State of Delaware

Delaware Comprehensive Assessment System 2010–2011

Volume 6

Score Interpretation Guide

American Institutes for Research

American Institutes for Research

TABLE OF CONTENTS

1. DELAWARE SCORE REPORTS

The State of Delaware implemented a new online assessment for operational use during the 2010–2011 school year. This new test, referred to as the Delaware Comprehensive Assessment System (DCAS), replaced the paper-and-pencil test, the Delaware State Testing Program (DSTP). In 2010–2011, students who were enrolled in various grades in public and charter schools were required to take the online assessment. The paper-and-pencil version was available as an accommodation for students with special needs.

Students participate in the DCAS test once in the fall, once in the winter, and finally in the spring. Only the scores from the spring test window are used for accountability purposes. Because pre-equating is used for the online test, test scores can be provided to students immediately upon completion of a test. Teachers also have access to multiple score types for each student in an easy-to-access electronic data warehouse including total scores, performance levels, and scores at the reporting category level.

Volume 6, Score Interpretation Guide, documents the features of the DCAS reporting system that are designed to help stakeholders understand and appropriately use the results of the state assessments.

1.1 OVERVIEW OF DELAWARE'S SCORE REPORTS

DCAS Reading and Mathematics are administered three times a year—fall, winter, and spring. DCAS Science and Social Studies are given once a year in the spring. Test scores from each online assessment are provided to students immediately upon their completion of the test through the DCAS Online Reporting System (DCAS ORS). Answer sheets for the paper-and-pencil version are scanned, the data are entered into the reporting system, and the test scores are immediately accessible on DCAS ORS, as in online testing. DCAS ORS provides information on student performance and an aggregated summary at various levels—state, district, school, class, and teacher. In addition, a DCAS Family Report is distributed to parents in paper format and to each school district for individual students in the spring.

To ensure the appropriate use and interpretation of DCAS results, AIR conducted focus groups to understand what data the stakeholders—parents, principals, teachers, district staff—would find useful, what elements they would find confusing, and what they wanted to do with the data. Information collected from the focus groups helped guide the development of score reporting.

In February 2010, two focus groups were convened on the Online Reporting System. Participants in the focus groups represented all the different geographic regions from the entire state. The first focus group, with seven participants, included principals, district test coordinators, and school test coordinators. The second focus group had nine participants; in addition to principals and test coordinators, the second focus group also included an inclusion specialist and a counselor. At the focus group meetings, participants reviewed sample reports and discussed how they would use the reports and, in general, how they want to use the assessment data.

In mid-January 2011, AIR also conducted three one- to two-hour focus groups meetings, one in each county, on the Family Reports. Each focus group had six to eight parent representatives. Mockup copies of the Family Report were presented to the participants to collect comments on

the report layout, language used, presentation of test scores, and the intended use of the reports by parents.

The information collected from educators provided guidance for the development of the DCAS Online Reporting System. The system is designed so that users can explore the assessment data along the *Who*, *What*, and *When* dimensions.

- The *Who* dimension allows users to see the data for larger and smaller groups (district, school, class, teacher, student).
- The *What* dimension allows users to explore the data for more or less specific content details (subject, reporting categories, topics).
- The *When* dimension allows users to see the data as a snapshot of current performance or as a trend over time.

The information collected from parents provided guidance for the development of the Family Report. For example, the Family Report is printed in full color with a consistent color scheme that allows parents to easily identify their child's test score on the reporting scale and in the performance level and to compare their child's achievement with that of the school, the school district, and the state. The Family Report also includes a "Next Step" section that provides parents with information on how to improve their child's performance.

The DCAS Online Reporting System (DCAS ORS) is a web-based application (http://de.portal.airast.org) that provides users access to the DCAS results at multiple levels. The access to the online reports depends on the user's level and responsibility. The following six types of users have access to the system:

- State users: Access to all data at the district, school, teacher, class, and student-levels.
- District users: Access to all data for their district and for the schools, teachers, classes, and students within that district. There are two types of district-level users:
  - District Administrator (DA)
  - District Test Coordinator (DTC)
- School users: Access to all data for their school and for the teachers, classes, and students in the school. There are two groups of school-level users:
  - School Administrators (SA)
  - Test Administrators with Reporting Access (TAR)
- Teachers: Access to data for all the classes and students associated with the teacher.

Access to the reports is password-protected, and password management is handled by the Delaware Department of Education (DDOE). Users are able to access the data at their level and below. For example, a school administrator can access the School Reports, Teacher Reports, Class Reports, and Student Reports for his or her school but not for a different school.

1.2 PERFORMANCE LEVELS AND REPORTING CATEGORIES

The DCAS reading and math report two scores for each student in each test administration: an accountability score and an instructional score. The *accountability score* is based on the items measuring on-grade content only, and the *instructional score* is based on all items presented to the student—both on- and off-grade. The accountability score from the spring administration is used to meet the NCLB requirements. The instructional score provides more information about the skill and knowledge level that a student knows. Section 8.1 of Volume 1 describes how each score is computed. The DCAS science and social studies report the accountability score only. In addition, the associated performance level is reported at the student level for each test. Technical details on the scoring algorithm are also presented in Volume 1, Section 8.

The DCAS Online Reporting System uses the accountability scores to classify student performance into four levels: *Well Below Standard*, *Below Standard*, *Meets Standard*, and *Advanced*. The cut scores of these various levels are listed by subjects in the following tables. For example, in Table 1, in grade 2, a score of 587 or above but below 623 is classified as *Below Standard* in reading. A score that is under 587 falls into the level of *Well Below Standard*, and a score that is at or above 682 is considered *Advanced*.

Table 1: Cut Scores for Reading, by Grade

Grade	Below Standard	Meets Standard	Advanced
2	587	623	682
3	651	690	737
4	682	721	772
5	700	739	798
6	725	758	818
7	744	776	827
8	764	800	844
9	767	811	853
10	775	820	859

Table 2: Cut Scores for Mathematics, by Grade

Grade	Below Standard	Meets Standard	Advanced
2	503	577	688
3	593	659	750
4	649	700	793
5	690	732	811
6	716	757	836
7	740	779	850
8	767	800	862
9	775	812	872
10	792	830	897

Table 3: Cut Scores for Science, by Grade

Grade	Below Standard	Meets Standard	Advanced
5	352	400	435
8	360	400	435
10	381	400	435

Table 4: Cut Scores for Social Studies, by Grade

Grade	Below Standard	Meets Standard	Advanced
4	383	400	435
7	371	400	435

Reporting categories are also available on content strands in the online report. Reporting categories for mathematics and reading are based on instructional scores; for science and social studies, they are based on accountability scores. Table 5 shows reporting categories by subject.

