

TECHNICAL REPORT



**for the
Pennsylvania
System of School Assessment**

**2010 Grade 12 Fall Retest
Mathematics, Reading, Science, and Writing**

**Provided by
Data Recognition Corporation**

April 2011

Table of Contents

Glossary of Common Terms	i
PSSA: The Pennsylvania System of School Assessment	1
Grade 12 PSSA Fall Retest	1
Item Analysis	2
<i>Multiple-Choice (MC) Items</i>	2
<i>Open-Ended (OE) Items</i>	2
Raw-to-Scaled Score Conversions	3
Summary of the Grade 12 Retest Results	4
<i>Scaled Score Results</i>	4
<i>Performance Level Results</i>	9
Appendix A: 2010 Grade 12 Fall Mathematics Retest Multiple-Choice Item Statistics	11
Appendix B: 2010 Grade 12 Fall Mathematics Retest Multiple-Choice Rasch Item Statistics	13
Appendix C: 2010 Grade 12 Fall Mathematics Retest Open-Ended Item Statistics	15
Appendix D: 2010 Grade 12 Fall Reading Retest Multiple-Choice Item Statistics	16
Appendix E: 2010 Grade 12 Fall Reading Retest Multiple-Choice Rasch Item Statistics	18
Appendix F: 2010 Grade 12 Fall Reading Retest Open-Ended Item Statistics	19
Appendix G: 2010 Grade 12 Fall Science Retest Multiple-Choice Item Statistics	20
Appendix H: 2010 Grade 12 Fall Science Retest Multiple-Choice Rasch Item Statistics	22
Appendix I: 2010 Grade 12 Fall Science Retest Open-Ended Item Statistics	23
Appendix J: 2010 Grade 12 Fall Writing Retest Multiple-Choice Item Statistics	24
Appendix K: 2010 Grade 12 Fall Writing Retest Multiple-Choice Rasch Item Statistics	25
Appendix L: 2010 Grade 12 Fall Writing Retest Prompt Statistics	26
Appendix M: 2010 Grade 12 Fall Writing Retest Percentage Agreement	27
Appendix N: 2010 Grade 12 Fall Mathematics Retest Raw-to-Scaled Score Conversion Table	28
Appendix O: 2010 Grade 12 Fall Reading Retest Raw-to-Scaled Score Conversion Table	30
Appendix P: 2010 Grade 12 Fall Science Retest Raw-to-Scaled Score Conversion Table	32
Appendix Q: 2010 Grade 12 Fall Writing Retest Raw-to-Scaled Score Conversion Table	34

Glossary of Common Terms

The following table contains some terms used in this technical report and their meanings. Some of these terms are used universally in the assessment community, and some of these terms are used commonly by psychometric professionals.

Table G–1. Glossary of Terms

Term	Common Definition
Ability	In Rasch scaling, ability is a generic term indicating the level of an individual on the construct measured by an exam. As an example for the PSSA, a student’s reading ability is measured by how the student performed on the PSSA Reading test. A student who answered more items correctly has a higher ability than a student who answered fewer items correctly.
Adjacent Agreement	A score/rating difference of one (1) point in value usually assigned by two different raters under the same conditions (e.g., two independent raters give the same paper scores that differ by one point).
Alternate Forms	Two or more versions of a test that are considered exchangeable, i.e., they measure the same constructs in the same ways, are intended for the same purposes, and are administered using the same directions. More specific terminology applies depending on the degree of statistical similarity between the test forms (e.g., parallel forms, equivalent forms, and comparable forms) where parallel forms refers to the situation in which the test forms have the highest degree of similarity to each other.
Average	A measure of central tendency in a score distribution that usually refers to the arithmetic mean of a set of scores. In this case, it is determined by adding all the scores in a distribution and then dividing the obtained value by the total number of scores. Sometimes people use the word average to refer to other measures of central tendency such as the median (the score in the middle of a distribution) or mode (the score value with the greatest frequency).
Bias	In a statistical context, bias refers to any source of systematic error in the measurement of a test score. In discussing test fairness, bias may refer to construct-irrelevant components of test scores that differentially affect the performance of different groups of test takers (e.g., gender, ethnicity, etc.). Attempts are made to reduce bias by conducting item fairness reviews and various differential item functioning (DIF) analyses, detecting potential areas of concern, and either removing or revising the flagged test items prior to the development of the final operational form of the test (see also Differential Item Functioning).
Constructed-Response Item	See Open-Ended Item.
Content Validity Evidence	Evidence regarding the extent to which a test provides an appropriate sampling of a content domain of interest (e.g., assessable portions of a state’s Grade 6 mathematics curriculum in terms of the knowledge, skills, objectives, and processes sampled.)

Term	Common Definition
Core-Linking Item	Items that are utilized during the linking process (see also Linking). They are a subset of the PSSA operational items and so they 1) are the same on all test forms for any grade/subject area test and 2) contribute to student total raw scores and scaled scores.
Criterion-Referenced Interpretation	When a score is interpreted as a measure of a student’s performance with respect to an expected level of mastery, educational objective, or standard. The types of resulting score interpretations provide information about what a student knows or can do with respect to a given content area.
Cut Score	A specified point on a score scale such that scores at or above that point are interpreted or acted upon differently from scores below that point (e.g., a score designated as the minimum level of performance needed to pass a competency test). One or more cut scores can be set for a test that results in dividing the score range into various proficiency level ranges. Methods for establishing cut scores vary. For the PSSA, three cut scores are used to place students into one of four performance levels (see also Performance Level Setting).
Decision Consistency	The extent to which classifications based on test scores would match the decisions based on scores from a second, parallel form of the same test. It is often expressed as the proportion of examinees who are classified the same way from the two test administrations.
Differential Item Functioning (DIF)	A statistical property of a test item in which different groups of test takers (who have the same total test score) have different average item scores. In other words, students with the same ability level but different group memberships do not have the same probability of answering the item correctly (see also Bias).
Distractor	An incorrect option in a multiple-choice item (also called a foil).
Equating	The strongest of several linking methods used to establish comparability between scores from multiple tests. Equated test scores should be considered exchangeable. Consequently, the criteria needed to refer to a linkage as equating are strong and somewhat complex (equal construct and precision, equity, and invariance). In practical terms, it is often stated that it should be a matter of indifference to a student if he/she takes any of the equated tests (see also Linking).
Equating Block (EB) Items	The PSSA uses multiple test forms for each grade/subject area test. Each form is composed of operational (OP) items, equating block (EB) items, and field test (FT) items. EB items are utilized during the linking process (see also Linking). Each test form includes a set of EB items. EB items are not part of any student scores.
Error of Measurement	The amount by which the score actually received (an observed score) differs from a hypothetical true score (see also Standard Error of Measurement).
Exact Agreement	When identical scores/ratings are assigned by two different raters under the same conditions (e.g., two independent raters give a paper the same score).

Term	Common Definition
Field Test (FT) Items	The PSSA uses multiple test forms for each grade/subject area test. Each form is composed of operational (OP) items, equating block (EB) items, and field test (FT) items. An FT item is a newly-developed item that is ready to be tried out to determine its statistical properties (see also <i>P</i> -value and Point-Biserial Correlation). Each test form includes a set of FT items. FT items are not part of any student scores.
Frequency	The number of times that a certain value or range of values (score interval) occurs in a distribution of scores.
Frequency Distribution	A tabulation of scores from low to high or high to low showing the number and/or percent of individuals who obtain each score or who fall within each score interval or category.
Infit/Outfit	Statistical indicators of the agreement of the data and the measurement model (see also Outfit/Infit).
Item Difficulty	For the Rasch model, the dichotomous item difficulty represents the point along the latent trait continuum where an examinee has a 0.50 probability of making a correct response. For a polytomous item, the difficulty is the average of the item's step difficulties (see also Step Difficulty).
Key	The correct response option or answer to a test item.
Linking	A generic term referring to one of a number of processes by which scores from different tests are made comparable to some degree. Linking includes several classes of transformations (equating, scale alignment, prediction, etc.). Equating is associated with the strongest degree of comparability (exchangeable scores). Other linkages may be very strong but fail to meet one or more of the strict criteria required of equating (see also Equating).
Logit	In Rasch scaling, logits are units used to express both examinee ability and item difficulty. When expressing examinee ability, a student who answers more items correctly has a higher logit value than a student who answers fewer items correctly. Logits are transformed into Scaled Scores through a linear transformation. When expressing item difficulty, logits are transformed <i>p</i> -value (see also <i>P</i> -value). The logit difficulty scale is inversely related to <i>p</i> -values. A higher logit value would represent a relatively harder item, while a lower logit value would represent a relatively easier item.
Mean	Also referred to as the arithmetic mean of a set of scores, is found by adding all the score values in a distribution and dividing by the total number of scores. For example, the mean of the set {66, 76, 85, 97} is 81. The value of a mean can be influenced by extreme values in a score distribution.
Measure	In Rasch scaling, measure generally refers to a specific estimate of an examinee's ability (often expressed as logits) or an item's difficulty (again, often expressed as logits). As an example for the PSSA, a student's reading measure might be equal to 0.525 logits. Or, a PSSA Reading test item might have logit equal to -0.905.

Term	Common Definition
Median	The middle point or score in a set of rank-ordered observations that divides the distribution into two equal parts such that each part contains 50 percent of the total data set. More simply put, half of the scores are below the median value and half of the scores are above the median value. As an example, the median for the following ranked set of scores {2, 3, 6, 8, 9} is 6.
Multiple-Choice Item	A type of item format that requires the test taker to select a response from a group of possible choices, one of which is the correct answer (or key) to the question posed (see also Open-Ended Item).
N-count	Sometimes designated as N or n , it is the number of observations (usually individuals or students) in a particular group. Some examples include the number of students tested, the number of students tested from a specific subpopulation (e.g., females), the number of students who attained a specific score, etc. In the follow set {23, 32, 56, 65, 78, 87}, $n = 6$.
Open-ended item	An open-ended (OE) item—referred to by some as a constructed-response (CR) item—is an item format that requires examinees to create their own responses, which can be expressed in various forms (e.g., written paragraph, created table/graph, formulated calculation, etc.). Such items are frequently scored using more than two score categories, that is, polytomously (e.g., 0, 1, 2, and 3). This format is in contrast to when students make a choice from a supplied set of answers options (e.g., multiple-choice (MC) items which are typically dichotomously scored as right = 1 or wrong = 0.) When interpreting item difficulty and discrimination indices it is important to consider whether an item is polytomously or dichotomously scored.
Operational Item	The PSSA uses multiple test forms for each grade/subject area test. Each form is composed of operational (OP) items, equating block (EB) items, and field test (FT) items. OP items are the same on all forms for any grade/subject area test. Student total raw scores and scaled scores are based exclusively on the OP items.
Outfit/Infit	Statistical indicators of the agreement of the data and the measurement model. Infit and Outfit are highly correlated, and both are highly correlated with the point-biserial correlation. Underfit can be caused when low-ability students correctly answer difficult items (perhaps by guessing or atypical experience) or high-ability students incorrectly answer easy items (perhaps because of carelessness or gaps in instruction). Any model expects some level of variability, so overfit can occur when nearly all low-ability students miss an item while nearly all high-ability students get the item correct.
Percent Correct	When referring to an individual item, the percent correct is the item's p -value expressed as a percent (instead of a proportion). When referring to a total test score, it is the percentage of the total number of points that a student received. The percent correct score is obtained by dividing the student's raw score by the total number of possible points and multiplying the result by 100. Percent Correct scores are often used in criterion-referenced interpretations and are generally more helpful if the overall difficulty of a test is known. Sometimes Percent Correct scores are incorrectly interpreted as Percentile Ranks.

