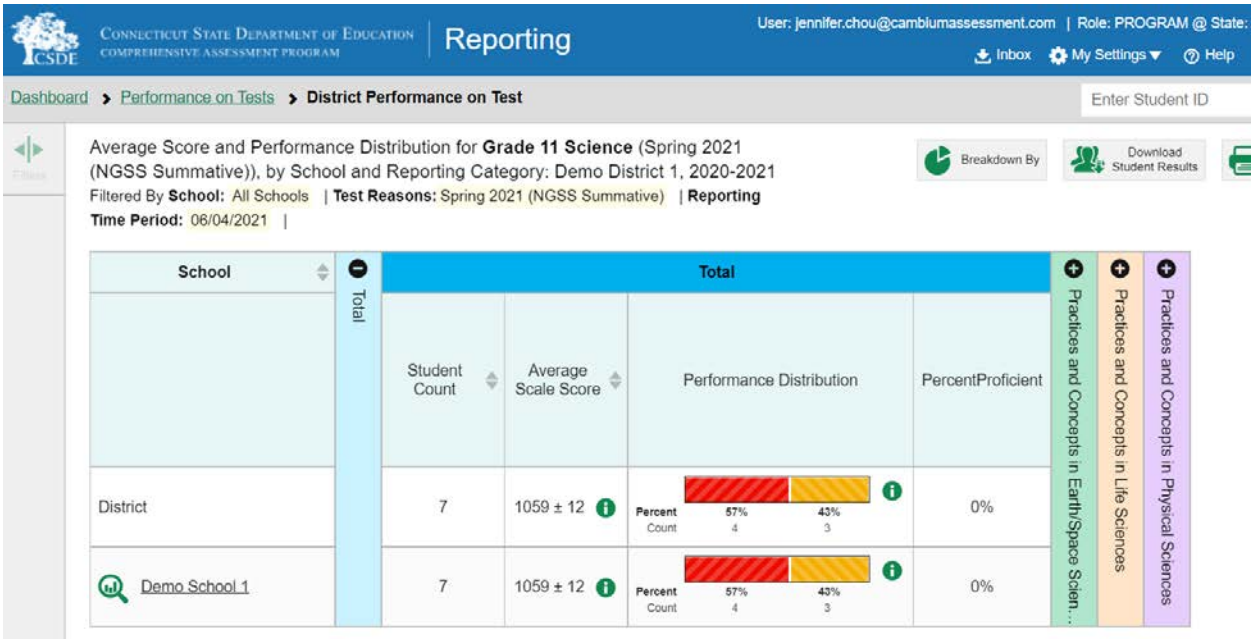
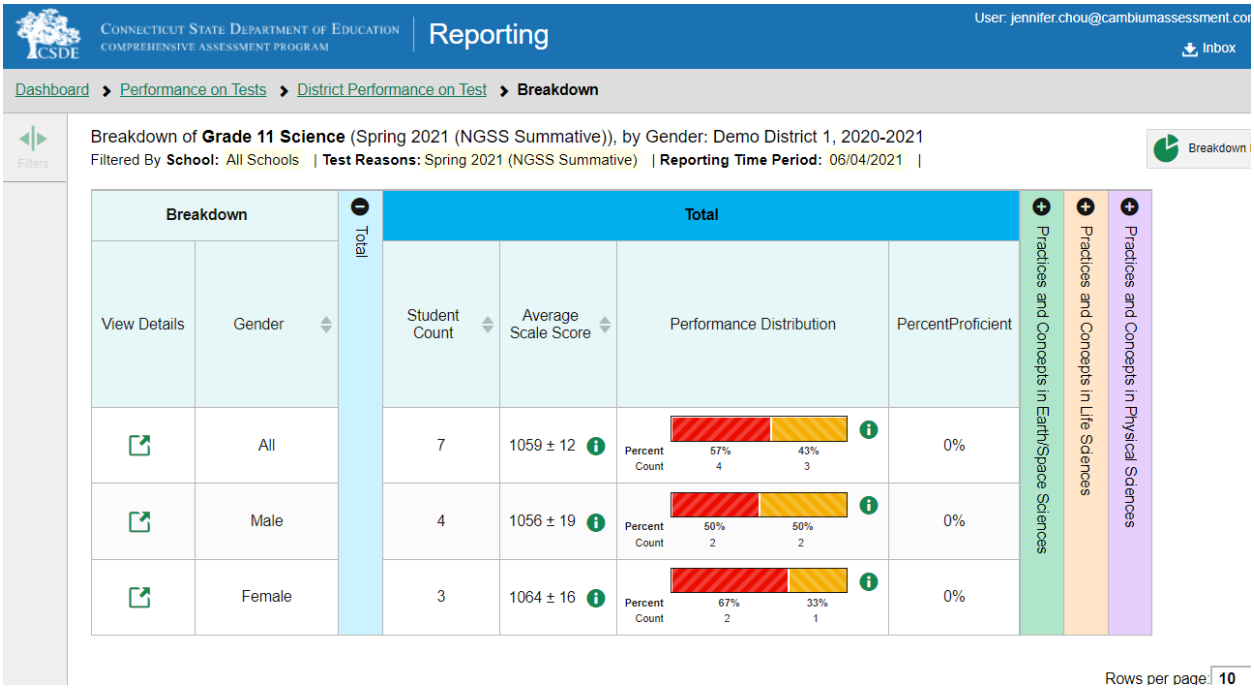