Table 5: Reporting Category for DCAS 2010–2011

Subject	Reporting Category*
Reading, Grades 3–10	Reading Comprehension Literary Text
Mathematics Grades 3–10	Numeric Reasoning Algebraic Reasoning Geometric Reasoning Quantitative Reasoning
Science Grades 5, 8, 10	Life Science Earth Science Physical Science
Social Studies Grades 4, 7	Civics Economics Geography History

*Note: Reporting categories with less than eight items are not reported. Please refer to Section 3.1.6 for a list of grades and suppressed reporting categories.

1.3 AVAILABLE REPORTS IN THE ONLINE REPORTING SYSTEM

The DCAS ORS is hierarchically structured. The home page has a drop-down menu with a list of aggregated units—e.g., districts, schools within a district, or teachers within a school—to choose from. An authorized user can view reports at his or her own aggregated unit and below. For example, a school user can view the reports at the school, teacher, and student levels at his or her school. In addition to the aggregated units, DCAS ORS also allows users to create customized

rosters. For example, a teacher may create a group (roster) of students who participate in an after-school intervention program.

Table 6 summarizes the types of score reports available in DCAS ORS and the levels at which the reports can be viewed. Description of each report (page) is provided below.

Table 6: Delaware Online Score Reports Summary

Type of Report Page	Level of Aggregation	Description
Home	District, school	Summary of performance and participation (%Proficient and %Tested) across grades and subjects
Subject Detail	District, school, teacher, and roster	Overall performance for a subject and a grade for all students and by subgroups
Reporting Category Detail	District, school, teacher, and roster	Performance on the reporting category for a subject and a grade for all students and by subgroups
Topics	District, school, teacher, and roster	Relative performance by topics for a group of students. Individual students and reports by subgroups are not available.
Trend	District, school, teacher, and roster	Longitudinal comparison of scores across windows within a year
Roster	Student	List of students
Student Detail Report	Student	Student performance for a selected subject

1.3.1 The Home Page

Once a user has logged in, the home page appears. The home page has two tabs for users to choose from, the Participation Report tab and the Performance Report tab.

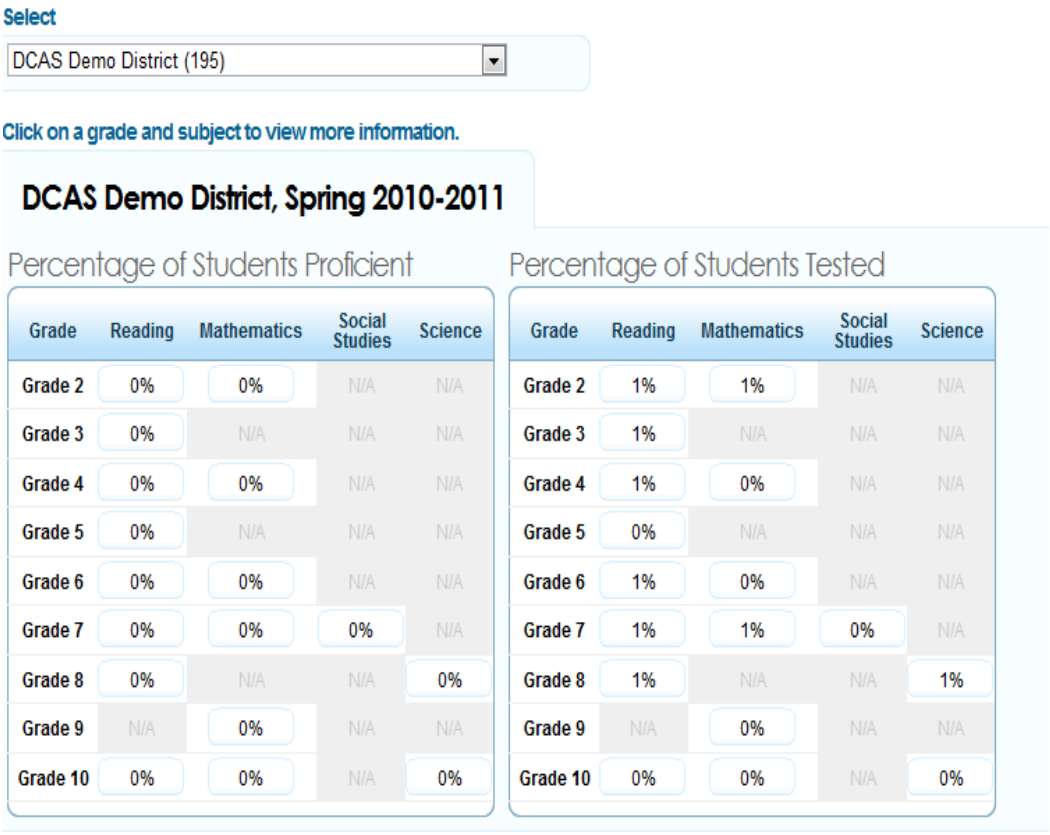
The Participation Report allows teachers, principals, and district staff to see which students still need to complete their tests. Users can select from a series of options to customize the group of students whose participation status in a particular grade/subject (e.g., have started but have not completed, have not started) is to be reviewed. Users can export the list into an Excel file and download the file.

DDOE staff, district staff, principals, and teachers can access student performance in DCAS through the Performance Report tab. Users must first choose the administration window (fall 2010–2011, winter 2010–2011, and spring 2010–2011) and, for DDOE and district level staff, the aggregated level. For example, a district staff member who is authorized to view data of multiple districts can select one district at a time for the spring 2010–2011 DCAS scores. Once the aggregated level and the administration window are selected, the Percentage of Students Proficient and the Percentage of Students Tested tables, as shown in Figure 1 below, will be filled with actual percentages. Principals and teachers can only choose the administration window as they are not allowed to view results of other schools.