Term	Common Definition
Percentile	The score or point in a score distribution at or below which a given percentage of scores fall. It should be emphasized that it is a value on the score scale, not the associated percentage (although sometimes in casual usage this misinterpretation is made). For example, if 72 percent of the students score at or below a Scaled Score of 1500 on a given test, then the Scaled Score of 1500 would be considered the 72nd percentile. As another example, the median is the 50th percentile.
Percentile Rank	The percentage of scores in a specified distribution falling at/below a certain point on a score distribution. Percentile Ranks range in value from 1 to 99, and indicate the status or relative standing of an individual within a specified group, by indicating the percent of individuals in that group who obtained equal or lower scores. An individual's percentile rank can vary depending on which group is used to determine the ranking. As suggested above, Percentiles and Percentile Rank are sometimes used interchangeably; however strictly speaking, a percentile is a value on the score scale.
Performance Level Descriptors	Descriptions of an individual's competency in a particular content area, usually defined as ordered categories on a continuum, often labeled from Below Basic to Advanced, that constitute broad ranges for classifying performance. The exact labeling of these categories, and narrative descriptions, may vary from one assessment or testing program to another.
Performance Level Setting	Also referred to as standard setting, a procedure used in the determination of the cut scores for a given assessment that is used to measure students' progress towards certain performance standards. Standard setting methods vary (e.g., modified Angoff, Bookmark Method, etc.), but most use a panel of educators and expert judgments to operationalize the level of achievement students must demonstrate in order to be categorized within each performance level.
Point-Biserial Correlation	In classical test theory this is an item discrimination index. It is the correlation between a dichotomously scored item and a continuous criterion, usually represented by the total test score (or the corrected total test score with the reference item removed). It reflects the extent to which an item differentiates between high-scoring and low-scoring examinees. This discrimination index ranges from -1.00 to $+1.00$. The higher the discrimination index (the closer to $+1.00$), the better the item is considered to be performing. For multiple-choice items scored as 0 or 1, it is rare for the value of this index to exceed 0.5.
<i>P</i> -value	An index indicating an item's difficulty for some specified group (perhaps grade). It is calculated as the proportion (sometimes percent) of students in the group who answer an item correctly. <i>P</i> -values range from 0.0 to 1.0 on the proportion scale. Lower values correspond to more difficult items and higher values correspond to easier items. <i>P</i> -values are usually provided for multiple-choice items or other items worth one point. For open-ended items or items worth more than one point, difficulty on a <i>p</i> -value-like scale can be estimated by dividing the item mean score by the maximum number of points possible for the item (see also Logit).

Term	Common Definition
Raw Score	Sometimes abbreviated by RS—it is an unadjusted score usually determined by tallying the number of questions answered correctly, or by the sum of item scores (i.e., points). (Some rarer situations might include formula-scoring, the amount of time required to perform a task, the number of errors, application of basal/ceiling rules, etc.). Raw scores typically have little or no meaning by themselves and require additional information—like the number of items on the test, the difficulty of the test items, norm-referenced information, or criterion-referenced information.
Reliability	The expected degree to which test scores for a group of examinees are consistent over exchangeable replications of an assessment procedure, and therefore, are considered dependable and repeatable for an individual examinee. A test that produces highly consistent, stable results (i.e., relatively free from random error) is said to be highly reliable. The reliability of a test is typically expressed as a reliability coefficient or by the standard error of measurement derived by that coefficient.
Reliability Coefficient	A statistical index that reflects the degree to which scores are free from random measurement error. Theoretically, it expresses the consistency of test scores as the ratio of true score variance to total score variance (true score variance plus error variance). This statistic is often expressed as correlation coefficient (e.g., correlation between two forms of a test) or with an index that resembles a correlation coefficient (e.g., calculation of a test’s internal consistency using Coefficient Alpha). Expressed this way, the reliability coefficient is a unitless index. The higher the value of the index (closer to 1.0), the greater the reliability of the test (see also Standard Error of Measurement).
Scaled Score	A mathematical transformation of a raw score developed through a process called scaling. Scaled scores are most useful when comparing test results over time. Several different methods of scaling exist, but each is intended to provide a continuous and meaningful score scale across different forms of a test.
Selected-Response Item	See Multiple-Choice Item.
Spiraling	A packaging process used when multiple forms of a test exist and it is desired that each form be tested in all classrooms (or other grouping unit (e.g., schools)) participating in the testing process. This process allows for the random distribution of test booklets to students. For example, if a package has four test forms labeled A, B, C, and D, the order of the test booklets in the package would be A, B, C, D, A, B, C, D, A, B, C, D, etc.

Term	Common Definition
Standard Deviation (SD)	A statistic that measures the degree of spread or dispersion of a set of scores. The value of this statistic is always greater than or equal to zero. If all of the scores in a distribution are identical, the standard deviation is equal to zero. The further the scores are away from each other in value, the greater the standard deviation. This statistic is calculated using the information about the deviations (distances) between each score and the distribution's mean. It is equivalent to the square root of the variance statistic. The standard deviation is a commonly used method of examining a distribution's variability since the standard deviation is expressed in the same units as the data.
Standard Error of Measurement (SEM)	It is the amount an observed score is expected to fluctuate around the true score. As an example, across replications of a measurement procedure, the true score will not differ by more than plus or minus one standard error from the observed score about 68 percent of the time (assuming normally distributed errors). The SEM is frequently used to obtain an idea of the consistency of a person's score in actual score units, or to set a confidence band around a score in terms of the error of measurement. Often a single SEM value is calculated for all test scores. On other occasions, however, the value of the SEM can vary along a score scale. Conditional standard errors of measurement (CSEMs) provide an SEM for each possible scaled score.
Step Difficulty	Step difficulty is a parameter estimate in Master's partial credit model (PCM) that represents the relative difficulty of each score step (e.g., going from a score of 1 to a score of 2). The higher the value of a particular step difficulty, the more difficult a particular step is relative to other score steps (e.g., is it harder to go from a 1 to a 2, or to go from a 2 to a 3).
Strand	On score reports, a strand often refers to a set of items on a test measuring the same contextual area (e.g., Number Sense in Mathematics). Items developed to measure the same reporting category would be used to determine the strand score (sometimes called "subscale" score).
Technical Advisory Committee (TAC)	A group of individuals, most often professionals in the field of testing, who are either appointed or selected to make recommendations for and to guide the technical development of a given testing program.
Validity	The degree to which accumulated evidence and theory support specific interpretations of test scores entailed by the purposed uses of a test. There are various ways of gathering validity evidence.

PSSA: The Pennsylvania System of School Assessment

The purposes of the 2010 statewide assessment component of the Pennsylvania System of School Assessment (PSSA), as specified in the Chapter 4 Regulations, include providing:

- (1) an understanding of the students' achievement of the academic standards to students, parents, educators and community citizens,
- (2) a measure of the degree to which school programs enable students to attain the academic standards,
- (3) results to school districts, charter schools and Area Vocational Technical Schools, Intermediate Units, Private Residential Rehabilitative Institutions, Approved Private Schools, non-public and private schools for use in their strategic plans,
- (4) information to the general public and state policymakers regarding school achievement of the academic standards, and
- (5) aggregate results for all students and, in compliance with federal No Child Left Behind regulations, disaggregated results for various demographic and special needs groups.

The broad purpose of the State Assessments is to provide information to teachers and schools to guide the improvement of curricula and instructional strategies to enable students to achieve the academic standards. The areas assessed in 2010 were mathematics and reading at Grades 3-8 and 11, science at Grades 4, 8, and 11, and writing at Grades 5, 8, and 11. *The Department strongly discourages the use of this testing information for "ranking" schools.*

Grade 12 PSSA Fall Retest

Chapter 4 Regulations state that students who score at the *Proficient* or *Advanced* level on the state assessments in mathematics, reading, science, and writing administered in Grade 11 or Grade 12 are eligible to receive Certificates of *Proficiency* and/or Certificates of *Distinction*. The purpose of the Grade 12 Retest is to provide students who did not achieve a *Proficient* level or higher on the Grade 11 assessments the opportunity to improve their PSSA scores and receive certificates.

The Grade 12 Retest is not a mandatory assessment, so a student may choose not to participate without parental request for exclusion and school/district officials are not required to authorize student exclusions. The Pennsylvania Department of Education (PDE) recommends schools that do not require student retest participation to encourage eligible students to discuss the retest with parents/guardians. Though the final decision about whether a student should participate in the retest is made by the student and his/her parents/guardians, the district must provide eligible students with the opportunity to participate.

A Grade 12 student is ELIGIBLE for the Grade 12 Retest if:

- Student achieved *Basic* or *Below Basic* performance level on that specific subject assessment, **OR**
- Student did not participate in the 2010 PSSA, **OR**
- Student's PSSA performance level is *unknown*, and attempts to determine student's performance level by contacting the student's former school *cannot confirm* that the student achieved *Proficient* or *Advanced* performance level.

A Student is NOT ELIGIBLE for the Grade 12 Retest if:

- Student achieved *Proficient* or *Advanced* performance level on that specific subject assessment, **OR**
- Student participated in the PASA, **OR**
- Student is not currently in Grade 12.

For each content area, only one test form was administered to all the eligible students. This technical report provides the retest results for PSSA mathematics, reading, science, and writing, including Item Analysis, Raw-to-Scaled Score Conversions, and Performance Levels results.

Item Analysis

Multiple-Choice (MC) Items

The most familiar indices of item performance for MC items are those that reflect item difficulty (i.e., *proportion correct*, generally referred to as a "*p*-value") and those that reflect item discrimination (often represented by the *point-biserial correlation* coefficient). The point-biserial correlation for an item is the Pearson product-moment correlation between students' item scores and their total test scores. It is expected that students who respond to the item correctly should have a higher total test score mean than students who respond incorrectly. An item that performs as expected should have a positive point-biserial correlation coefficient.

The item-level analyses done for the Grade 12 retests' MC items also included statistics for the incorrect responses (i.e., distractors) such as proportion of students selecting each distractor, and the point-biserial correlation for each distractor. The results from distractor analyses provide additional information for understanding the item's behavior. For example, the percent selecting each response is an indicator of which responses are particularly attractive.

Item level statistics for the MC items for mathematics, reading, science, and writing can be found in Appendices A, D, G, and J respectively. These statistics include the number of students attempting each item, *p*-values, proportions of students selecting each response, item-total correlations, and point-biserial correlations for each response category. The tabled values indicate that the MC items on the PSSA retests performed as expected.

Open-Ended (OE) Items

A first step when evaluating OE item performance is to examine the item's score-point distribution (percentages of students in each scoring category) as this can provide a rough "snapshot" of an item's performance. For example, a four-point OE item with a vast majority of

students receiving *ones* or *fours* with virtually no other scores occurring would be unusual. Another useful statistic is the correlation between the item scores and total test scores. Similar to the MC item's point-biserial index, this correlation reflects how an OE item discriminates between low scoring and high scoring students. The students with higher test scores are expected to have higher mean score on the item.

Item level statistics for the mathematics, reading, science, and writing OE items can be found in Appendices C, F, I, and L respectively. In the appendices, the "B" code denotes a blank non-response, the "F" code denotes a response in a foreign language, the "K" code denotes an off-task response, and the "U" code denotes an unreadable response. The score-point distributions and the item-total score correlations indicate that all the OE items performed as expected.

Raw-to-Scaled Score Conversions

A *scaled score*, in the simplest sense, is a transformed raw score. For the PSSA retests, this transformation was done in two steps. First, the students attempting the Grade 12 retests were scored using the Rasch scaling model by anchoring the Rasch item difficulties at the values calibrated from the 2010 spring operational data. This scoring transformed student raw scores into Rasch logit scores which typically fall between -5.0 to 5.0. This transformation is non-linear and often referred to as the "Raw-to-Logit conversion". Appendices B, E, H, and K present the anchored Rasch item logit difficulties, their corresponding standard errors, and fit statistics for all the mathematics, reading, science, and writing MC items, respectively.

The second step is to convert these logit scores into PSSA score scales using linear transformations. Table 1 gives the linear logit-to-scaled score conversion functions for Grade 12 PSSA mathematics, reading, science, and writing.

Table 1: Logit-to-Scaled Score Conversions

Content	Transformation
Mathematics	$206.42X + 1203.10$
Reading	$245.45X + 1115.20$
Science	$101.81X + 1194.69$
Writing	$100.00X + 1244.30$

Note. X denotes the Rasch logit ability values

Scaled scores have several interpretive advantages over raw scores, as illustrated in the following example. A raw score of 30, for instance, is almost meaningless unless the reader is also given how many points are possible. The same score has a different meaning if it is based on a thirty-item test or on a sixty-item test. Total points attained are transformed to percent correct scores to remove the effect of test length. In the same way, a score based on sixty difficult items is different from the same score based on sixty easy items. Total points attained are transformed to scaled scores to remove the effects of test length and item difficulty.

Since 2002, a lowest obtainable scaled score (LOSS) of 700 has been implemented for most PSSA mathematics, reading, and writing exams. One of the exceptions is PSSA science where the LOSS is 1050 for Grade 11. Scores lower than the LOSS values are converted to the LOSS value. However, the highest obtainable scale scores for PSSA tests are not fixed. They are allowed to float for each subject and grade. The RS-SS conversion tables for mathematics,

reading, science and writing can be found in Appendices N, O, P, and Q, respectively. The students' raw scores were transformed to the scaled scores based on those tables.