Figure 3. District Aggregate-Level Subject Report for Grade 11 Science by Gender



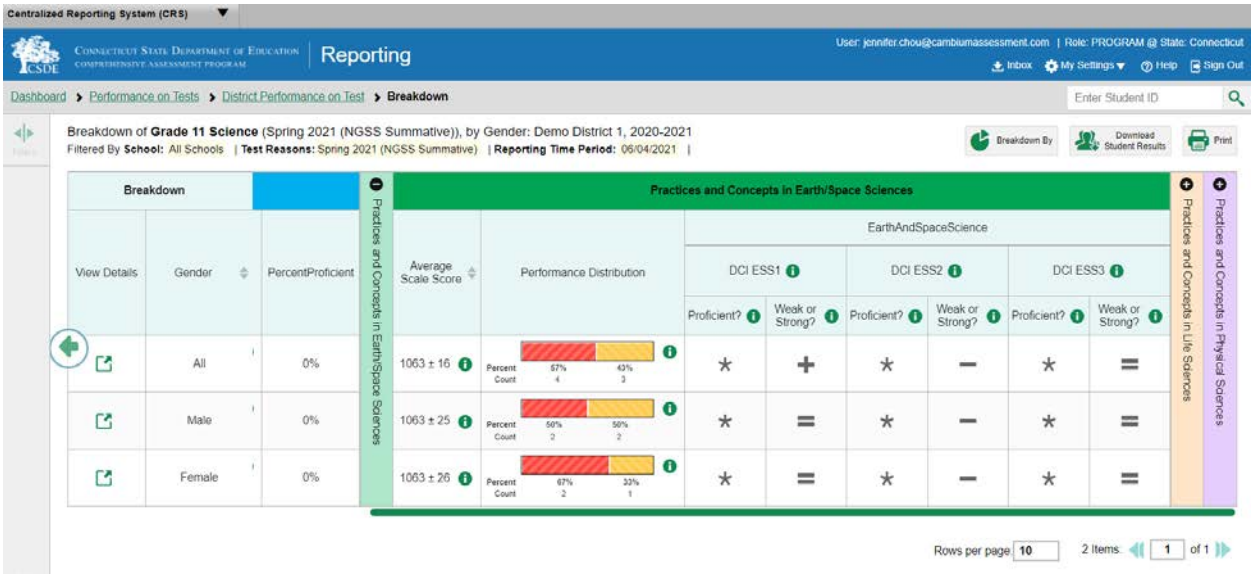
1.4.4 Aggregate-Level Discipline-Level Report

The Aggregate-Level Discipline-Level Report provides the aggregate summaries on student performance in each discipline level for each grade. The summaries on the Aggregate-Level Discipline-Level Report include

- number of students;
- average scale score and standard error of the average scale score;
- percentage of students meeting standard; and
- percentage of students in each performance category for each of the disciplines.

Similar to the Aggregate-Level Subject Report, the Aggregate-Level Discipline-Level Report presents the summary results for the selected aggregate unit, as well as the summary results for the aggregate unit above the selected aggregate. In addition, summaries can be presented for all students within an aggregate and by students within a defined subgroup. Figure 4 presents an example of the District Aggregate-Level Discipline-Level Report for grade 11 science. Reports by subgroups are also available for the Aggregate-Level Discipline-Level Report, similar to Figure 3.

Figure 4. District Aggregate-Level Discipline-Level Report for Grade 11 Science



1.4.5 Aggregate-Level Items and Disciplinary Core Ideas Level Report

The Aggregate-Level Disciplinary Core Ideas Level Report lists data on the performance of student groups on each standard of a subject for the current window and reports Areas Where Performance Indicates Proficiency and Areas of Strongest and Weakest Performance. For Areas Where Performance Indicates Proficiency, a performance indicator produces information on how a group of students in a class, school, or district performed on the standard compared to the proficiency cuts. It shows whether performance on this standard for this group was above, no different than, or below what is expected of students at the proficient level. This indicator shows strengths and weaknesses for a group of students and is provided only at an aggregate level since it is unstable at the individual level. For Areas of Strongest and Weakest Performance, the expected performance is determined based on the students' overall performance on the entire test.

Figure 5 demonstrates examples of the Aggregate-Level Disciplinary Core Ideas Report for grade 8 science.