The Percentage of Students Proficient table shows the percentage of students who achieved a score of Proficient out of all students who completed the spring 2010–2011 administration by grade and subject. For example, in Figure 1 (with fake data), zero percent of students out of those who completed testing for Grade 3 Reading tested at or above Proficient. The Percentage of Students Tested table shows the percentage of students who completed the test out of the total

number of students enrolled. For example, as Figure 1 shows, only 1 percent of students out of those enrolled in grade 3 in school year 2010–2011 completed their Grade 3 Reading test. If no student took the test, the cell is filled with "NA." A zero percent indicates that there are completed tests, but the percentage is rounded down to zero.

Figure 1: Performance Report Screenshot



Grade	Reading	Mathematics	Social Studies	Science
Grade 2	0%	0%	N/A	N/A
Grade 3	0%	N/A	N/A	N/A
Grade 4	0%	0%	N/A	N/A
Grade 5	0%	N/A	N/A	N/A
Grade 6	0%	0%	N/A	N/A
Grade 7	0%	0%	0%	N/A
Grade 8	0%	N/A	N/A	0%
Grade 9	N/A	0%	N/A	N/A
Grade 10	0%	0%	N/A	0%

Grade	Reading	Mathematics	Social Studies	Science
Grade 2	1%	1%	N/A	N/A
Grade 3	1%	N/A	N/A	N/A
Grade 4	1%	0%	N/A	N/A
Grade 5	0%	N/A	N/A	N/A
Grade 6	1%	0%	N/A	N/A
Grade 7	1%	1%	0%	N/A
Grade 8	1%	N/A	N/A	1%
Grade 9	N/A	0%	N/A	N/A
Grade 10	0%	0%	N/A	0%

1.3.2 Subject Detail Page

When the user clicks a particular percentage shown in Figure 1, the Subject Detail page will open (see Figure 2 below) with the chosen window of administration (fall, winter, spring), subject, and grade. The table on this page has the following data:

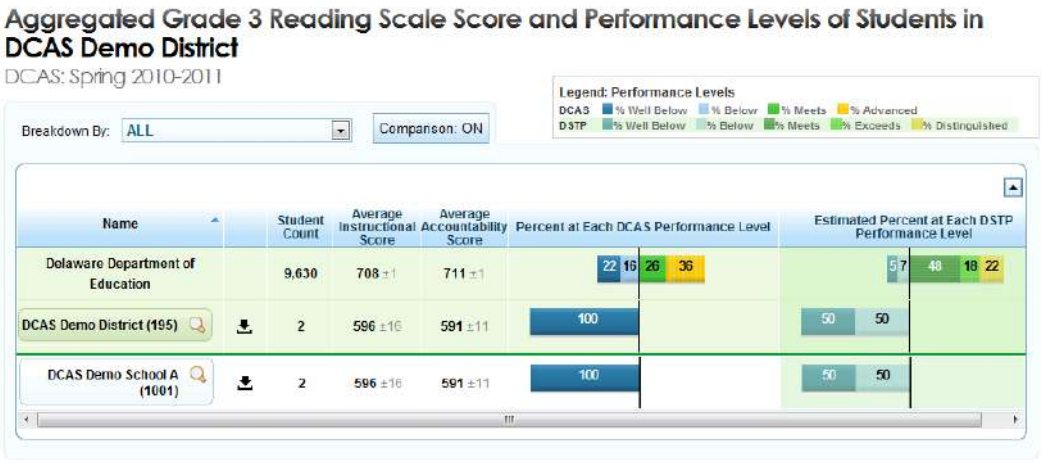
- Student Count: The number of students who have completed the test at the time the reporting is pulled for the chosen grade and subject.
- Average Instructional Score: The average instructional scale score of those students who have completed the selected assessment and aggregated level. The ± represents the boundaries within which this estimated average may fall. For example, in Figure 2, the average instructional score for the state of Delaware ranges between 707 and 709. For a more detailed description of the boundary, please see Section 3.2.2.
- Average Accountability Score: The average accountability scale score of those students who have completed the selected assessment. The ± represents the boundaries within

which this estimated average may fall. For example, in Figure 2, the average accountability score for the state of Delaware ranges between 710 and 712. For a more detailed description of the boundary and the calculation of the average, please see Section 3.2.2.

- Percent at Each DCAS Performance Level: This is the distribution of students across the DCAS performance levels of Well Below, Below Proficiency, Meets Proficiency, and Advanced in the selected assessment and aggregated level. Performance levels are based on accountability scores, not on instructional scores.
- Estimated Percent at Each DTSP Performance Level: This column provides an estimated percentage of students who would have placed in each of the DSTP performance levels.

Also available on this page is the "Breakdown By" box, which allows users to break down the results into race/ethnicity groups, gender, CD504 status, special education status, and English Language Learner status. Users can also show or hide the comparison groups (at a higher aggregated level) by clicking the Comparison On/Off button.