Summary of the Grade 12 Retest Results

Scaled Score Results

The performance of students attempting the fall retests was compared with the performance of students attempting 2010 spring operational tests. Table 2 summarizes the spring and fall test results for these two groups of students including the mean, standard deviation (SD), maximum, and minimum scaled scores as well as the reliability of the assessments. The mean scaled scores on the fall retest were lower than the mean scores on the spring test, indicating that the students who took the fall retest did not perform as well as the students who took the previous spring test. These results are expected in a retest situation since the group taking the retest is typically comprised of students who had not performed well on the previous administration.

The standard deviations were also lower for the retest group. Smaller standard deviations were the result of a more homogeneous score distribution and an artifact of the aforementioned group of retesters. The relatively lower test reliabilities (based on Coefficient Alpha) for mathematics, reading, and science can also be attributed to the decreased variability in test scores. Reliabilities for the writing inventory are given in the form of stratified alpha coefficients, and the two reliabilities were similar for operational and retest. Scorer agreement percentages for the prompt scores are provided in Appendix M. These are generally consistent with historic values.

Table 2: Operational and Retest Summary Statistics (Scaled Score Metric)

	Mathematics		Reading		Science		Writing	
	Oper.	Retest	Oper.	Retest	Oper.	Retest	Oper.	Retest
N	129910	27850	133291	22173	129926	24676	130352	8816
Mean	1372.2	1127.0	1363.2	1102.8	1242.6	1181.0	1483.7	1169.3
St. Dev.	276.0	166.9	280.4	214.6	96.7	76.5	282.3	231.9
Min	700	700	700	700	1050	1050	700	700
Max	2377	2121	2520	2215	1862	1480	2382	2382
Reli.	0.94	0.89	0.90	0.86	0.92	0.88	0.81	0.82

Figures 1, 2, 3, and 4 contrast the fall retest frequency distributions against the spring operational frequency distributions for mathematics, reading, science, and writing test scores, respectively. As seen from Figures 1, 2, and 3, the distributions of scaled scores for the fall mathematics reading, and science retests are positively-skewed relative to their operational counterparts with lower test scores occurring with much greater frequency than higher scores. In contrast, the spring operational test scores are more negatively distributed. As shown in Figure 4, both the operational and fall distributions for writing have a ‘roller-coaster’ pattern with a major mode and several minor modes. This pattern likely results from the weighting given to the writing prompts in scoring. This is described in the 2010 PSSA Technical Report.

Figure 1: Mathematics Operational and Retest Scaled Score Frequency Distributions