Figure 5. District Aggregate-Level Disciplinary Core Idea Report for Grade 11 Science

Practices and Concepts in Earth/Space Sciences								
Average Scale Score	Performance Distribution	EarthAndSpaceScience						
		DCI ESS1		DCI ESS2		DCI ESS3		
		Proficient?	Weak or Strong?	Proficient?	Weak or Strong?	Proficient?	Weak or Strong?	
1063 ± 16	Percent 57% 43% Count 4 3	*	+	*	−	*	=	
1063 ± 25	Percent 50% 50% Count 2 2	*	=	*	−	*	=	
1063 ± 26	Percent 67% 33% Count 2 1	*	=	*	−	*	=	

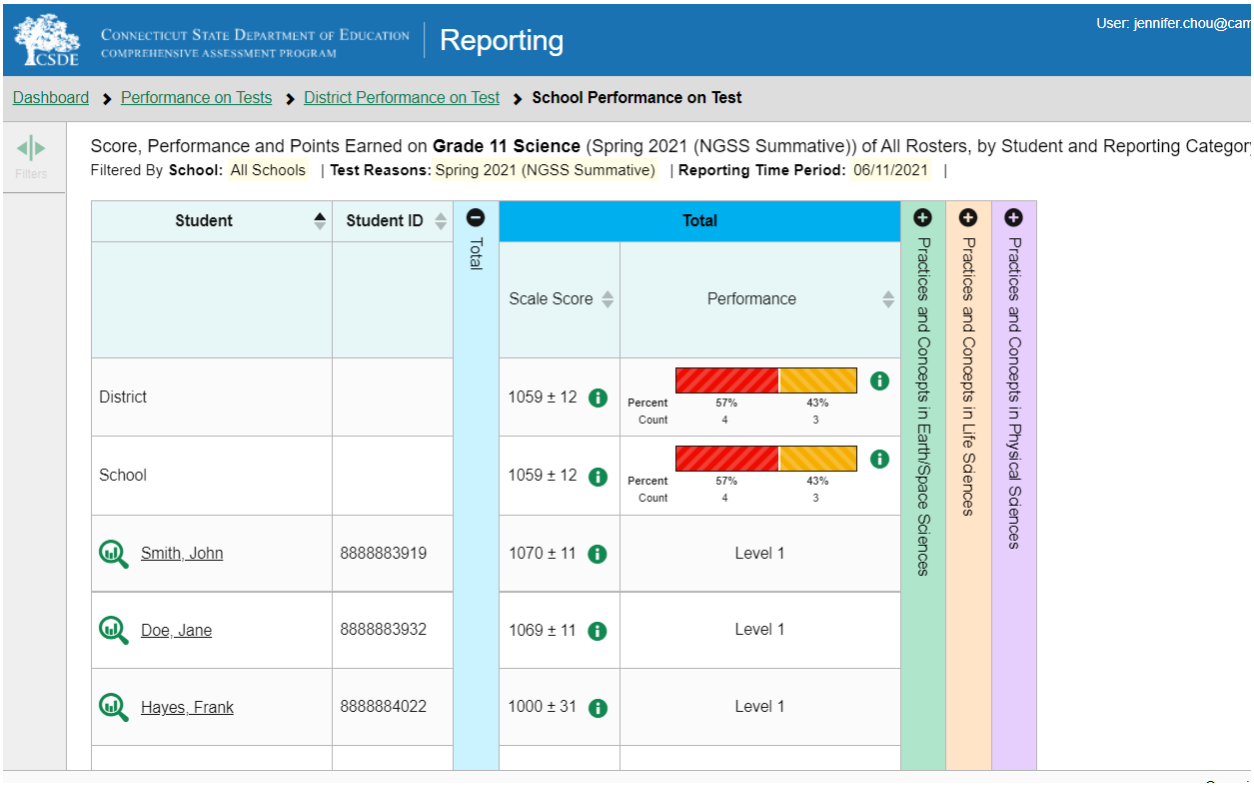
1.4.6 Student-Level Subject Report

The Student-Level Subject Report lists all students who belong to the selected aggregate level, such as a school, and reports the following measures for each student:

- Scale score
- Overall subject performance level

Figure 6 demonstrates an example of the Student-Level Subject Report for grade 11 science.