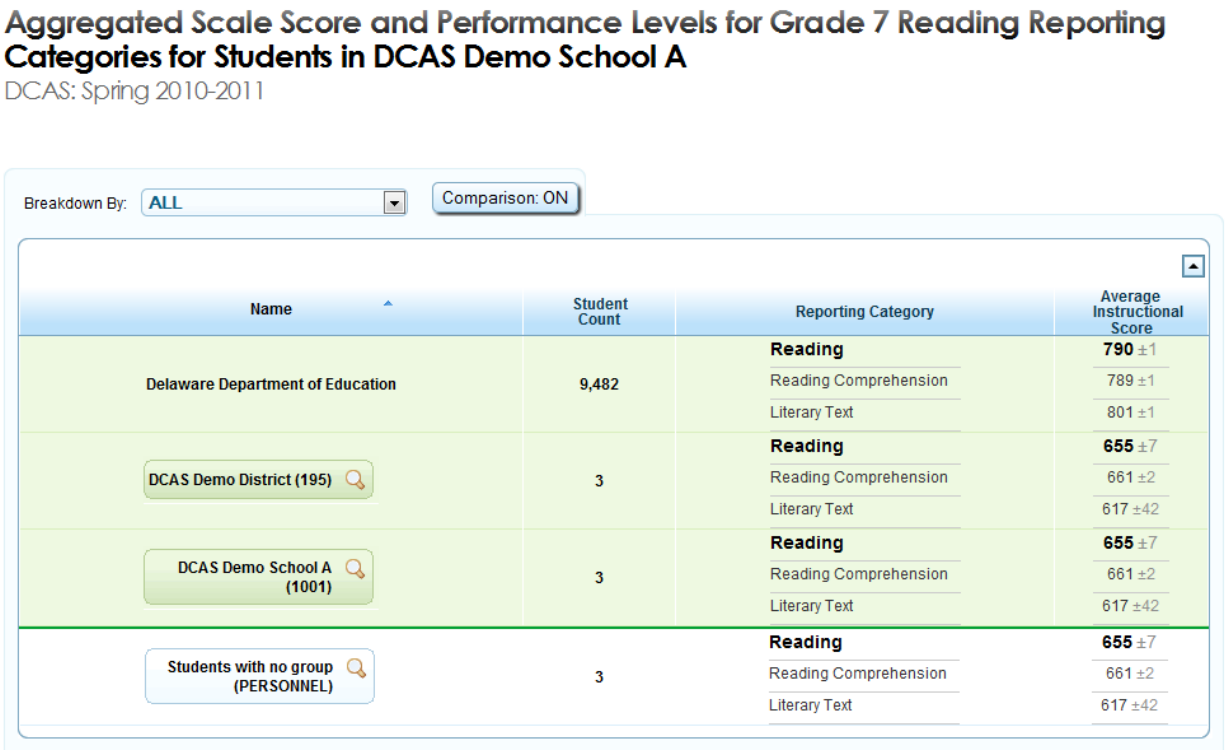
Figure 2: A Sample Page of Aggregated Subject Detail Report



1.3.3 Reporting Category Detail Page

By clicking the magnifying glass to the right of the district/school name, a user can access the Reporting Category Detail page, which shows, at the chosen aggregated level, how students performed in each reporting category for the selected assessment (see Figure 3). At the Reporting Category Detail page, the average instructional score is presented for Reading and Mathematics, and the average score is presented for Science and Social Studies.

Figure 3: A Sample Reporting Category Detail Page



1.3.4 Topics Page

From the Reporting Category Detail page, users can access the Topics page. This page displays the relative strengths and weaknesses of a particular group of aggregated level of students for a topic, which is a group of grade-level expectations. On this page, relative strengths and weaknesses are indicated by four symbols: +, =, –, and ?. A "+" sign next to a topic means that this topic is a strength for the students selected; they have performed better on this topic than on the test as a whole. A "–" sign means the opposite; these students have not performed as well on this topic as on the test as a whole. An "=" signifies that the students have performed just as well on this topic as on the test as a whole. A "?" indicates that there was insufficient information to make any determination.

With paper forms, the reporting of topics is typically not advisable. Any given form may contain a very small number of items related to a given topic, and all students within a class respond to the same items. Hence, there is no broad sampling of the content domain within a class when all students respond to the same items. With an adaptive test, it is still true that any given student will respond to a small number of items within a topic. However, students within a class are typically exposed to a varied set of items. Consequently, there is much larger sampling of the content domain.

These topics (now referred to as Benchmarks) are reported based on the calculations described below.

Computation for Multiple-Choice Items

Let $\Pr_{si}$ $(z_{si} = 1)$ represent the probability that student *s* will respond correctly to item *i* and where *z* is the observed response of student *s* to item *i*. Because the Rasch model is used, we have:

$$\Pr_{si}(z_{si} = 1) = \frac{1}{1 + \exp\left[-(\theta_s - b_i)\right]}$$

From this we compute $\delta_{si} = z_{si} - \Pr_{si}$.

Computation for Polytomous Items

Let $\Pr_{sik}$ (x_{si}) be the probability that student *s* scores in category *k* for item *i* and where x_{si} is the observed score for student *s* on item *i*. Because Master's Partial Credit Model is used, we have:

$$\Pr_{sik} = \frac{\exp \sum_{k=1}^{x_i} (\theta - \delta_{ki})}{1 + \sum_{j=1}^{m_i} \exp \sum_{k=1}^{j} (\theta - \delta_{ki})}$$

thus yielding a probability of scoring within each of the possible score categories $k = 1, \ldots, m_i$. From this, we compute:

$$\delta_{si} = x_{si} - \left(\Pr_{si,1} + 2\Pr_{si,2} + 3\Pr_{si,3} +, \ldots, m_i \Pr_{si,m_i}\right)$$

These deviations from the conditional expectations are subsequently aggregated to form benchmark reporting scores as:

$$\bar{\delta}_B = N^{-1} \sum_{s=1}^{N} \left(\frac{\sum_{i \in B} \delta_{si}}{\sum_{i \in B} m_i} \right)$$

where *B* denotes benchmark category and *m* is the maximum number of score points on item *i*. The outer summation is generic in the sense that it can be used for any level of aggregation. For example, for teacher-level reporting we have $s = 1, \ldots, N$ where *N* denotes all students within the teacher's class. For school reporting we still have $s = 1, \ldots, N$ where *N* denotes all students within the school.