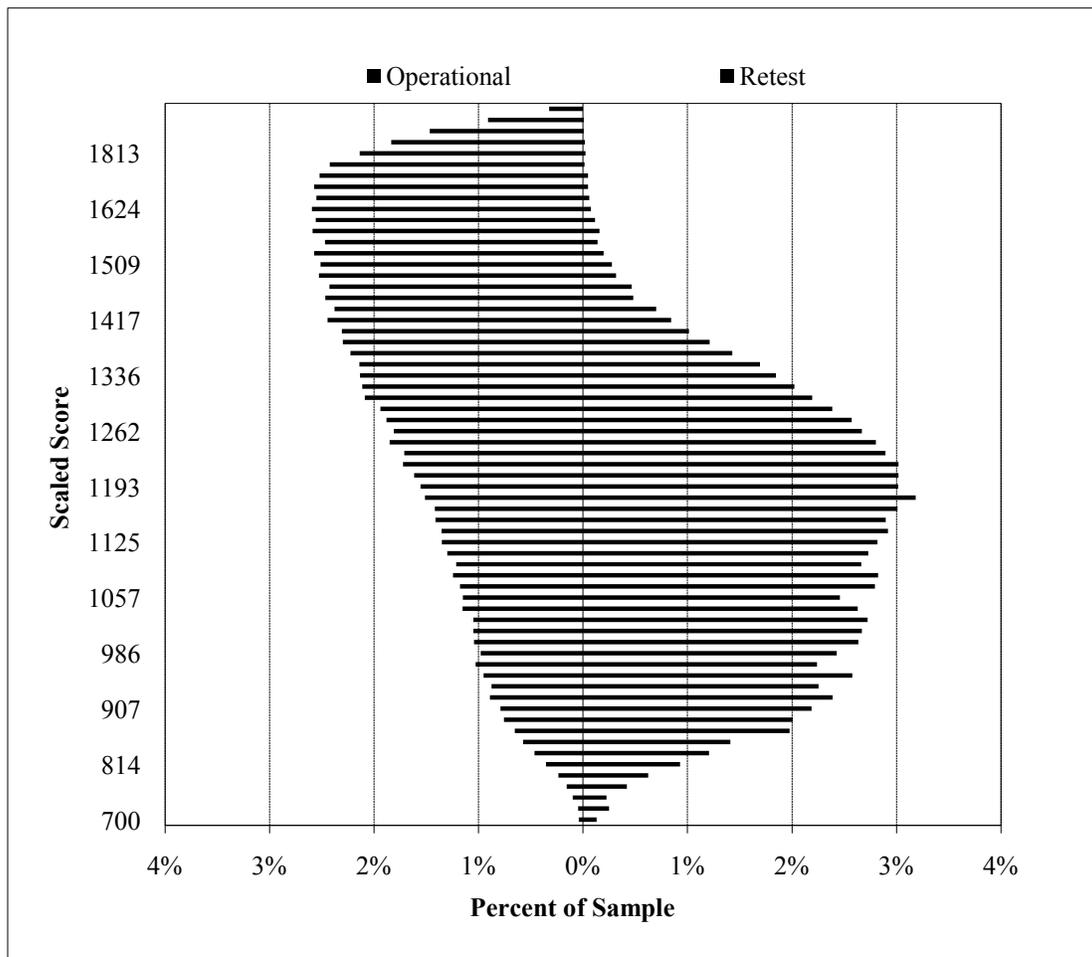


Figure 2: Reading Operational and Retest Scaled Score Frequency Distribution

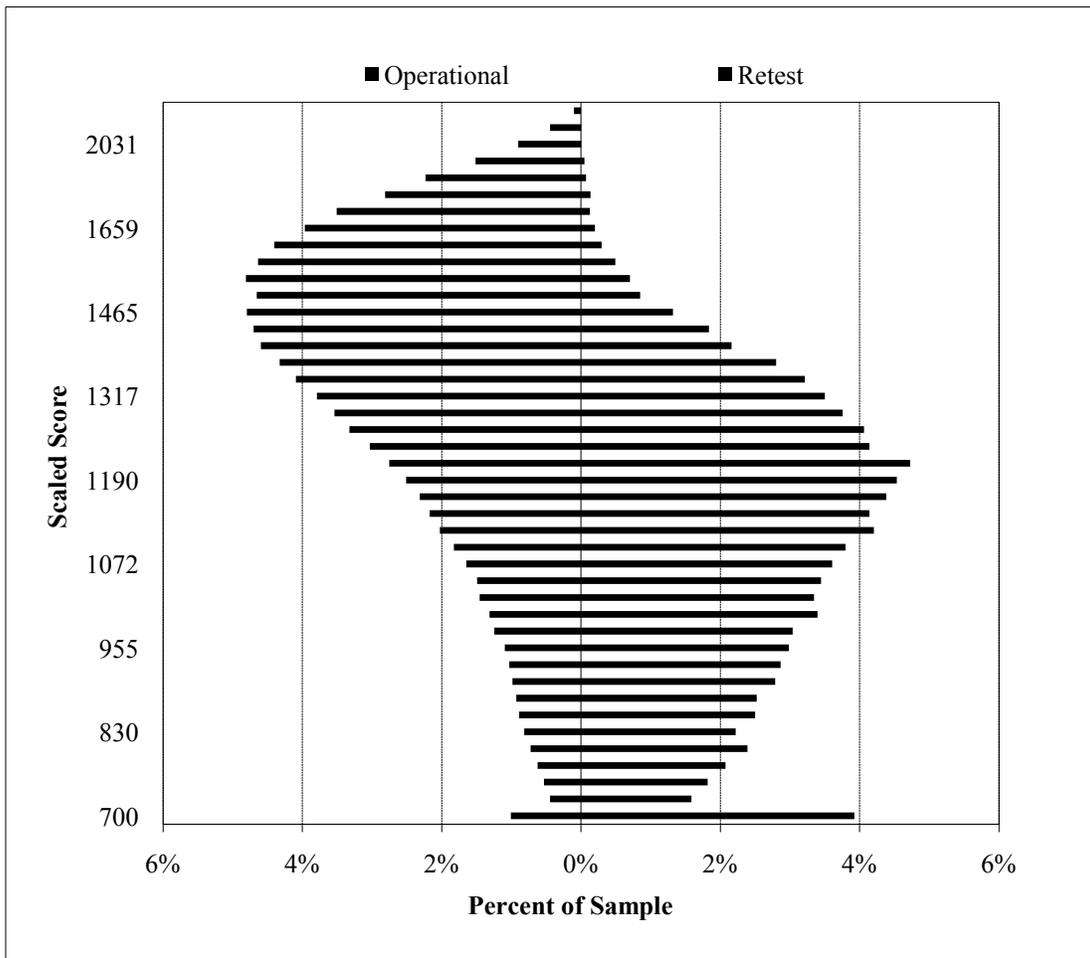


Figure 3: Science Operational and Retest Scaled Score Frequency Distribution

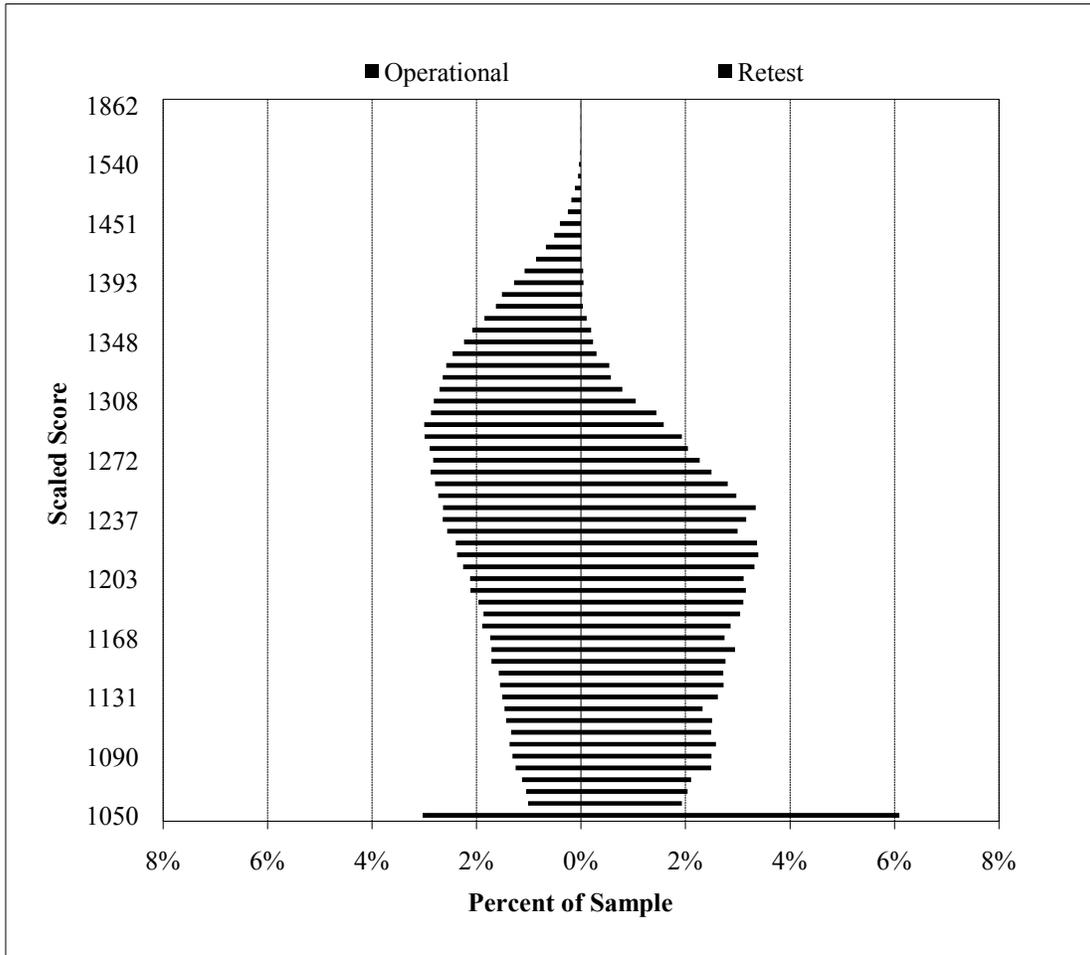
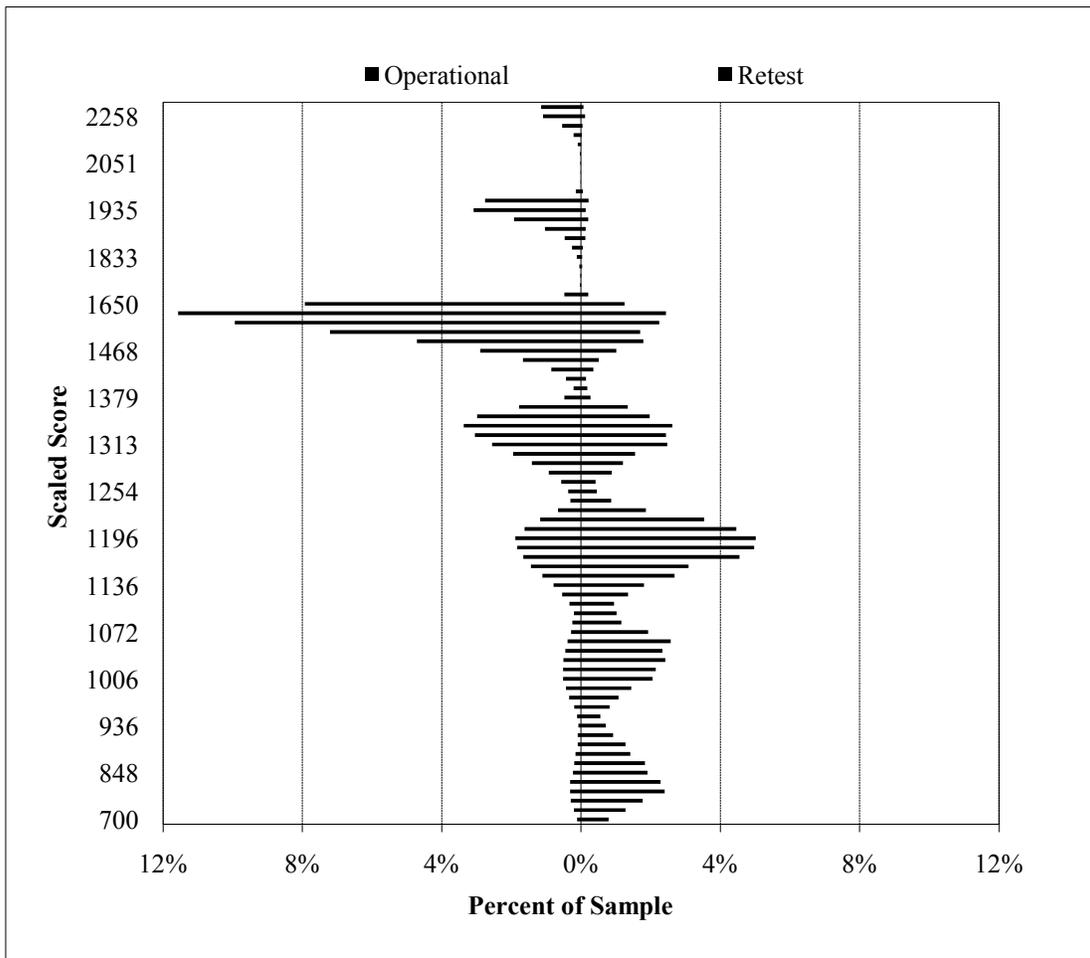


Figure 4: Writing Operational and Retest Scaled Score Frequency Distribution



Performance Level Results

Performance levels descriptors (PLDs) are another way to attach meaning to the scaled score metric. PLDs associate precise quantitative ranges of scaled scores with verbal, qualitative descriptions of student status. While much less precise, the qualitative description of the levels is one way for parents and teachers to interpret the student scores. They are also useful in assessing the status of the school. The Pennsylvania General Performance Level Descriptors, as developed by PDE and teacher panels, are given below. These are also included on student score reports.

- **Advanced:** The Advanced Level reflects superior academic performance. Advanced work indicates an in-depth understanding and exemplary display of the skills included in the Pennsylvania Academic Content Standards.
- **Proficient:** The Proficient Level reflects satisfactory academic performance. Proficient work indicates a solid understanding and adequate display of the skills included in the Pennsylvania Academic Content Standards.
- **Basic:** The Basic Level reflects marginal academic performance. Basic work indicates a partial understanding and limited display of the skills included in the Pennsylvania Academic Content Standards. This work is approaching satisfactory performance, but has not yet reached it. There is a need for additional instructional opportunities and/or increased student academic commitment to achieve the Proficient Level.
- **Below Basic:** The Below Basic Level reflects inadequate academic performance. Below Basic work indicates little understanding and minimal display of the skills included in the Pennsylvania Academic Content Standards. There is a major need for additional instructional opportunities and/or increased student academic commitment to achieve the Proficient Level.

The scores that correspond with each performance level are located in Table 3. The cumulative percentage of students who achieved a *Proficient* or *Advanced* performance level on the mathematics, reading, science, and writing retests are 15.4, 25.6, 11.0, and 29.6, respectively. (Numbers may be slightly different than tabled numbers due to rounding.) Approximately 70%-90% of the students who took the retest still scored in the *Basic* or *Below Basic* levels for each subject level.

Table 3: Grade 12 Retest Performance Standards

Mathematics			
Performance Level	Scaled Score	Frequency	Percent
Advanced	1509 and up	331	1.2
Proficient	1304-1508	3960	14.2
Basic	1167-1303	7119	25.6
Below Basic	1166 and below	16440	59.0

Reading			
Performance Level	Scaled Score	Frequency	Percent
Advanced	1492 and up	650	2.9
Proficient	1257-1491	5022	22.6
Basic	1112-1256	5793	26.1
Below Basic	1111 and below	10708	48.3

Science			
Performance Level	Scaled Score	Frequency	Percent
Advanced	1347 and up	176	0.7
Proficient	1275-1346	2528	10.2
Basic	1150-1274	13299	53.9
Below Basic	1149 and below	8673	35.1

Writing			
Performance Level	Scaled Score	Frequency	Percent
Advanced	1806 and up	111	1.3
Proficient	1236-1805	2499	28.3
Basic	952-1235	4693	53.2
Below Basic	951 and below	1513	17.2

Note. Numbers may not add exactly to 100% due to rounding.

Of the students with scores for both the spring operational and the fall retest administrations, 62.8% of the students remained at the same performance level in mathematics, while 23.2% transitioned to a higher level and 13.9% regressed to a lower level. For reading, 52.7% of the students stayed at the same level, 33.9% improved and 13.5% regressed. For science, 69.0% of the students stayed at the same level, 15.6% improved and 15.4% regressed. For writing, 56.6% of the students remained at the same level, 29.7% improved and 13.7% regressed.

Appendix A: 2010 Grade 12 Fall Mathematics Retest Multiple-Choice Item Statistics

Item Description			Proportions							Point Biserials				
Seq.	Key	N	P-Value	A	B	C	D	-	*	Tot. Corr.	A	B	C	D
1	D	27850	0.561	0.080	0.253	0.106	0.561	0.001	0.000	0.362	-0.234	-0.227	-0.054	0.362
2	A	27850	0.520	0.520	0.040	0.430	0.009	0.001	0.000	0.280	0.280	-0.172	-0.195	-0.091
3	B	27850	0.772	0.058	0.772	0.131	0.038	0.001	0.000	0.258	-0.163	0.258	-0.145	-0.105
4	B	27850	0.771	0.052	0.771	0.094	0.082	0.001	0.000	0.304	-0.137	0.304	-0.188	-0.151
5	D	27850	0.560	0.115	0.169	0.153	0.560	0.002	0.000	0.331	-0.243	-0.058	-0.178	0.331
6	C	27850	0.424	0.377	0.176	0.424	0.022	0.001	0.000	0.309	-0.176	-0.137	0.309	-0.097
7	A	27850	0.563	0.563	0.147	0.223	0.062	0.004	0.000	0.353	0.353	-0.251	-0.163	-0.069
8	B	27850	0.421	0.314	0.421	0.108	0.154	0.003	0.001	0.258	-0.059	0.258	-0.173	-0.125
9	C	27850	0.488	0.156	0.234	0.488	0.117	0.004	0.000	0.271	-0.145	-0.162	0.271	-0.042
10	A	27850	0.595	0.595	0.262	0.046	0.095	0.001	0.000	0.225	0.225	-0.029	-0.207	-0.179
11	B	27850	0.273	0.418	0.273	0.191	0.113	0.004	0.001	0.169	-0.149	0.169	-0.008	0.005
12	C	27850	0.515	0.178	0.201	0.515	0.101	0.006	0.000	0.189	0.027	-0.163	0.189	-0.128
13	D	27850	0.547	0.162	0.122	0.163	0.547	0.006	0.001	0.285	-0.068	-0.149	-0.184	0.285
14	A	27850	0.225	0.225	0.225	0.134	0.412	0.003	0.001	0.230	0.230	-0.118	-0.086	-0.033
15	B	27850	0.619	0.103	0.619	0.130	0.146	0.001	0.000	0.316	-0.235	0.316	-0.110	-0.122
16	C	27850	0.391	0.111	0.328	0.391	0.164	0.005	0.000	0.114	-0.140	0.007	0.114	-0.035
17	B	27850	0.492	0.156	0.492	0.209	0.138	0.004	0.000	0.341	-0.058	0.341	-0.256	-0.124
18	C	27850	0.494	0.177	0.243	0.494	0.083	0.003	0.000	0.241	-0.118	-0.089	0.241	-0.130
19	C	27850	0.460	0.101	0.283	0.460	0.150	0.006	0.000	0.283	-0.154	-0.145	0.283	-0.076
20	A	27850	0.744	0.744	0.076	0.080	0.098	0.002	0.000	0.337	0.337	-0.221	-0.244	-0.069
21	C	27850	0.493	0.095	0.114	0.493	0.296	0.003	0.000	0.368	-0.181	-0.169	0.368	-0.165
22	D	27850	0.731	0.121	0.088	0.058	0.731	0.002	0.000	0.355	-0.096	-0.230	-0.253	0.355
23	B	27850	0.494	0.214	0.494	0.153	0.135	0.004	0.000	0.262	-0.081	0.262	-0.162	-0.110
24	C	27850	0.494	0.182	0.157	0.494	0.164	0.003	0.000	0.392	-0.103	-0.158	0.392	-0.261
27	B	27850	0.601	0.191	0.601	0.168	0.036	0.003	0.000	0.279	-0.101	0.279	-0.222	-0.069
28	C	27850	0.435	0.404	0.081	0.435	0.079	0.001	0.000	0.166	0.065	-0.244	0.166	-0.174
29	B	27850	0.597	0.112	0.597	0.146	0.141	0.004	0.000	0.340	-0.113	0.340	-0.231	-0.137
30	D	27850	0.644	0.108	0.079	0.166	0.644	0.002	0.000	0.399	-0.183	-0.241	-0.180	0.399
31	A	27850	0.534	0.534	0.205	0.140	0.117	0.003	0.000	0.336	0.336	-0.288	-0.134	-0.011
32	A	27850	0.604	0.604	0.103	0.173	0.116	0.003	0.001	0.478	0.478	-0.202	-0.244	-0.245
33	D	27850	0.541	0.096	0.116	0.244	0.541	0.002	0.001	0.343	-0.191	-0.215	-0.101	0.343
34	D	27850	0.578	0.041	0.310	0.070	0.578	0.001	0.001	0.308	-0.205	-0.106	-0.235	0.308
35	D	27850	0.270	0.098	0.199	0.429	0.270	0.004	0.001	0.257	-0.204	-0.216	0.071	0.257
36	B	27850	0.657	0.144	0.657	0.103	0.093	0.003	0.000	0.336	-0.161	0.336	-0.092	-0.251