Figure 6. Student Roster Subject Report for Grade 11 Science



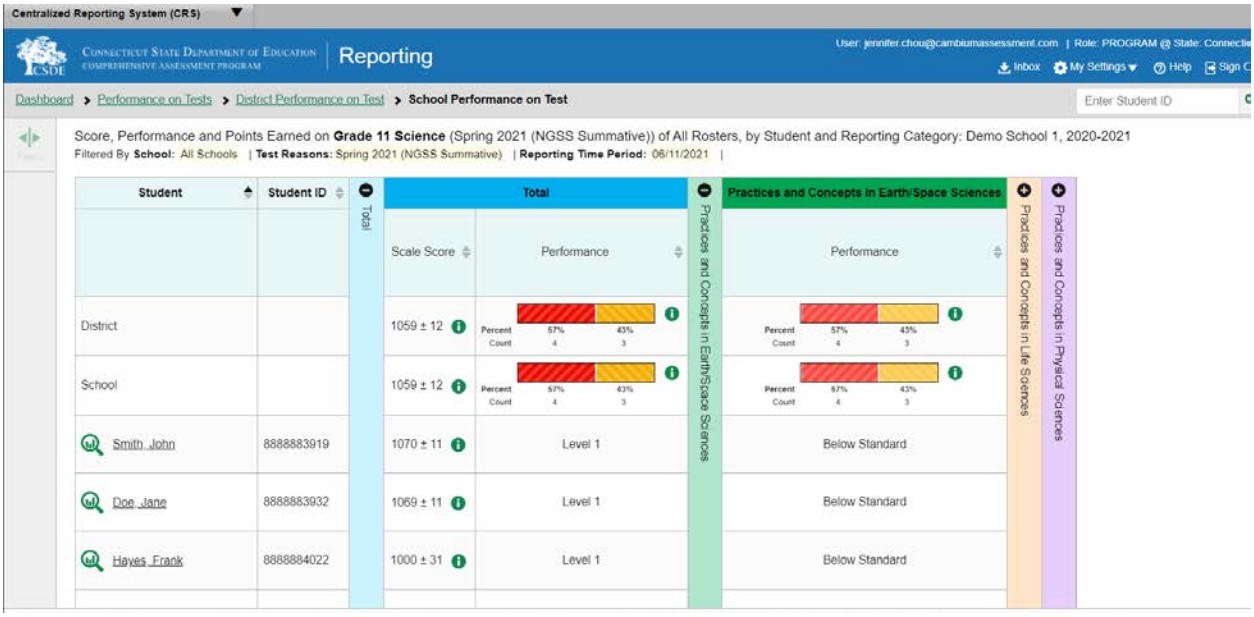
1.4.7 Student-Level Discipline-Level Report

The Student-Level Discipline-Level Report lists all students who belong to the selected aggregate level, such as a school, and reports the following measures for each student:

- Scale score
- Overall subject performance level
- Discipline performance category (i.e., Earth and Space Sciences, Life Sciences, and Physical Sciences)

Figure 7 presents an example of the Student-Level Discipline Report for grade 11 science.

Figure 7. Student Roster Discipline Report for Grade 11 Science



1.4.8 Individual Student Report

When a student receives a valid test score, an Individual Student Report (ISR) can be generated in the Centralized Reporting System (CRS). The ISR contains the following measures:

- Scale score and standard error of measurement (SEM)
- Overall subject performance level
- Average scale scores for student's district and school
- Performance category in each discipline (science)

On the top of the report, information includes:

- Student's name
- Scale score with SEM
- Performance level

In the middle section of the report, information includes:

- Barrel chart with student's scale score and SEM (using a sign of "±")
- PLDs with cut scores at each performance level
- Average scale scores and standard errors for district and school aggregation levels

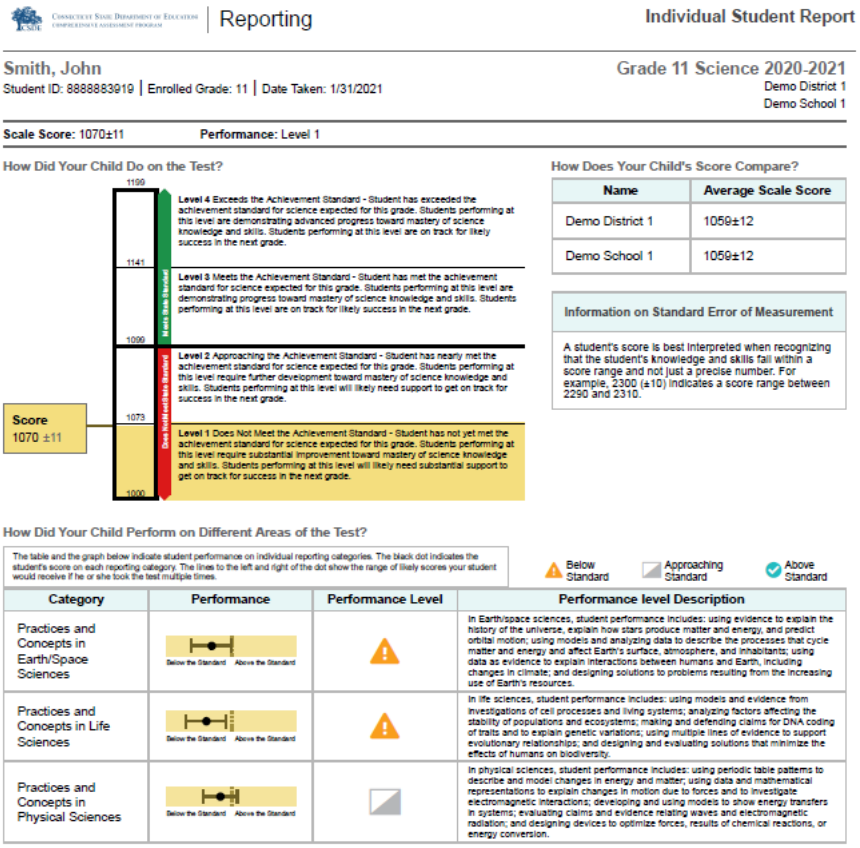
- Note: the "±" next to the student's scale score is the standard error of measurement of the scale score, whereas the "±" next to the average scale scores for aggregate levels represents the standard error of the average scale scores.