1.3.5 Student Detail Page

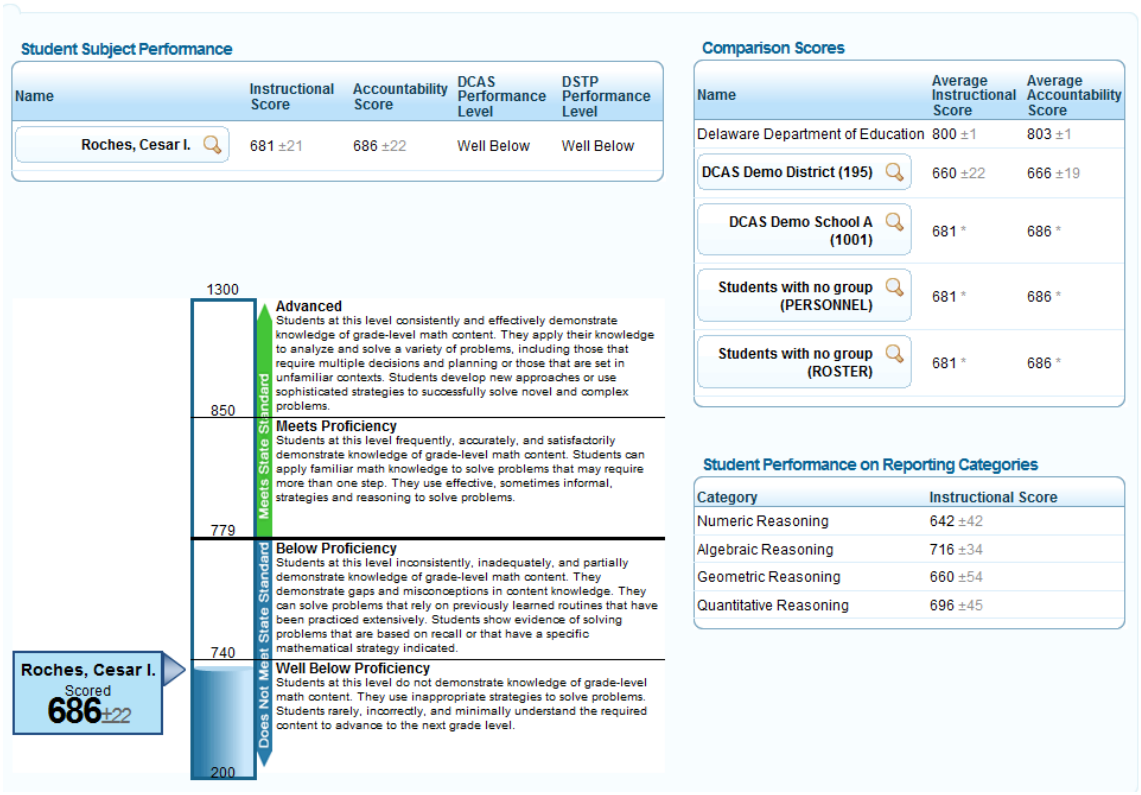
The Student Detail page presents detailed information on the performance of a particular student (see Figure 4 below). It is divided into four sections: the Student Subject Performance, the Barrel Chart, the Comparison Scores, and the Student Performance on Reporting Category section.

The *Student Subject Performance* section shows the student's instructional score and accountability score. The standard error of measurement (SEM) is shown as the ± next to each of the scores (refer to Section 3.2.2 for a detailed description and interpretation of the SEM). The individual student's DCAS performance, which is based on his or her accountability score only, and the estimated DSTP performance level, which is also based on the accountability score, are presented in this section.

Below the *Student Subject Performance* section is the *Barrel Chart*, which displays the student's accountability score and its associated SEM in the entire scale. This chart gives readers a sense of where the student stands in relation to the highest and lowest score possible for the test and the range within which the student's true score falls. For example, in Figure 4, the student scored 686 with a bandwidth of 657 and 715. The barrel chart, which has a minimum score of 200 and a maximum score of 1,300, is demarcated by different performance-level cut scores. For example, at or above 779 and below 850 are the scores that fall into the Meets Proficiency level for Grade 7 Mathematics. The text next to the barrel chart provides an explanation of the level of mastery at each performance level.

The *Comparison Scores* section shows how the student's score compares with the average scores of a roster, class (teacher), school, district, and the state. The chart of *Student Performance on Reporting Categories* shows the student's instructional scores for each reporting category.