2010 PSSA Grade 12 Retest Technical Report for Mathematics, Reading, Science, and Writing

Item Description			Proportions							Point Biserials				
Seq.	Key	N	P-Value	A	B	C	D	-	*	Tot. Corr.	A	B	C	D
37	C	27850	0.616	0.133	0.148	0.616	0.099	0.004	0.001	0.417	-0.211	-0.218	0.417	-0.169
38	C	27850	0.487	0.192	0.231	0.487	0.085	0.004	0.000	0.356	-0.182	-0.171	0.356	-0.111
39	A	27850	0.531	0.531	0.159	0.149	0.156	0.005	0.001	0.485	0.485	-0.212	-0.276	-0.169
40	D	27850	0.437	0.076	0.359	0.123	0.437	0.004	0.001	0.260	-0.215	-0.017	-0.183	0.260
41	C	27850	0.565	0.125	0.173	0.565	0.132	0.005	0.001	0.369	-0.188	-0.203	0.369	-0.118
42	B	27850	0.652	0.177	0.652	0.096	0.070	0.004	0.001	0.396	-0.185	0.396	-0.217	-0.195
43	D	27850	0.331	0.341	0.138	0.185	0.331	0.005	0.001	0.234	-0.024	-0.108	-0.145	0.234
44	D	27850	0.656	0.089	0.127	0.123	0.656	0.005	0.000	0.456	-0.214	-0.226	-0.231	0.456
45	C	27850	0.352	0.129	0.337	0.352	0.177	0.005	0.000	0.143	-0.187	0.027	0.143	-0.037
46	D	27850	0.163	0.399	0.297	0.136	0.163	0.005	0.001	0.163	0.077	-0.090	-0.151	0.163
47	A	27850	0.715	0.715	0.069	0.145	0.066	0.005	0.001	0.480	0.480	-0.255	-0.274	-0.201
48	A	27850	0.278	0.278	0.201	0.188	0.327	0.005	0.001	0.193	0.193	-0.038	-0.153	-0.013
49	C	27850	0.484	0.187	0.240	0.484	0.084	0.005	0.001	0.228	-0.096	-0.083	0.228	-0.126
50	A	27850	0.437	0.437	0.199	0.209	0.147	0.007	0.001	0.366	0.366	-0.158	-0.157	-0.138
51	A	27850	0.444	0.444	0.143	0.107	0.299	0.005	0.001	0.309	0.309	-0.159	-0.222	-0.051
52	A	27850	0.481	0.481	0.144	0.240	0.127	0.007	0.001	0.339	0.339	-0.210	-0.114	-0.122
53	C	27850	0.608	0.155	0.155	0.608	0.076	0.006	0.001	0.342	-0.116	-0.163	0.342	-0.227
54	A	27850	0.550	0.550	0.193	0.142	0.107	0.007	0.000	0.343	0.343	-0.140	-0.229	-0.098
55	D	27850	0.420	0.159	0.287	0.126	0.420	0.007	0.001	0.392	-0.154	-0.133	-0.215	0.392
56	B	27850	0.434	0.231	0.434	0.263	0.064	0.008	0.001	0.198	-0.128	0.198	-0.009	-0.140
57	B	27850	0.435	0.155	0.435	0.227	0.175	0.007	0.001	0.301	-0.091	0.301	-0.119	-0.158
58	C	27850	0.422	0.199	0.234	0.422	0.134	0.010	0.001	0.294	-0.120	-0.171	0.294	-0.054
59	A	27850	0.384	0.384	0.196	0.230	0.181	0.008	0.001	0.334	0.334	-0.151	-0.124	-0.113
60	A	27850	0.492	0.492	0.128	0.174	0.197	0.008	0.001	0.328	0.328	-0.222	-0.198	-0.019
61	A	27850	0.576	0.576	0.136	0.216	0.064	0.007	0.001	0.308	0.308	-0.131	-0.146	-0.166
62	D	27850	0.492	0.132	0.218	0.148	0.492	0.009	0.001	0.298	-0.152	-0.107	-0.130	0.298

Note. "-" denotes omits; "*" denotes multiple marks.

Appendix B: 2010 Grade 12 Fall Mathematics Retest Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit		OutFit	
			MS	ZSTD	MS	ZSTD
1	-0.5469	0.0129	0.96	-9.9	0.95	-9.8
2	-0.3412	0.0128	1.03	7.9	1.04	6.7
3	-1.5454	0.0144	0.93	-9.9	0.93	-7.4
4	-1.4890	0.0142	0.88	-9.9	0.86	-9.9
5	-0.3659	0.0128	0.99	-1.8	0.99	-1.4
6	0.2147	0.0132	1.06	9.9	1.08	9.9
7	-0.5912	0.0129	0.97	-8.1	0.96	-7.1
8	0.3360	0.0134	1.13	9.9	1.18	9.9
9	-0.1508	0.0129	1.05	9.9	1.06	9.9
10	-0.3303	0.0128	1.09	9.9	1.11	9.9
11	1.0830	0.0153	1.26	9.9	1.45	9.9
12	-0.0708	0.0129	1.15	9.9	1.19	9.9
13	-0.5985	0.0129	1.03	6.3	1.03	5.1
14	0.9317	0.0148	0.98	-3.0	1.06	5.5
15	-0.5796	0.0129	0.98	-5.8	0.98	-4.2
16	0.3224	0.0134	1.20	9.9	1.30	9.9
17	-0.3260	0.0128	0.98	-4.2	0.98	-3.7
18	-0.0246	0.0130	1.11	9.9	1.13	9.9
19	-0.2479	0.0129	1.03	6.0	1.03	6.1
20	-1.6043	0.0146	0.96	-5.8	0.95	-4.7
21	-0.5229	0.0129	0.97	-6.3	0.96	-7.1
22	-1.4048	0.0140	0.90	-9.9	0.91	-9.9
23	-0.1375	0.0129	1.06	9.9	1.09	9.9
24	-0.5003	0.0129	0.95	-9.9	0.94	-9.9
27	-0.8243	0.0131	1.02	4.9	1.03	5.2
28	0.4164	0.0135	1.27	9.9	1.37	9.9
29	-1.0034	0.0133	1.02	3.9	1.02	2.3
30	-1.1033	0.0134	0.94	-9.9	0.92	-9.9
31	-0.4594	0.0128	0.98	-3.9	0.98	-3.2
32	-1.1893	0.0136	0.97	-6.1	0.94	-8.1
33	-0.5611	0.0129	0.98	-5.0	0.98	-4.4
34	-0.5629	0.0129	0.99	-1.3	1.00	-0.5
35	0.8148	0.0144	1.04	6.3	1.08	7.3
36	-1.2604	0.0137	1.01	2.0	1.03	3.7
37	-1.2199	0.0136	1.01	2.6	0.98	-2.2
38	-0.5421	0.0129	0.99	-2.8	0.98	-3.0
39	-0.9956	0.0133	0.99	-2.7	0.96	-5.7
40	0.1989	0.0132	1.10	9.9	1.13	9.9
41	-0.9053	0.0131	1.01	1.3	0.99	-0.8
42	-1.3679	0.0139	1.03	4.1	1.00	0.5
43	0.0637	0.0130	1.00	0.6	1.01	2.4
44	-1.4231	0.0141	1.00	-0.7	0.96	-4.2
45	0.3297	0.0134	1.13	9.9	1.19	9.9
46	1.2518	0.0160	0.93	-7.6	1.07	5.0
47	-1.9192	0.0156	1.14	9.9	1.01	0.7
48	0.6331	0.0140	1.04	5.6	1.12	9.9
49	-0.1281	0.0129	1.09	9.9	1.11	9.9
50	-0.4924	0.0129	0.99	-3.0	0.98	-3.2
51	-0.3255	0.0128	1.01	1.4	1.01	1.6
52	-0.4978	0.0129	1.00	-1.0	1.00	-0.4
53	-0.7137	0.0130	0.96	-9.9	0.95	-8.5

Seq.	Anchored Measure	Measure SE	InFit		OutFit	
			MS	ZSTD	MS	ZSTD
54	-0.5142	0.0129	0.97	-6.2	0.97	-4.8
55	-0.2846	0.0128	0.93	-9.9	0.92	-9.9
56	0.1968	0.0132	1.15	9.9	1.21	9.9
57	-0.3119	0.0128	1.01	2.5	1.02	3.0
58	-0.1980	0.0129	1.01	1.7	1.01	2.6
59	-0.0032	0.0130	0.96	-9.0	0.96	-6.4
60	-0.4128	0.0128	0.99	-1.3	0.99	-1.1
61	-0.7020	0.0130	1.00	0.2	1.01	1.7
62	-0.1015	0.0129	1.04	9.4	1.05	9.7

Appendix C: 2010 Grade 12 Fall Mathematics Retest Open-Ended Item Statistics

Item Description			Proportions									Correlations					
Seq.	Max	N	Mean	0	1	2	3	4	B	K	U	Tot. Corr.	0	1	2	3	4
25	4	27850	0.586	0.607	0.263	0.081	0.035	0.014	0.050	0.000	0.000	0.440	-0.414	0.200	0.224	0.189	0.159
26	4	27850	1.496	0.241	0.193	0.427	0.108	0.031	0.057	0.000	0.000	0.515	-0.447	-0.020	0.157	0.269	0.216
63	4	27850	0.600	0.503	0.430	0.035	0.026	0.005	0.152	0.000	0.000	0.369	-0.312	0.174	0.181	0.180	0.103

Note. B = blank; K = off task; U = unreadable.

Appendix D: 2010 Grade 12 Fall Reading Retest Multiple-Choice Item Statistics

Item Description			Proportions							Point Biserials				
Seq.	Key	N	P-Value	A	B	C	D	-	*	Tot. Corr.	A	B	C	D
1	A	22173	0.593	0.593	0.324	0.033	0.049	0.001	0.000	0.312	0.312	-0.225	-0.167	-0.084
2	B	22173	0.776	0.043	0.776	0.058	0.122	0.000	0.000	0.381	-0.204	0.381	-0.233	-0.192
3	D	22173	0.776	0.136	0.041	0.044	0.776	0.001	0.000	0.377	-0.169	-0.230	-0.249	0.377
4	A	22173	0.673	0.673	0.155	0.087	0.083	0.001	0.000	0.415	0.415	-0.259	-0.196	-0.162
5	A	22173	0.426	0.426	0.088	0.415	0.069	0.002	0.000	0.156	0.156	-0.213	0.047	-0.150
6	D	22173	0.292	0.065	0.564	0.077	0.292	0.002	0.000	0.109	-0.130	0.022	-0.103	0.109
7	C	22173	0.716	0.055	0.082	0.716	0.144	0.002	0.000	0.255	-0.254	-0.107	0.255	-0.075
9	B	22173	0.655	0.136	0.655	0.150	0.058	0.001	0.000	0.334	-0.216	0.334	-0.173	-0.096
10	C	22173	0.635	0.139	0.106	0.635	0.119	0.002	0.000	0.363	-0.155	-0.192	0.363	-0.188
11	B	22173	0.752	0.083	0.752	0.064	0.100	0.001	0.000	0.403	-0.214	0.403	-0.263	-0.163
12	D	22173	0.418	0.201	0.180	0.198	0.418	0.001	0.000	0.250	-0.156	-0.146	-0.007	0.250
13	A	22173	0.529	0.529	0.099	0.304	0.065	0.002	0.001	0.390	0.390	-0.188	-0.177	-0.220
14	C	22173	0.667	0.118	0.148	0.667	0.064	0.002	0.000	0.315	-0.173	-0.125	0.315	-0.188
15	D	22173	0.593	0.167	0.168	0.068	0.593	0.004	0.000	0.315	-0.096	-0.152	-0.240	0.315
16	C	22173	0.554	0.157	0.141	0.554	0.145	0.003	0.000	0.277	-0.196	-0.165	0.277	-0.019
17	C	22173	0.400	0.367	0.185	0.400	0.044	0.004	0.000	0.127	-0.005	-0.046	0.127	-0.190
18	A	22173	0.448	0.448	0.276	0.111	0.161	0.004	0.000	0.324	0.324	-0.166	-0.191	-0.063
19	A	22173	0.441	0.441	0.321	0.067	0.167	0.004	0.000	0.296	0.296	-0.052	-0.266	-0.143
20	D	22173	0.367	0.142	0.164	0.321	0.367	0.005	0.000	0.237	-0.158	-0.135	-0.012	0.237
21	C	22173	0.688	0.107	0.097	0.688	0.103	0.004	0.001	0.457	-0.205	-0.260	0.457	-0.219
22	A	22173	0.591	0.591	0.110	0.199	0.094	0.005	0.000	0.416	0.416	-0.249	-0.227	-0.107
23	B	22173	0.622	0.149	0.622	0.110	0.113	0.005	0.001	0.374	-0.141	0.374	-0.244	-0.159
24	B	22173	0.614	0.165	0.614	0.116	0.100	0.006	0.000	0.445	-0.167	0.445	-0.277	-0.207
26	C	22173	0.452	0.150	0.322	0.452	0.076	0.001	0.000	0.205	-0.207	0.035	0.205	-0.167
27	D	22173	0.534	0.074	0.252	0.139	0.534	0.001	0.000	0.408	-0.191	-0.145	-0.258	0.408
28	D	22173	0.628	0.116	0.118	0.135	0.628	0.002	0.001	0.453	-0.182	-0.255	-0.222	0.453
29	B	22173	0.314	0.055	0.314	0.350	0.279	0.002	0.001	0.165	-0.254	0.165	-0.098	0.067
30	D	22173	0.362	0.301	0.187	0.145	0.362	0.004	0.001	0.189	-0.024	-0.096	-0.116	0.189
31	B	22173	0.293	0.289	0.293	0.288	0.126	0.003	0.001	0.066	0.077	0.066	-0.071	-0.090
32	A	22173	0.396	0.396	0.211	0.214	0.174	0.004	0.001	0.234	0.234	-0.159	-0.041	-0.078
33	C	22173	0.511	0.193	0.092	0.511	0.200	0.003	0.001	0.323	-0.134	-0.266	0.323	-0.072
34	B	22173	0.369	0.156	0.369	0.237	0.234	0.004	0.000	0.251	-0.142	0.251	-0.161	0.001