On the bottom of the report, information includes:

- Detailed information on student performance on each discipline level
  - Note: Bar charts in the Discipline table show how students performed on each discipline (black bar), relative to the discipline-level performance standard (dashed white line). Green boxes show the score range the student would likely fall within if they took the test multiple times.

Figure 8 presents an example ISR for grade 11 science. An example of the printed ISRs is displayed in Appendix B, Sample Printed Individual Student Report.

Figure 8. Individual Student Report for Grade 11 Science



Reporting

Individual Student Report

Smith, John

Student ID: 8888883919 | Enrolled Grade: 11 | Date Taken: 1/31/2021

Grade 11 Science 2020-2021

Demo District 1

Demo School 1

Scale Score: 1070±11 **Performance:** Level 1

How Did Your Child Do on the Test?

1199

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

1141

Level 3 Meets the Achievement Standard - Student has met the achievement standard for science expected for this grade. Students performing at this level are demonstrating progress toward mastery of science knowledge and skills. Students performing at this level are on track for likely success in the next grade.

Meets the Standard

1099

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

1073

Score 1070 ±11

Does Not Meet the Standard

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

1000

How Does Your Child's Score Compare?

Name	Average Scale Score
Demo District 1	1059±12
Demo School 1	1059±12

Information on Standard Error of Measurement

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.

How Did Your Child Perform on Different Areas of the Test?

The table and the graph below indicate student performance on individual reporting categories. The black dot indicates the student's score on each reporting category. The lines to the left and right of the dot show the range of likely scores your student would receive if he or she took the test multiple times.

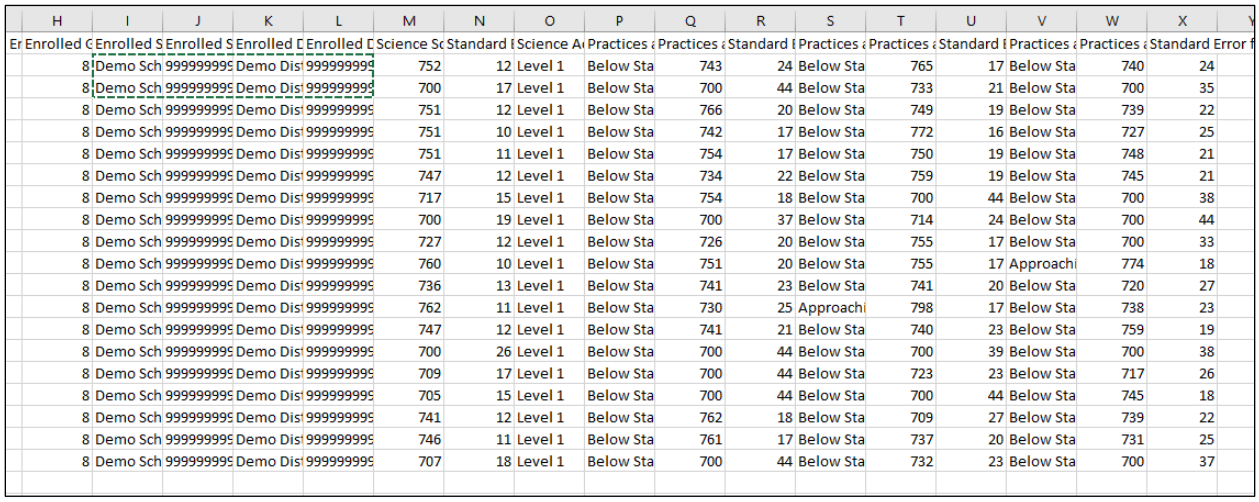
Below Standard | Approaching Standard | Above Standard

Category	Performance	Performance Level	Performance level Description
Practices and Concepts in Earth/Space Sciences	Below the Standard Above the Standard	⚠	In Earth/space sciences, student performance includes: using evidence to explain the history of the universe, explain how stars produce matter and energy, and predict orbital motion; using models and analyzing data to describe the processes that cycle matter and energy and affect Earth's surface, atmosphere, and inhabitants; using data as evidence to explain interactions between humans and Earth, including changes in climate; and designing solutions to problems resulting from the increasing use of Earth's resources.
Practices and Concepts in Life Sciences	Below the Standard Above the Standard	⚠	In life sciences, student performance includes: using models and evidence from investigations of cell processes and living systems; analyzing factors affecting the stability of populations and ecosystems; making and defending claims for DNA coding of traits and to explain genetic variations; using multiple lines of evidence to support evolutionary relationships; and designing and evaluating solutions that minimize the effects of humans on biodiversity.
Practices and Concepts in Physical Sciences	Below the Standard Above the Standard	◪	In physical sciences, student performance includes: using periodic table patterns to describe and model changes in energy and matter; using data and mathematical representations to explain changes in motion due to forces and to investigate electromagnetic interactions; developing and using models to show energy transfers in systems; evaluating claims and evidence relating waves and electromagnetic radiation; and designing devices to optimize forces, results of chemical reactions, or energy conversion.