Figure 4: A Sample Student Detail Report, Grade 7 Mathematics



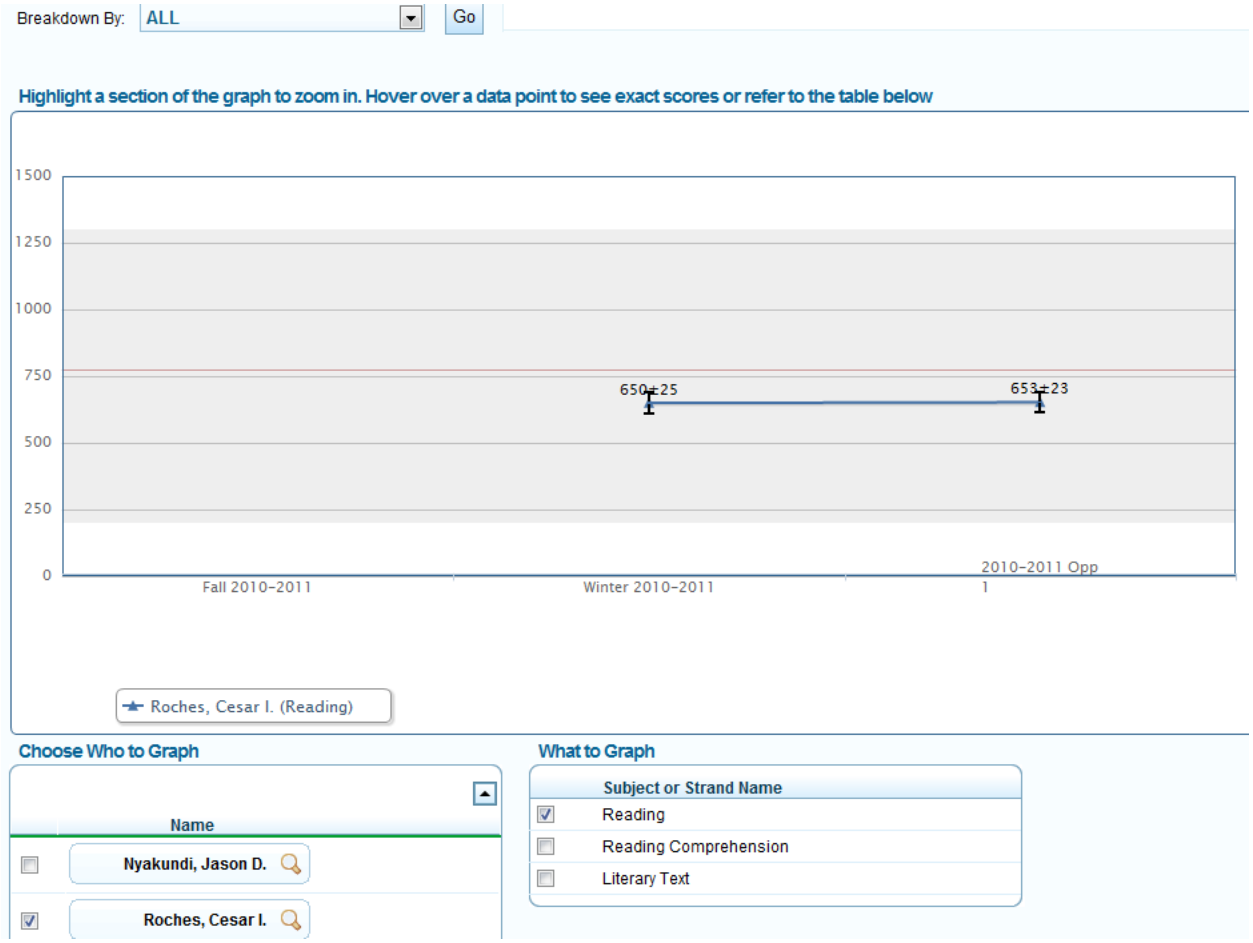
1.3.6 Trend Page

The performance of individual students or a group of students can be monitored using the trend report. The Trend page, shown in Figure 5 below, displays a student's scale score across the fall, winter, and spring 2010–2011 test windows. When the student has completed the 2011–2012 test, data for the reported year's test windows will be shown along with the scores from previous years' spring test windows.

In the "Choose Who to Graph" section, the user can choose trends for up to five students or groups to compare by marking the checkboxes. When the user places the mouse over a score on

the graph, the student's name and score will be displayed. Error bars are plotted for each data point. The user can also zoom in to a particular section of the graph for a clear view. In the "What to Graph" section, the user can choose to graph the progress at the reporting category level. A sample trend report is presented in Figure 5 below. Trend data are also available for the school and district level as well.

Figure 5: A Sample Trend Page for an Individual Student



1.3.7. Roster Page

In addition to the administrative hierarchy of state, district, school, and teacher, a user can create a group of students within a school using the Manage Rosters tool. For example, a teacher may want to create a roster of students who participated in an after-school intervention program to track their performance as a group. Once created, the roster will be available as an aggregated unit for Subject Details Reports, Reporting Category Reports, Topic Reports, and Trend Reports. Rosters created in this way do not affect accountability and can be edited or deleted at any time.

1.4 FAMILY REPORTS

Family Reports are delivered as printed materials by AIR directly to parents. The primary purpose of the Family Report is to provide a document to enable parents to understand their

child's performance in each subject and what their child can do in relation to the state standards. The Family Report is a folded color document printed on both sides that presents information that enables parents to

- learn how their child's performance compares with that of other children and groups;
- understand the trend in their child's proficiency over time;
- understand what is expected of their child to achieve or maintain proficiency.

The Family Report was delivered to the parents and the district between July 22 and 29, 2011, about seven weeks after the third test window closed. A single Family Report displays all four subjects—Reading, Mathematics, Science (for the relevant grades), and Social Studies (for the relevant grades).

The Family Report uses the barrel chart discussed in the Student Detail page (Section 1.4.5) to display the student's accountability scores and his or her associated performance level in Reading and Mathematics and (as applicable) Science and Social Studies. To avoid overburdening parents with information, the report does not include standard errors of measurement and instructional scores. Also similar to the Student Detail page, average scores from the student's school and district and from the state are included. Below the barrel chart is a line graph that shows the student's score over time for Reading and Mathematics.

Each Family Report differs from the Student Detail page in that it has a "Next Step" section that provides pointers that parents can do to further their child's skills and knowledge. Text in the Next Step changes with the performance level of the student—it is customized to the performance levels attained. A sample of the Family Report is included in Appendix A.

In order to identify which next steps are included on a student's report, the following statistical procedures are used to identify whether such areas of relative weakness exist. These areas of relative weakness are derived from an Ordinary Least Squares (OLS) regression in which the scores for each reporting category are regressed on the other scores for the remaining reporting categories. Technically, the statistical model takes the following general form:

$$Y = \mu + \sum_{j-1}^{K} \beta_j X_j + e$$

$$e \sim N(0, \sigma^2)$$

where the outcome variable (Y) is a single reporting Category and the remaining reporting Categories are the independent variables (X_j). This approach determines whether individual students, classes, or schools performed higher or lower than conditionally expected given their performance on the other strands. Hence, it is a relative strength or a relative weakness.

As a concrete illustration, consider the four Reporting Categories in mathematics: Numeric Reasoning (NR), Algebraic Reasoning (AR), Quantitative Reasoning (QR), and Geometric Reasoning (GR). At the student level, we perform an OLS regression, which uses the estimated individual student's scores in numeric reasoning as the outcome variable. The student's scores for the remaining reporting categories are the regressors, used as follows:

$$Y_{NR} = \mu_{NR} + \beta_{AR}X_{AR} + \beta_{QR}X_{QR} + \beta_{GR}X_{GR} + e$$

Given the estimates of the fixed effects, we then form the conditional prediction for each individual as:

$$\hat{Y}_{NR} = \hat{\mu}_{NR} + \hat{\beta}_{AR}X_{AR} + \hat{\beta}_{QR}X_{QR} + \hat{\beta}_{GR}X_{GR}$$

where *X* is the student's observed score for the various strands. The standard error of the estimate (SEE) is computed from the regression as:

$$SEE = \sqrt{[\boldsymbol{x}_s'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_s + 1] * \sigma^2}$$

where $\boldsymbol{x}_s$ is the row of the model matrix $\boldsymbol{X}$ corresponding to student *s***.** We then perform the following *t*-test:

$$t = \frac{\hat{Y}_{NR} - Y_{NR}}{SEE}$$

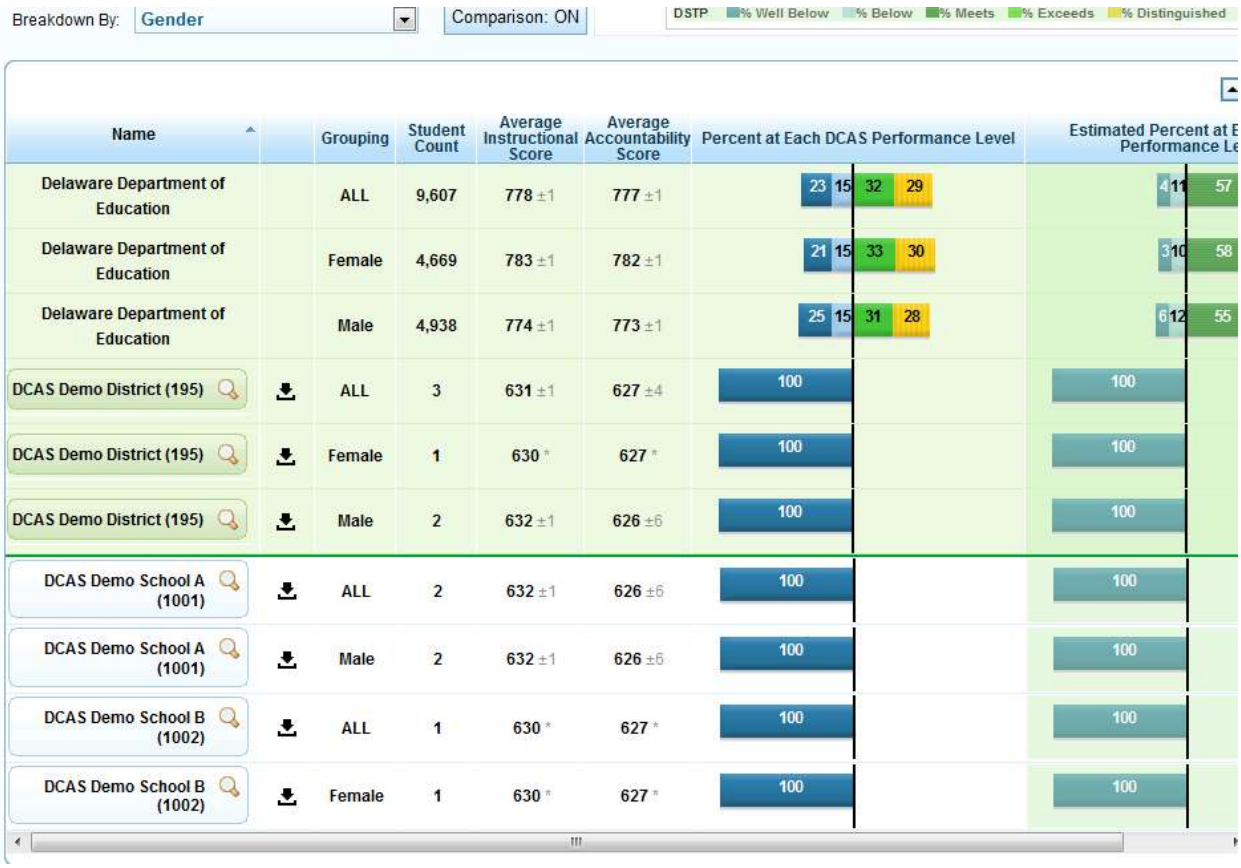
Here, the critical value is –1 instead of –1.96. When the *t* value is less than –1, the student is identified as having a relative weakness—the observed score for the reporting category is more than one SEE lower than the expected score, given the student's performance in the other reporting categories.

Text for the Next Steps section is derived from the identified relative weakness and the student's overall performance classification. For instance, a student with a relative weakness in Quantitative Reasoning but scoring overall at the Meets Proficiency level has a different Step than a student with a relative weakness in Numeric Reasoning who is scoring overall in the Advanced performance level.

2. SUBGROUP REPORTING

In the Subject Detail page (see Section 1.4.2, Figure 2), information displayed at the aggregated level can be further broken down by demographic subgroups, as displayed in Figure 6 below.

Figure 6: Score Reporting by Gender



The following subgroups are reported:

- Gender
- Race/ethnicity, which is divided into American Indian/Alaskan Native, Asian/Pacific Islander, African-American, Hispanics, white, and multi-racial
- Special education status
- CD504 status
- Title I (economically disadvantaged) status
- English Language Learners (ELL) status

Demographic characteristics for a student are reported as they are recorded in the administration window. Subgroup statistics are not reported in the Family Report.

3. INTERPRETATION OF REPORTED SCORES

The DCAS ORS offers a variety of resources to help parents and educators understand student test scores and use the data and associated information to improve student learning and classroom instruction. These resources are available at http://de.portal.airast.org/, and a "Definitions" button leads users to the interpretation of terms used in reports. In addition, AIR conducted a series of in-person workshops in early 2011 to introduce various users to the ORS site. These workshops were conducted with the live system. Samples of training materials are included as Appendix B of this volume. This section provides guidance for appropriate interpretations and uses of test results.

3.1 BUSINESS RULES

Appropriate interpretation of reported scores is based on the understanding of the business rules—the rules that are applied for aggregated test results and which students are to be included or excluded from aggregation.

3.1.1 Aggregation

Test data are collected and kept at the individual student level during the test period. Aggregations to a higher unit—such as class, school, or district—are calculated directly from the student level for the current window. More specifically, state, district, school, and class-level aggregates are calculated by aggregating all the students in the state, in the district, in the school, and in the class, respectively. For example, the percentage Proficient for a district is based on a count of the students who scored at or above the Meets Standard level in that district, rather than on the percentage Proficient of each school in the district.

3.1.2 Student Mobility Rules

If a student transfer from one district (A) to another district (B) is reported to Delaware Student Information System (DELSIS), the student's test record (his or her test scores) will be removed from the report for the previous school district and will appear on the new school district. Therefore, this student will not be included in the aggregation for district A. However, district B will be able to track this student's test records for trend analysis.

3.1.3 Student Inclusion in Trend Data

The Trend Report presents longitudinal data for individual students as well as at the aggregated levels. If a student moved to Delaware after the first administration window was closed, he or she is not included for aggregation in the trend report of any administration window for the year.