2010 PSSA Grade 12 Retest Technical Report for Mathematics, Reading, Science, and Writing

Item Description			Proportions							Point Biserials				
Seq.	Key	N	P-Value	A	B	C	D	-	*	Tot. Corr.	A	B	C	D
36	D	22173	0.411	0.101	0.089	0.393	0.411	0.007	0.000	0.276	-0.274	-0.272	0.065	0.276
37	B	22173	0.574	0.251	0.574	0.065	0.102	0.008	0.000	0.276	0.008	0.276	-0.289	-0.202
38	C	22173	0.357	0.356	0.144	0.357	0.135	0.008	0.001	0.185	0.084	-0.227	0.185	-0.119
39	B	22173	0.475	0.141	0.475	0.249	0.125	0.009	0.001	0.316	-0.156	0.316	-0.048	-0.226
40	B	22173	0.686	0.091	0.686	0.090	0.124	0.009	0.001	0.407	-0.179	0.407	-0.280	-0.147
41	C	22173	0.357	0.093	0.322	0.357	0.218	0.009	0.001	0.192	-0.203	0.034	0.192	-0.095
42	A	22173	0.325	0.325	0.141	0.305	0.218	0.010	0.001	0.190	0.190	-0.202	-0.018	-0.004
43	B	22173	0.340	0.101	0.340	0.268	0.280	0.009	0.001	0.191	-0.133	0.191	0.019	-0.112

Note. "-" denotes omits; "*" denotes multiple marks.

Appendix E: 2010 Grade 12 Fall Reading Retest Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit		OutFit	
			MS	ZSTD	MS	ZSTD
1	-0.5445	0.0149	1.03	5.5	1.03	3.5
2	-1.6180	0.0177	0.98	-2.1	0.93	-4.6
3	-1.5706	0.0175	0.95	-5.6	0.94	-4.0
4	-0.9367	0.0155	0.93	-9.9	0.90	-9.9
5	0.4336	0.0148	1.17	9.9	1.26	9.9
6	1.5659	0.0178	1.50	9.9	1.96	9.9
7	-0.6532	0.0150	0.97	-5.0	0.99	-1.4
9	-0.7122	0.0151	0.98	-4.0	0.97	-4.0
10	-0.7395	0.0152	0.98	-2.6	0.96	-4.7
11	-1.3875	0.0168	0.93	-8.3	0.89	-8.5
12	0.7872	0.0154	1.21	9.9	1.30	9.9
13	-0.3964	0.0147	0.98	-3.6	0.98	-3.4
14	-0.6691	0.0150	0.97	-5.4	0.98	-2.0
15	-0.4452	0.0148	1.01	1.9	1.02	2.6
16	-0.0504	0.0146	1.06	9.9	1.07	9.9
17	0.5408	0.0150	1.19	9.9	1.31	9.9
18	0.1084	0.0146	1.00	0.3	1.01	1.1
19	0.0426	0.0146	1.03	5.3	1.04	5.5
20	0.5585	0.0150	1.06	9.9	1.11	9.9
21	-1.0532	0.0158	0.91	-9.9	0.86	-9.9
22	-0.6514	0.0150	0.97	-5.9	0.94	-7.1
23	-0.6618	0.0150	0.97	-5.2	0.96	-5.2
24	-0.7661	0.0152	0.94	-9.0	0.92	-9.7
26	0.4966	0.0149	1.17	9.9	1.25	9.9
27	-0.3129	0.0147	0.95	-9.9	0.95	-7.2
28	-1.0552	0.0158	1.03	4.0	1.00	0.2
29	1.1613	0.0163	1.23	9.9	1.48	9.9
30	0.6175	0.0151	1.11	9.9	1.19	9.9
31	1.5085	0.0176	1.48	9.9	2.04	9.9
32	0.4968	0.0149	1.08	9.9	1.14	9.9
33	-0.3874	0.0147	1.05	9.5	1.05	7.5
34	0.4692	0.0149	1.03	6.3	1.08	9.8
36	0.4349	0.0148	1.06	9.9	1.08	9.5
37	-0.3217	0.0147	1.05	8.7	1.06	7.8
38	0.7404	0.0153	1.13	9.9	1.26	9.9
39	0.0427	0.0146	1.01	2.6	1.03	5.0
40	-1.4103	0.0169	1.14	9.9	1.11	8.2
41	0.6401	0.0151	1.09	9.9	1.21	9.9
42	0.7173	0.0152	1.07	9.9	1.17	9.9
43	0.7191	0.0152	1.09	9.9	1.20	9.9

Appendix F: 2010 Grade 12 Fall Reading Retest Open-ended Item Statistics

Item Description			Proportions									Correlations				
Seq.	Max	N	Mean	0	1	2	3	B	F	K	U	Tot. Corr.	0	1	2	3
8	3	22173	1.549	0.110	0.297	0.527	0.066	0.034	0.000	0.001	0.000	0.535	-0.398	-0.236	0.351	0.230
25	3	22173	1.117	0.228	0.454	0.293	0.026	0.072	0.000	0.012	0.000	0.567	-0.504	0.017	0.383	0.181
35	3	22173	1.303	0.212	0.347	0.367	0.074	0.066	0.000	0.002	0.000	0.599	-0.488	-0.125	0.387	0.276
44	3	22173	1.501	0.163	0.304	0.403	0.130	0.073	0.000	0.002	0.000	0.557	-0.451	-0.175	0.307	0.286

Note.: B = blank; F = foreign language; K = off task; U = unreadable.

Appendix G: 2010 Grade 12 Fall Science Retest Multiple-Choice Item Statistics

Item Description			Proportions							Point Biserials				
Seq.	Key	N	P-Value	A	B	C	D	-	*	Tot. Corr.	A	B	C	D
1	A	24676	0.577	0.577	0.069	0.293	0.059	0.002	0.000	0.262	0.262	-0.115	-0.180	-0.068
2	B	24676	0.655	0.154	0.655	0.115	0.073	0.002	0.000	0.267	-0.048	0.267	-0.239	-0.120
3	C	24676	0.670	0.061	0.219	0.670	0.048	0.001	0.000	0.294	-0.141	-0.165	0.294	-0.164
4	B	24676	0.802	0.036	0.802	0.043	0.118	0.001	0.000	0.343	-0.166	0.343	-0.190	-0.205
5	B	24676	0.572	0.142	0.572	0.189	0.095	0.002	0.000	0.320	-0.121	0.320	-0.157	-0.180
6	C	24676	0.674	0.100	0.110	0.674	0.115	0.002	0.000	0.355	-0.207	-0.193	0.355	-0.134
7	D	24676	0.827	0.113	0.029	0.029	0.827	0.001	0.000	0.327	-0.184	-0.191	-0.187	0.327
8	A	24676	0.791	0.791	0.098	0.052	0.058	0.001	0.000	0.305	0.305	-0.180	-0.188	-0.119
9	B	24676	0.496	0.149	0.496	0.225	0.127	0.002	0.000	0.273	-0.162	0.273	-0.117	-0.090
10	A	24676	0.711	0.711	0.069	0.147	0.070	0.002	0.000	0.399	0.399	-0.175	-0.261	-0.165
11	D	24676	0.551	0.109	0.131	0.202	0.551	0.007	0.000	0.362	-0.182	-0.179	-0.145	0.362
12	C	24676	0.775	0.074	0.091	0.775	0.059	0.001	0.000	0.379	-0.228	-0.230	0.379	-0.135
13	C	24676	0.592	0.285	0.065	0.592	0.057	0.001	0.000	0.282	-0.058	-0.257	0.282	-0.208
14	D	24676	0.469	0.119	0.214	0.192	0.469	0.007	0.000	0.290	-0.114	-0.071	-0.193	0.290
16	A	24676	0.556	0.556	0.126	0.267	0.049	0.002	0.000	0.028	0.028	-0.215	0.167	-0.069
17	D	24676	0.421	0.119	0.192	0.268	0.421	0.001	0.000	0.262	-0.027	-0.116	-0.166	0.262
18	D	24676	0.390	0.126	0.209	0.271	0.390	0.004	0.000	0.287	-0.137	-0.177	-0.046	0.287
19	B	24676	0.532	0.096	0.532	0.226	0.143	0.003	0.000	0.337	-0.185	0.337	-0.152	-0.132
21	A	24676	0.406	0.406	0.201	0.184	0.186	0.021	0.001	0.339	0.339	-0.074	-0.093	-0.233
22	D	24676	0.351	0.148	0.182	0.315	0.351	0.004	0.000	0.301	-0.125	-0.274	0.023	0.301
23	A	24676	0.440	0.440	0.139	0.211	0.194	0.016	0.000	0.233	0.233	-0.172	-0.124	0.006
25	A	24676	0.507	0.507	0.202	0.167	0.114	0.010	0.000	0.387	0.387	-0.222	-0.221	-0.060
26	C	24676	0.619	0.134	0.127	0.619	0.117	0.003	0.000	0.335	-0.139	-0.229	0.335	-0.112
27	A	24676	0.563	0.563	0.123	0.252	0.059	0.003	0.000	0.238	0.238	-0.179	-0.035	-0.179
28	C	24676	0.633	0.108	0.184	0.633	0.069	0.006	0.000	0.322	-0.262	-0.066	0.322	-0.175
29	A	24676	0.607	0.607	0.115	0.110	0.166	0.002	0.000	0.456	0.456	-0.283	-0.286	-0.110
30	C	24676	0.577	0.139	0.209	0.577	0.070	0.005	0.000	0.299	-0.221	-0.013	0.299	-0.249
31	D	24676	0.556	0.145	0.128	0.170	0.556	0.001	0.000	0.404	-0.136	-0.267	-0.165	0.404
32	C	24676	0.455	0.164	0.082	0.455	0.296	0.002	0.000	0.276	-0.107	-0.206	0.276	-0.087
33	D	24676	0.405	0.194	0.283	0.115	0.405	0.002	0.000	0.360	-0.179	-0.094	-0.193	0.360
34	C	24676	0.605	0.151	0.149	0.605	0.092	0.003	0.000	0.462	-0.165	-0.242	0.462	-0.270
35	D	24676	0.380	0.264	0.210	0.136	0.380	0.010	0.000	0.319	-0.137	-0.086	-0.172	0.319
36	A	24676	0.527	0.527	0.237	0.188	0.046	0.002	0.000	0.318	0.318	-0.128	-0.178	-0.159

2010 PSSA Grade 12 Retest Technical Report for Mathematics, Reading, Science, and Writing

Item Description			Proportions							Point Biserials				
Seq.	Key	N	P-Value	A	B	C	D	-	*	Tot. Corr.	A	B	C	D
37	A	24676	0.310	0.310	0.260	0.284	0.145	0.002	0.000	0.114	0.114	-0.088	0.002	-0.040
41	B	24676	0.316	0.321	0.316	0.148	0.209	0.006	0.000	0.000	0.017	0.000	-0.137	0.105
42	B	24676	0.314	0.216	0.314	0.238	0.225	0.007	0.000	0.101	-0.039	0.101	-0.104	0.040
43	D	24676	0.373	0.158	0.238	0.225	0.373	0.006	0.000	0.223	-0.100	-0.058	-0.105	0.223
44	C	24676	0.384	0.157	0.336	0.384	0.116	0.007	0.000	0.139	-0.213	0.171	0.139	-0.212
46	A	24676	0.256	0.256	0.209	0.206	0.328	0.001	0.000	0.144	0.144	-0.204	-0.214	0.229
47	C	24676	0.690	0.128	0.092	0.690	0.085	0.004	0.000	0.447	-0.190	-0.242	0.447	-0.251
48	D	24676	0.530	0.135	0.148	0.176	0.530	0.011	0.000	0.434	-0.243	-0.179	-0.178	0.434
49	D	24676	0.371	0.176	0.192	0.259	0.371	0.001	0.000	0.225	-0.142	-0.076	-0.056	0.225
50	B	24676	0.474	0.118	0.474	0.171	0.234	0.002	0.000	0.356	-0.266	0.356	-0.209	-0.029
51	C	24676	0.730	0.083	0.129	0.730	0.056	0.002	0.000	0.403	-0.305	-0.195	0.403	-0.121
52	B	24676	0.457	0.147	0.457	0.326	0.066	0.004	0.000	0.187	-0.043	0.187	-0.091	-0.138
53	D	24676	0.381	0.171	0.284	0.160	0.381	0.004	0.000	0.294	-0.150	-0.063	-0.153	0.294
54	B	24676	0.401	0.184	0.401	0.294	0.113	0.008	0.000	0.278	-0.157	0.278	-0.062	-0.137
55	C	24676	0.460	0.262	0.151	0.460	0.122	0.005	0.000	0.234	0.034	-0.157	0.234	-0.222
56	B	24676	0.592	0.163	0.592	0.119	0.120	0.005	0.000	0.391	-0.176	0.391	-0.236	-0.151
57	B	24676	0.797	0.118	0.797	0.061	0.020	0.004	0.000	0.249	-0.190	0.249	-0.082	-0.113

Note. "-" denotes omits; "*" denotes multiple marks.