1.4.9 Data File

CRS users have the option to quickly generate a comprehensive data file of their students' scores. Data files (See Figure 9) can be downloaded in Microsoft Excel or CSV format and contain a wide variety of data, which includes scale scores, reporting discipline scores, demographic data, and performance levels. Data files can be useful as a resource for further analysis. Data files can be generated at the district, school, teacher, or roster level.

Figure 9. Data File



H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Er	Enrolled C	Enrolled S	Enrolled S	Enrolled D	Enrolled D	Science Sc	Standard E	Science Ac	Practices a	Practices a	Standard E	Practices a	Practices a	Standard E	Practices a	Practices a	Standard Error f
	8	Demo Sch	999999995	Demo Dist	999999995	752	12	Level 1	Below Sta	743	24	Below Sta	765	17	Below Sta	740	24	
	8	Demo Sch	999999995	Demo Dist	999999995	700	17	Level 1	Below Sta	700	44	Below Sta	733	21	Below Sta	700	35	
	8	Demo Sch	999999995	Demo Dist	999999995	751	12	Level 1	Below Sta	766	20	Below Sta	749	19	Below Sta	739	22	
	8	Demo Sch	999999995	Demo Dist	999999995	751	10	Level 1	Below Sta	742	17	Below Sta	772	16	Below Sta	727	25	
	8	Demo Sch	999999995	Demo Dist	999999995	751	11	Level 1	Below Sta	754	17	Below Sta	750	19	Below Sta	748	21	
	8	Demo Sch	999999995	Demo Dist	999999995	747	12	Level 1	Below Sta	734	22	Below Sta	759	19	Below Sta	745	21	
	8	Demo Sch	999999995	Demo Dist	999999995	717	15	Level 1	Below Sta	754	18	Below Sta	700	44	Below Sta	700	38	
	8	Demo Sch	999999995	Demo Dist	999999995	700	19	Level 1	Below Sta	700	37	Below Sta	714	24	Below Sta	700	44	
	8	Demo Sch	999999995	Demo Dist	999999995	727	12	Level 1	Below Sta	726	20	Below Sta	755	17	Below Sta	700	33	
	8	Demo Sch	999999995	Demo Dist	999999995	760	10	Level 1	Below Sta	751	20	Below Sta	755	17	Approachi	774	18	
	8	Demo Sch	999999995	Demo Dist	999999995	736	13	Level 1	Below Sta	741	23	Below Sta	741	20	Below Sta	720	27	
	8	Demo Sch	999999995	Demo Dist	999999995	762	11	Level 1	Below Sta	730	25	Approachi	798	17	Below Sta	738	23	
	8	Demo Sch	999999995	Demo Dist	999999995	747	12	Level 1	Below Sta	741	21	Below Sta	740	23	Below Sta	759	19	
	8	Demo Sch	999999995	Demo Dist	999999995	700	26	Level 1	Below Sta	700	44	Below Sta	700	39	Below Sta	700	38	
	8	Demo Sch	999999995	Demo Dist	999999995	709	17	Level 1	Below Sta	700	44	Below Sta	723	23	Below Sta	717	26	
	8	Demo Sch	999999995	Demo Dist	999999995	705	15	Level 1	Below Sta	700	44	Below Sta	700	44	Below Sta	745	18	
	8	Demo Sch	999999995	Demo Dist	999999995	741	12	Level 1	Below Sta	762	18	Below Sta	709	27	Below Sta	739	22	
	8	Demo Sch	999999995	Demo Dist	999999995	746	11	Level 1	Below Sta	761	17	Below Sta	737	20	Below Sta	731	25	
	8	Demo Sch	999999995	Demo Dist	999999995	707	18	Level 1	Below Sta	700	44	Below Sta	732	23	Below Sta	700	37	

1.5 TEST INFORMATION DISTRIBUTION ENGINE

Test Completion Rate Reports were available on the Test Information Distribution Engine (TIDE) website (https://ct.tide.cambiumassessment.com). These reports indicate the students who completed or need to complete computer-based testing and allow users to view participation summary statistics (counts and percentages) of students who have tested.

Once a user logs in, he or she is directed to the homepage, which allows users to access the Test Completion Rate Reports.

The Test Completion Rate Report allows teachers, principals, and district staff to see which students have not yet completed their tests. Users can select from a series of options to customize the group of students whose participation status is to be reviewed for a particular grade and subject, such as those who started but have not completed their test or those who have not yet begun their test. Users can export the list into a Microsoft Excel file and download the file.

1.6 PAPER INDIVIDUAL STUDENT REPORTS FOR FAMILIES

Individual Student Reports (ISRs) were delivered as printed materials to the districts where students were enrolled as of June 4, 2021, at 11:59:59 p.m. The primary purpose of the ISR was to provide a document that enabled families to understand their child's performance in the subject in which he or she tested. The ISR also presented information that indicated how a student's performance compared to that of other students who took the same test. The report is organized as follows:

- **Top of Report.** The student's name, student ID, test grade, test date, school, and district are identified. Science reports included a frequently asked questions section here.
- **Connecticut NGSS Assessment Scores**. The student's scale score and corresponding performance level are displayed graphically and explained in accompanying text. A range of scores that is ±SEM is given with explanatory text.
- **Student Performance Compared.** Included with the Connecticut NGSS Assessment scores graphic, this section provides a comparison between the student's scale score and that of the student's school and district.
- **Discipline Level Scores.** Discipline Level tables show how students performed on each discipline level. This section includes graphical displays of the Performance Category (*Below Standard, Approaching Standard, or Above Standard*), or relative strength/weakness, for each of the discipline levels assessed. These results are explained in greater detail next to the graphics, including "Next Steps" that families and students may take to improve student performance.