3.1.4 Incomplete Test and Invalid Test Score

If a student does not complete all the answers by the end of each testing window, no valid score is assigned to the student. In extreme circumstances, such as cheating, the score is considered invalid. Invalid scores are excluded from reporting,

3.1.5 Minimum Group Size

There is no restriction on the cell size for group-level data, and no suppression rule is applied with respect to the number of students in a group.

3.1.6 Minimum Items Presented in a Reporting Category

A reporting category requires a minimum of eight operational items (or score points) per category. Since the mathematics blueprint requires fewer than eight items for some reporting categories, the following Reporting Categories are not available for these four grades.

1. Grade 3 Mathematics: Quantitative Reasoning
2. Grade 4 Mathematics: Quantitative Reasoning
3. Grade 9 Mathematics: Numeric Reasoning, Geometric Reasoning
4. Grade 10 Mathematics: Numeric Reasoning

3.2 INTERPRETATION OF DCAS SCORES

3.2.1 Scale Scores

DCAS reading and mathematics scores are reported on vertical scales that allow comparison of student performance across administration windows within the school year and across grades over time. The vertical scales also make it possible to measure student growth. For science and social studies, the scores are reported on a horizontal scale and cannot be used to measure the student's growth between grades over time. The technical details of transforming the raw scores into a horizontal scale for Science and Social Studies, and into a vertical scale for Reading and Mathematics, are presented in Volume 1, Section 6.2 and 6.3, respectively.

Scale scores are grouped into performance levels that describe how much students must know and be able to do. Table 7 below provides the description of these levels. The cut scores for these levels are presented in Tables 1–4.

Table 7: Descriptions of Mastery by Performance Levels for DCAS Scores, by Subjects

Performance Level	Reading	Mathematics	Science	Social Studies
Well Below Standard	Students rarely, incorrectly, or minimally apply the knowledge and skills outlined in the state content standards for Reading.	Students rarely, incorrectly, or minimally apply the knowledge and skills outlined in the state content standards for Mathematics.	Students rarely, incorrectly, or minimally apply the knowledge and skills outlined in the state content standards for Science.	Students rarely, incorrectly, or minimally apply the knowledge and skills outlined in the state content standards for Social Studies.
Below Standard	Students inconsistently, inadequately, or partially apply the knowledge and skills outlined in the state content standards for Reading.	Students inconsistently, inadequately, or partially apply the knowledge and skills outlined in the state content standards for Mathematics.	Students inconsistently, inadequately, or partially apply the knowledge and skills outlined in the state content standards for Science.	Students inconsistently, inadequately, or partially apply the knowledge and skills outlined in the state content standards for Social Studies.
Meets Standard	Students frequently, accurately, and satisfactorily apply the knowledge and skills outlined in the state content standards for Reading.	Students frequently, accurately, and satisfactorily apply the knowledge and skills outlined in the state content standards for Mathematics.	Students frequently, accurately, and satisfactorily apply the knowledge and skills outlined in the state content standards for Science.	Students frequently, accurately, and satisfactorily apply the knowledge and skills outlined in the state content standards for Social Studies.
Advanced	Students consistently, accurately, effectively, and skillfully apply the knowledge and skills outlined in the state content standards for Reading.	Students consistently, accurately, effectively, and skillfully apply the knowledge and skills outlined in the state content standards for Mathematics.	Students consistently, accurately, effectively, and skillfully apply the knowledge and skills outlined in the state content standards for Science.	Students consistently, accurately, effectively, and skillfully apply the knowledge and skills outlined in the state content standards for Social Studies.

3.2.2 Reported Statistics

The following key statistics are presented in the DCAS Online Reporting System and Family Report:

- Percentage of Students Tested: The percentage of students who have tested within a test window at the time when the report was pulled out of the total number of enrolled students for that grade and subject for school year 2010–2011. For example, 10 percent means that 10 percent of the students completed the DCAS test of the selected grade and subject. Because the reports are presented in real time, the percentage of students tested should increase throughout the test window until it closes. The percentage is rounded to the nearest integer.
- Percentage of Students Proficient: The percentage of tested students who scored in the Meets Standard or Advanced level out of those students who completed the test in the selected administration window, grade, and subject. For example, a 15 percent at Grade 3 Reading means that out of those students who completed their test, 15 percent were tested at or above Meets Standard for grade 3. The percentage of students Proficient may

increase, decrease, or remain unchanged throughout the same test window. The percentage is rounded to the nearest integer.

- Student Count: The number of students who have taken the DCAS in that grade and subject.
- Accountability Scores: Scale scores generated only from the within-grade items that the students took. For a technical explanation of the derivation of the scores, please refer to Volume 1, Section 8. Student scores are rounded to the nearest integer.
- Average Accountability Score: The mean of the individual accountability scores of students who took the DCAS for the selected grade, subject, and aggregated level.
- Percentage at Each Performance Level: Performance levels are based on the accountability scores. It is the percentage distribution of students by performance levels.
- Estimated Percentage at Each DSTP Performance Level: An estimate of how students would have placed in each of the DSTP performance levels based on the DCAS. Using person-mean equating, DSTP cut scores were mapped onto the DCAS scale, thus estimating where DSTP performance levels might lie on the DCAS scale.
- Instructional Scores: Scale scores generated from both the within-grade and off-grade items. Student scores are rounded to the nearest integer.
- Average Instructional Scores: The mean of the instructional scores of the students who took the DCAS for the selected grade, subject, and aggregated level.
- Standard Error for the Mean of the Average Accountability/Instructional Scale Score (at the aggregated level): The ± range that is found for an aggregated unit's average scores. The boundaries of the band are 1 standard error above or below the averaged score. For example, an average accountability score of 275 ± 5 means that if the same group of students (school, district, white students, etc.) took the test multiple times, the average scores from these multiple tests would fall between 270 and 280 about 68 percent of the time.
- Standard Error of Measurement (SEM): This is the ± range that is found for an individual student's scores. The observed score is an estimate of the true score—the true ability of the student. The SEM provides information about the certainty, or confidence, of the scores reported for the student. The boundaries of the score band are 1 standard error above or below the student's observed score, representing the range of score values that are likely to have the true score. For example, 310 ± 10 means that if the student were tested multiple times, the true score of the student would fall between 300 and 320 about 68 percent of the time. Because students are administered different sets of items in each window, the standard error of measurement can be different for the same scale score.