Appendix H: 2010 Grade 12 Fall Science Retest Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit		OutFit	
			MS	ZSTD	MS	ZSTD
1	-0.3757	0.0137	1.03	6.3	1.03	5.5
2	-0.8435	0.0142	1.02	3.4	1.02	3.2
3	-0.7569	0.0141	0.96	-6.8	0.95	-7.1
4	-1.4300	0.0157	0.82	-9.9	0.77	-9.9
5	-0.3834	0.0137	0.99	-2.9	0.98	-3.4
6	-0.9512	0.0144	0.95	-8.2	0.93	-8.9
7	-1.9223	0.0177	0.94	-5.4	0.88	-7.9
8	-1.5717	0.0162	0.92	-9.8	0.93	-5.9
9	-0.1236	0.0136	1.03	6.8	1.04	6.9
10	-1.0949	0.0147	0.89	-9.9	0.86	-9.9
11	-0.3765	0.0137	0.96	-8.9	0.95	-8.9
12	-1.5286	0.0160	0.90	-9.9	0.85	-9.9
13	-0.5203	0.0138	1.01	3.0	1.02	3.1
14	0.0291	0.0136	1.02	3.6	1.02	3.8
16	0.0474	0.0136	1.26	9.9	1.32	9.9
17	0.2192	0.0137	1.03	6.9	1.04	6.9
18	0.3701	0.0139	1.01	1.8	1.02	3.5
19	-0.3004	0.0136	0.98	-3.8	0.98	-3.9
21	0.2435	0.0138	0.97	-7.5	0.96	-6.8
22	0.7535	0.0145	1.06	9.4	1.06	6.7
23	0.1545	0.0137	1.06	9.9	1.09	9.9
25	-0.1931	0.0136	0.94	-9.9	0.94	-9.9
26	-0.7998	0.0141	1.00	0.2	1.00	-0.3
27	-0.4284	0.0137	1.06	9.9	1.07	9.9
28	-0.8991	0.0143	1.01	2.4	1.04	4.9
29	-0.7418	0.0141	0.90	-9.9	0.89	-9.9
30	-0.2309	0.0136	1.00	-0.1	1.01	1.1
31	-0.6755	0.0140	0.98	-4.0	0.98	-3.2
32	0.0724	0.0136	1.02	5.5	1.04	7.6
33	0.2732	0.0138	0.95	-9.9	0.95	-8.4
34	-0.7144	0.0140	0.90	-9.9	0.87	-9.9
35	0.3116	0.0138	0.97	-7.0	0.98	-3.6
36	-0.0573	0.0136	1.00	0.3	1.01	2.3
37	0.8647	0.0148	1.15	9.9	1.33	9.9
41	0.9856	0.0151	1.30	9.9	1.60	9.9
42	1.0060	0.0152	1.24	9.9	1.43	9.9
43	0.6051	0.0142	1.09	9.9	1.16	9.9
44	0.8893	0.0149	1.32	9.9	1.46	9.9
46	1.0745	0.0154	1.08	9.9	1.22	9.9
47	-1.4031	0.0156	1.04	5.3	0.98	-1.7
48	-0.6813	0.0140	0.99	-1.5	0.98	-3.0
49	0.4485	0.0140	1.05	9.5	1.09	9.9
50	-0.0999	0.0136	0.96	-8.2	0.96	-6.9
51	-1.5420	0.0161	1.04	4.2	0.99	-0.8
52	0.1358	0.0137	1.10	9.9	1.12	9.9
53	0.2796	0.0138	0.98	-3.3	0.99	-1.9
54	0.3090	0.0138	1.01	3.1	1.04	5.9
55	-0.0292	0.0136	1.06	9.9	1.07	9.9
56	-0.8255	0.0142	1.01	1.1	0.99	-1.5
57	-1.4716	0.0158	0.92	-9.9	0.91	-7.9

Appendix I: 2010 Grade 12 Fall Science Retest Open-Ended Item Statistics

Item Description			Proportions										Correlations					
Seq.	Max	N	Mean	0	1	2	3	4	B	F	K	U	Tot. Corr.	0	1	2	3	4
15	2	24676	0.316	0.727	0.230	0.043			0.123	0.000	0.002	0.000	0.274	-0.257	0.186	0.178		
20	4	24676	0.664	0.477	0.397	0.113	0.011	0.002	0.164	0.000	0.003	0.000	0.415	-0.377	0.185	0.262	0.120	0.064
24	4	24676	0.541	0.614	0.263	0.094	0.025	0.004	0.155	0.000	0.003	0.000	0.459	-0.463	0.273	0.257	0.162	0.076
38	2	24676	0.793	0.424	0.359	0.217			0.119	0.000	0.005	0.000	0.480	-0.478	0.211	0.328		
39	2	24676	0.536	0.544	0.377	0.080			0.121	0.000	0.002	0.000	0.316	-0.307	0.213	0.182		
40	2	24676	0.356	0.716	0.211	0.073			0.146	0.000	0.003	0.000	0.361	-0.371	0.278	0.207		
45	4	24676	0.659	0.524	0.303	0.162	0.010	0.000	0.129	0.000	0.002	0.000	0.397	-0.389	0.194	0.253	0.110	0.024
58	2	24676	0.429	0.655	0.262	0.083			0.134	0.000	0.009	0.000	0.378	-0.378	0.267	0.226		
59	2	24676	0.318	0.715	0.252	0.033			0.186	0.000	0.003	0.000	0.353	-0.352	0.299	0.162		

Note. B = blank; F = foreign language; K = off task; U = unreadable.

Appendix J: 2010 Grade 12 Fall Writing Retest Multiple-Choice Item Statistics

Item Description			Proportions							Point Biserials				
Seq.	Key	N	P-Value	A	B	C	D	-	*	Tot. Corr.	A	B	C	D
1	B	8816	0.704	0.111	0.704	0.159	0.025	0.002	0.000	0.219	-0.118	0.219	-0.142	-0.068
2	D	8816	0.549	0.357	0.032	0.060	0.549	0.002	0.000	0.299	-0.177	-0.143	-0.161	0.299
3	A	8816	0.877	0.877	0.040	0.046	0.035	0.002	0.000	0.376	0.376	-0.230	-0.219	-0.169
4	D	8816	0.583	0.064	0.299	0.051	0.583	0.003	0.000	0.348	-0.191	-0.191	-0.165	0.348
5	D	8816	0.605	0.086	0.051	0.257	0.605	0.002	0.000	0.268	-0.173	-0.267	-0.053	0.268
6	A	8816	0.486	0.486	0.156	0.239	0.112	0.007	0.000	0.339	0.339	-0.171	-0.154	-0.124
7	A	8816	0.646	0.646	0.203	0.106	0.043	0.002	0.000	0.317	0.317	-0.115	-0.218	-0.185
8	C	8816	0.728	0.124	0.053	0.728	0.093	0.003	0.000	0.374	-0.168	-0.142	0.374	-0.268
9	A	8816	0.336	0.336	0.328	0.289	0.046	0.002	0.000	0.223	0.223	-0.136	-0.049	-0.088
10	B	8816	0.703	0.091	0.703	0.082	0.121	0.002	0.000	0.430	-0.221	0.430	-0.276	-0.172
11	C	8816	0.653	0.122	0.077	0.653	0.145	0.004	0.000	0.380	-0.202	-0.249	0.380	-0.134
12	C	8816	0.716	0.139	0.093	0.716	0.049	0.003	0.000	0.418	-0.209	-0.226	0.418	-0.226

Note. "-" denotes omits; "*" denotes multiple marks.

Appendix K: 2010 Grade 12 Fall Writing Retest Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit		OutFit	
			MS	ZSTD	MS	ZSTD
1	-0.0775	0.0282	2.45	9.9	9.90	9.9
2	0.8651	0.0314	2.99	9.9	9.90	9.9
3	-2.9594	0.0333	1.23	9.9	3.28	4.6
4	0.3714	0.0294	2.42	9.9	9.90	9.9
5	0.4654	0.0298	2.76	9.9	9.90	9.9
6	1.2947	0.0338	3.26	9.9	9.90	9.9
7	-0.1281	0.0281	2.17	9.9	6.24	9.9
8	-0.6803	0.0276	1.82	9.9	4.64	9.9
9	2.3846	0.0415	4.15	9.9	9.90	9.9
10	-0.7961	0.0276	1.68	9.9	3.89	9.9
11	-0.3402	0.0278	1.92	9.9	8.63	9.9
12	-1.0830	0.0277	1.59	9.9	5.33	9.9

Appendix L: 2010 Grade 12 Fall Writing Retest Prompt Statistics

Item Description			Proportions					Correlations				
Seq.	Max	N	Mean	1	2	3	4	Tot. Corr.	1	2	3	4
13	4	8840	2.265	0.096	0.560	0.328	0.016	0.597	-0.384	-0.267	0.465	0.218
13	4	8840	2.287	0.070	0.593	0.318	0.020	0.646	-0.371	-0.362	0.513	0.241
14	4	8840	2.011	0.196	0.601	0.198	0.005	0.550	-0.423	-0.017	0.419	0.137
14	4	8840	2.138	0.114	0.643	0.234	0.009	0.627	-0.416	-0.201	0.500	0.185

Appendix M: 2010 Grade 12 Fall Writing Retest Percentage Agreement

Prompt	Composition			Revising & Editing		
	% Exact Agreement	% Adjacent Agreement	% Exact + Adjacent	% Exact Agreement	% Adjacent Agreement	% Exact + Adjacent
1	89	11	100	86	14	100
2	89	11	100	87	13	100

Appendix N: 2010 Grade 12 Fall Mathematics Retest Raw-to-Scaled Score Conversion Table