An example of the printed ISRs is displayed in Appendix B, Sample Printed Individual Student Report.

2. INTERPRETATION OF REPORTED SCORES

A student's performance on a test is reported as a scale score and a performance level for the overall test and as a performance level for each discipline level. Students' scores and performance levels are summarized at the aggregate levels. This section describes how to interpret these scores.

2.1 SCALE SCORE

A scale score is used to describe how well a student performed on a test and can be interpreted as an estimate of a student's knowledge and skills as measured by their performance on the test. A scale score is the student's overall numeric score. These scores fall on a continuous scale. The Connecticut NGSS Assessment scale scores are not expressed on a vertical scale, which means that scores from different grades cannot be compared.

Scale scores can be used to illustrate a student's current levels of performance. Low scale scores indicate that a student does not possess sufficient knowledge and skills measured by the test. Conversely, high scale scores indicate that a student has proficient knowledge and skills measured by the test. When combined across a student population, scale scores can also describe school- and district-level changes in performance and reveal gaps in performance among different groups of students. In addition, scale scores can be averaged across groups of students, allowing educators to use group comparison. Interpretation of scale scores is more meaningful when the scale scores are used along with performance levels and PLDs. It should be noted that the utility of scale scores is limited when comparing smaller differences among scores (or averaged group scores), particularly when the difference among scores is within the SEM. Furthermore, the scale score of individual students should be cautiously interpreted when comparing two scale scores, because small differences in scores may not reflect real differences in performance.

2.2 STANDARD ERROR OF MEASUREMENT

A student's score is best interpreted when recognizing that the student's knowledge and skills fall within a score range and are not just precise numbers. A scale score (the observed score on any test) is an estimate of the true score. If a student takes a similar test several times, the resulting scale scores will vary across administrations; sometimes the scores will be a little higher, a little lower, or the same. The SEM represents the precision of the scale score, or the range in which the student would likely score if a similar test were administered several times. The SEM can be interpreted as the degree of uncertainty of a student's score based on a statistical analysis of the student's answers on a test. When interpreting scale scores, it is recommended to always consider the range of scale scores incorporating the SEM of the scale score.

The "±" next to a student's scale score provides information about the certainty, or confidence, of the score's interpretation. The boundaries of the score band are one SEM above and below the student's observed scale score, representing a range of score values that is likely to contain the true score. For example, "680 ± 10" indicates that if a student were tested again, it is likely that he or she would receive a score between 670 and 690.

2.3 PERFORMANCE LEVEL

Performance levels are proficiency categories on a test, which students fall into based on their scale scores. For the Connecticut NGSS Assessment, scale scores are mapped into four performance levels (*Does Not Meet Standard, Approaching Standard, Meets Standard, Exceeds Standard*) using performance standards (see Section 2.5, Cut Scores). PLDs are a description of content-area knowledge and skills that students at each performance level are expected to possess. Thus, performance levels can be interpreted based on PLDs. Students performing on the Connecticut NGSS Assessment at *Meets Standard* and *Exceeds Standard* are considered on track to demonstrate progress toward mastery of the knowledge and skills necessary for college and career readiness.

2.4 PERFORMANCE CATEGORY FOR DISCIPLINE LEVELS

Students' performance on each reporting discipline is reported for three performance categories: *Below Standard, Approaching Standard, and Above Standard*. Unlike the performance levels for the overall test, student performance on each of the discipline levels is evaluated with respect to the *Meets Standard* performance standard. Students performing at either *Below Standard* or *Above Standard* can be interpreted as having student performance that is clearly below or above the *Meets Standard* cut score for a specific discipline level. Students performing at *Approaching Standard* can be interpreted as having student performance that does not provide enough information to tell whether students reached the *Meets Standard* mark for the specific discipline level.

2.5 CUT SCORES

For all grades in the Connecticut NGSS Assessment, scale scores are mapped onto four performance levels (*Does Not Meet Standard, Approaching Standard, Meets Standard, Exceeds Standard*). For each performance level, there is a minimum and a maximum scale score that define the range of scale scores students in each performance level have achieved. Collectively, these

minimum and maximum scale scores are defined as cut scores and are the cut-off points for each performance level. Table 4 presents the cut scores for science for all grades.

Table 4. Connecticut NGSS Assessment Science Performance-Level Cut Scores

Grade	Does Not Meet Standard	Approaching Standard	Meets Standard	Exceeds Standard
5	400–467	468–497	498–534	535–599
8	700–771	772–797	798–841	842–899
11	1000–1072	1073–1098	1099–1140	1141–1199

2.6 AGGREGATED SCORES

Students' scale scores are aggregated at roster, teacher, school, and district levels to represent how a group of students perform on a test. When students' scale scores are aggregated, the aggregated scale scores can be interpreted as an estimate of knowledge and skills that a group of students possesses. This interpretation makes aggregated scores a powerful tool when comparing student performance across different groups of students, whether it be at a similar level of aggregation (e.g., school to school) or an analysis of a subgroup (e.g., comparing a teacher's roster to the overall school).

Given that student scale scores are estimates, the aggregated scale scores are also estimates and are subject to measures of uncertainty, as expressed using the calculated SEM for an aggregate average scale score. In addition to the aggregated scale scores, the percentage of students in each performance level is reported at the aggregate level to represent how well a group of students performs overall and by discipline level.

2.7 RELATIVE STRENGTHS AND WEAKNESSES FOR DISCIPLINARY CORE IDEAS

For Disciplinary Core Idea performance, relative strengths and weaknesses at each standard are reported for aggregate levels only (e.g., classroom, school, or district). Because an individual student responds to too few items within a standard to generate reliable data, the standard performance is produced by aggregating all items within a standard across students at an aggregate level.

The "Areas Where Performance Indicates Proficiency" for a standard shows how a group of students performed in each standard relative to the expected performance for proficiency. For summative tests, this is the expected level of performance necessary to achieve *Meets Standard* performance. This is a standards-based report with the group performance in each standard being compared to performance standard for that standard. Similar to the performance levels provided for the total test, this is an indication of students' performance in the standard with respect to the standards.

Since the "Areas Where Performance Indicates Proficiency" data for each standard is a comparison to the standards-based expectations, performance across groups can be compared.

For "Areas of Strongest and Weakest Performance," the expected performance is determined based on the students' overall performance on the entire test. It shows how a group of students performed in each standard relative to their performance on the test overall. Rather than comparing across groups, "Areas of Strongest and Weakest Performance" provides more information regarding the relative strength and weakness on different standards on the test within a group.

2.8 APPROPRIATE USES FOR SCORES AND REPORTS

Assessment results can be used to provide information on individual student performance on the test. Overall, assessment results tell what a student knows and is able to do in certain subject areas and gives further information on whether a student is on track to demonstrate the knowledge and skills necessary for college and career readiness. Additionally, assessment results can be used to identify a student's relative strengths and weaknesses in certain content areas. For example, performance categories for reporting disciplines can be used to identify an individual student's relative strengths and weaknesses among reporting categories within a content area.

Assessment results on student performance on the test can be used to help teachers or schools make decisions on how to support students' learning. Aggregate score reports at the teacher and school level provide information about the strengths and weaknesses of students and can be used to improve teaching and student learning. For example, a group of students may have performed very well overall, but possibly did not perform as well in several standards compared to their overall performance. In this case, teachers or schools can identify strengths and weaknesses of their students through the group performance by standards and promote instruction on specific areas where student performance is below their overall performance. Further, by narrowing down the student performance result by subgroup, teachers and schools can determine what strategies may be needed to improve teaching and student learning, particularly for students from disadvantaged subgroups. For example, teachers might see student assessment results by gender and observe that a particular group of students is struggling with Physical Sciences. Teachers can then provide additional instructions that focus on the Physical Sciences for these students.

In addition, assessment results can be used to compare student performance among different students and among different groups. Teachers can evaluate how their students perform compared with other students in schools and districts for overall scores and by discipline level. Although all students are administered different sets of items under the linear-on-the-fly test design, scale scores are comparable across students.

While assessment results provide valuable information to understand student performance, these scores and reports should be used with caution. It is important to note that scale scores are estimates of true scores and hence do not represent the precise measure for student performance. A student's scale score is associated with measurement error, and thus users need to consider measurement error when using student scores to make decisions about student performance. Moreover, although student scores may be used to help make important decisions about student placement and retention or teachers' instructional planning and implementation, the assessment results should not be used as the only source of information. Given that assessment results measured by a test provide limited information, other sources on student performance, such as classroom assessment and teacher evaluation, should be considered when making decisions on student learning. Finally, when student performance is compared across groups, users need to take into account the group

size. The smaller the group, the larger the measurement error related to these aggregate data, thus requiring a more cautious interpretation.

3. SUMMARY

Connecticut NGSS Assessment results are reported online via the Centralized Reporting System (CRS), as well as through printed Individual Student Reports (ISRs) sent to families. The results were released after the testing window closed and standard setting was completed.

The CRS is interactive. When educators or administrators log in, they see a summary of data about students for whom they are responsible (e.g., a principal will see the students in his or her school; a teacher will see students in his or her class). Users can then drill down through various levels of aggregation all the way to individual reports. The system allows them to tailor the content more precisely, moving from subject area through reporting categories, and even to standards-level reports for aggregates. Aggregate reports are available at every level, and authorized users can print these reports or download them (or the data on which they are based). ISRs can be produced individually or batched as PDF reports.

All authorized users can download files, including data about students for whom they are responsible, at any time. The various reports available may be used to inform stakeholders regarding student performance and instructional strategies.