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
0	-5.9544	1.8319	700	378	1	0.0	1	0.0	1
1	-4.7345	1.0114	700	209	0	0.0	1	0.0	1
2	-4.0185	0.7232	700	149	0	0.0	1	0.0	1
3	-3.5900	0.5970	700	123	0	0.0	1	0.0	1
4	-3.2792	0.5227	700	108	0	0.0	1	0.0	1
5	-3.0329	0.4725	700	98	1	0.0	2	0.0	1
6	-2.8271	0.4360	700	90	2	0.0	4	0.0	1
7	-2.6495	0.4080	700	84	5	0.0	9	0.0	1
8	-2.4922	0.3857	700	80	27	0.1	36	0.1	1
9	-2.3506	0.3675	718	76	69	0.2	105	0.4	1
10	-2.2212	0.3523	745	73	63	0.2	168	0.6	1
11	-2.1017	0.3395	769	70	117	0.4	285	1.0	1
12	-1.9902	0.3285	792	68	174	0.6	459	1.6	1
13	-1.8855	0.3190	814	66	259	0.9	718	2.6	2
14	-1.7865	0.3106	834	64	336	1.2	1054	3.8	3
15	-1.6923	0.3033	854	63	393	1.4	1447	5.2	4
16	-1.6022	0.2969	872	61	551	2.0	1998	7.2	6
17	-1.5158	0.2912	890	60	559	2.0	2557	9.2	8
18	-1.4325	0.2861	907	59	609	2.2	3166	11.4	10
19	-1.3520	0.2816	924	58	665	2.4	3831	13.8	13
20	-1.2739	0.2776	940	57	628	2.3	4459	16.0	15
21	-1.1978	0.2740	956	57	718	2.6	5177	18.6	17
22	-1.1236	0.2709	971	56	624	2.2	5801	20.8	20
23	-1.0510	0.2681	986	55	676	2.4	6477	23.3	22
24	-0.9798	0.2656	1001	55	734	2.6	7211	25.9	25
25	-0.9098	0.2635	1015	54	743	2.7	7954	28.6	27
26	-0.8409	0.2616	1030	54	759	2.7	8713	31.3	30
27	-0.7729	0.2600	1044	54	732	2.6	9445	33.9	33
28	-0.7056	0.2587	1057	53	685	2.5	10130	36.4	35
29	-0.6390	0.2576	1071	53	778	2.8	10908	39.2	38
30	-0.5728	0.2568	1085	53	787	2.8	11695	42.0	41
31	-0.5071	0.2561	1098	53	742	2.7	12437	44.7	43
32	-0.4416	0.2557	1112	53	760	2.7	13197	47.4	46
33	-0.3763	0.2554	1125	53	785	2.8	13982	50.2	49
34	-0.3110	0.2554	1139	53	813	2.9	14795	53.1	52
35	-0.2458	0.2555	1152	53	807	2.9	15602	56.0	55
36	-0.1805	0.2558	1166	53	838	3.0	16440	59.0	58
37	-0.1149	0.2563	1179	53	887	3.2	17327	62.2	61
38	-0.0491	0.2570	1193	53	840	3.0	18167	65.2	64
39	0.0172	0.2578	1207	53	841	3.0	19008	68.3	67
40	0.0839	0.2588	1220	53	841	3.0	19849	71.3	70
41	0.1511	0.2599	1234	54	806	2.9	20655	74.2	73
42	0.2190	0.2612	1248	54	781	2.8	21436	77.0	76
43	0.2876	0.2627	1262	54	743	2.7	22179	79.6	78
44	0.3570	0.2643	1277	55	716	2.6	22895	82.2	81
45	0.4273	0.2661	1291	55	664	2.4	23559	84.6	83
46	0.4987	0.2680	1306	55	611	2.2	24170	86.8	86
47	0.5710	0.2701	1321	56	563	2.0	24733	88.8	88
48	0.6446	0.2724	1336	56	514	1.8	25247	90.7	90
49	0.7194	0.2749	1352	57	472	1.7	25719	92.3	92
50	0.7957	0.2776	1367	57	398	1.4	26117	93.8	93

2010 PSSA Grade 12 Retest Technical Report for Mathematics, Reading, Science, and Writing

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
51	0.8736	0.2804	1383	58	338	1.2	26455	95.0	94
52	0.9531	0.2836	1400	59	282	1.0	26737	96.0	95
53	1.0345	0.2870	1417	59	235	0.8	26972	96.8	96
54	1.1180	0.2908	1434	60	195	0.7	27167	97.5	97
55	1.2038	0.2951	1452	61	134	0.5	27301	98.0	98
56	1.2923	0.2998	1470	62	130	0.5	27431	98.5	98
57	1.3837	0.3052	1489	63	88	0.3	27519	98.8	99
58	1.4788	0.3114	1509	64	77	0.3	27596	99.1	99
59	1.5779	0.3186	1529	66	55	0.2	27651	99.3	99
60	1.6821	0.3270	1550	67	39	0.1	27690	99.4	99
61	1.7922	0.3371	1573	70	44	0.2	27734	99.6	99
62	1.9098	0.3491	1597	72	32	0.1	27766	99.7	99
63	2.0367	0.3637	1624	75	21	0.1	27787	99.8	99
64	2.1753	0.3817	1652	79	17	0.1	27804	99.8	99
65	2.3295	0.4043	1684	83	13	0.0	27817	99.9	99
66	2.5044	0.4332	1720	89	13	0.0	27830	99.9	99
67	2.7083	0.4715	1762	97	4	0.0	27834	99.9	99
68	2.9549	0.5244	1813	108	7	0.0	27841	100.0	99
69	3.2698	0.6025	1878	124	5	0.0	27846	100.0	99
70	3.7085	0.7332	1969	151	2	0.0	27848	100.0	99
71	4.4453	1.0250	2121	212	2	0.0	27850	100.0	99
72	5.6890	1.8430	2377	380	0	0.0	27850	100.0	100

Appendix O: 2010 Grade 12 Fall Reading Retest Raw-to-Scaled Score Conversion Table

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
0	-5.5882	1.8340	700	450	0	0.0	0	0.0	0
1	-4.3628	1.0152	700	249	0	0.0	0	0.0	0
2	-3.6387	0.7288	700	179	2	0.0	2	0.0	1
3	-3.2018	0.6042	700	148	3	0.0	5	0.0	1
4	-2.8821	0.5313	700	130	11	0.0	16	0.1	1
5	-2.6264	0.4826	700	118	21	0.1	37	0.2	1
6	-2.4108	0.4475	700	110	44	0.2	81	0.4	1
7	-2.2227	0.4209	700	103	115	0.5	196	0.9	1
8	-2.0546	0.4000	700	98	154	0.7	350	1.6	1
9	-1.9014	0.3833	700	94	215	1.0	565	2.5	2
10	-1.7598	0.3696	700	91	305	1.4	870	3.9	3
11	-1.6275	0.3583	716	88	351	1.6	1221	5.5	5
12	-1.5026	0.3488	746	86	403	1.8	1624	7.3	6
13	-1.3838	0.3409	776	84	460	2.1	2084	9.4	8
14	-1.2700	0.3342	803	82	530	2.4	2614	11.8	11
15	-1.1602	0.3285	830	81	492	2.2	3106	14.0	13
16	-1.0539	0.3238	857	79	554	2.5	3660	16.5	15
17	-0.9503	0.3198	882	78	559	2.5	4219	19.0	18
18	-0.8491	0.3166	907	78	618	2.8	4837	21.8	20
19	-0.7498	0.3139	931	77	635	2.9	5472	24.7	23
20	-0.6520	0.3117	955	77	662	3.0	6134	27.7	26
21	-0.5553	0.3101	979	76	674	3.0	6808	30.7	29
22	-0.4595	0.3089	1002	76	753	3.4	7561	34.1	32
23	-0.3643	0.3082	1026	76	742	3.3	8303	37.4	36
24	-0.2695	0.3079	1049	76	764	3.4	9067	40.9	39
25	-0.1747	0.3079	1072	76	799	3.6	9866	44.5	43
26	-0.0797	0.3084	1096	76	842	3.8	10708	48.3	46
27	0.0156	0.3092	1119	76	932	4.2	11640	52.5	50
28	0.1115	0.3104	1143	76	918	4.1	12558	56.6	55
29	0.2084	0.3120	1166	77	972	4.4	13530	61.0	59
30	0.3063	0.3140	1190	77	1005	4.5	14535	65.6	63
31	0.4056	0.3164	1215	78	1048	4.7	15583	70.3	68
32	0.5066	0.3193	1240	78	918	4.1	16501	74.4	72
33	0.6097	0.3227	1265	79	901	4.1	17402	78.5	76
34	0.7150	0.3266	1291	80	833	3.8	18235	82.2	80
35	0.8231	0.3311	1317	81	776	3.5	19011	85.7	84
36	0.9344	0.3362	1345	83	712	3.2	19723	89.0	87
37	1.0493	0.3421	1373	84	621	2.8	20344	91.8	90
38	1.1686	0.3488	1402	86	479	2.2	20823	93.9	93
39	1.2930	0.3566	1433	88	407	1.8	21230	95.7	95
40	1.4232	0.3655	1465	90	293	1.3	21523	97.1	96
41	1.5606	0.3759	1498	92	188	0.8	21711	97.9	97
42	1.7064	0.3881	1534	95	156	0.7	21867	98.6	98
43	1.8625	0.4024	1572	99	109	0.5	21976	99.1	99
44	2.0312	0.4197	1614	103	66	0.3	22042	99.4	99
45	2.2161	0.4409	1659	108	44	0.2	22086	99.6	99
46	2.4220	0.4676	1710	115	28	0.1	22114	99.7	99
47	2.6565	0.5024	1767	123	30	0.1	22144	99.9	99
48	2.9323	0.5504	1835	135	16	0.1	22160	99.9	99
49	3.2732	0.6219	1919	153	11	0.0	22171	100.0	99
50	3.7323	0.7442	2031	183	1	0.0	22172	100.0	99

2010 PSSA Grade 12 Retest Technical Report for Mathematics, Reading, Science, and Writing

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
51	4.4794	1.0267	2215	252	1	0.0	22173	100.0	99
52	5.7214	1.8406	2520	452	0	0.0	22173	100.0	100

Appendix P: 2010 Grade 12 Fall Science Retest Raw-to-Scaled Score Conversion Table

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
0	-5.7383	1.8337	1050	187	1	0.0	1	0.0	1
1	-4.5140	1.0144	1050	103	0	0.0	1	0.0	1
2	-3.7916	0.7275	1050	74	0	0.0	1	0.0	1
3	-3.3569	0.6022	1050	61	0	0.0	1	0.0	1
4	-3.0399	0.5286	1050	54	2	0.0	3	0.0	1
5	-2.7873	0.4790	1050	49	6	0.0	9	0.0	1
6	-2.5754	0.4430	1050	45	11	0.0	20	0.1	1
7	-2.3916	0.4154	1050	42	23	0.1	43	0.2	1
8	-2.2283	0.3935	1050	40	47	0.2	90	0.4	1
9	-2.0807	0.3756	1050	38	75	0.3	165	0.7	1
10	-1.9452	0.3607	1050	37	112	0.5	277	1.1	1
11	-1.8197	0.3482	1050	35	199	0.8	476	1.9	2
12	-1.7023	0.3374	1050	34	258	1.0	734	3.0	2
13	-1.5917	0.3280	1050	33	373	1.5	1107	4.5	4
14	-1.4868	0.3198	1050	33	397	1.6	1504	6.1	5
15	-1.3869	0.3126	1053	32	477	1.9	1981	8.0	7
16	-1.2911	0.3063	1063	31	504	2.0	2485	10.1	9
17	-1.1991	0.3006	1073	31	520	2.1	3005	12.2	11
18	-1.1103	0.2955	1082	30	615	2.5	3620	14.7	13
19	-1.0243	0.2910	1090	30	616	2.5	4236	17.2	16
20	-0.9408	0.2869	1099	29	637	2.6	4873	19.7	18
21	-0.8595	0.2833	1107	29	615	2.5	5488	22.2	21
22	-0.7802	0.2800	1115	29	619	2.5	6107	24.7	23
23	-0.7026	0.2771	1123	28	574	2.3	6681	27.1	26
24	-0.6266	0.2744	1131	28	646	2.6	7327	29.7	28
25	-0.5520	0.2720	1138	28	674	2.7	8001	32.4	31
26	-0.4786	0.2699	1146	27	672	2.7	8673	35.1	34
27	-0.4062	0.2680	1153	27	683	2.8	9356	37.9	37
28	-0.3348	0.2664	1161	27	728	3.0	10084	40.9	39
29	-0.2643	0.2649	1168	27	678	2.7	10762	43.6	42
30	-0.1944	0.2636	1175	27	707	2.9	11469	46.5	45
31	-0.1253	0.2626	1182	27	752	3.0	12221	49.5	48
32	-0.0566	0.2617	1189	27	766	3.1	12987	52.6	51
33	0.0117	0.2609	1196	27	778	3.2	13765	55.8	54
34	0.0797	0.2604	1203	27	768	3.1	14533	58.9	57
35	0.1473	0.2600	1210	26	820	3.3	15353	62.2	61
36	0.2149	0.2597	1217	26	838	3.4	16191	65.6	64
37	0.2823	0.2596	1223	26	831	3.4	17022	69.0	67
38	0.3497	0.2597	1230	26	740	3.0	17762	72.0	70
39	0.4172	0.2600	1237	26	780	3.2	18542	75.1	74
40	0.4849	0.2604	1244	27	825	3.3	19367	78.5	77
41	0.5528	0.2609	1251	27	734	3.0	20101	81.5	80
42	0.6211	0.2616	1258	27	693	2.8	20794	84.3	83
43	0.6897	0.2625	1265	27	617	2.5	21411	86.8	86
44	0.7590	0.2636	1272	27	561	2.3	21972	89.0	88
45	0.8288	0.2649	1279	27	505	2.0	22477	91.1	90
46	0.8993	0.2663	1286	27	476	1.9	22953	93.0	92
47	0.9707	0.2680	1294	27	391	1.6	23344	94.6	94
48	1.0430	0.2699	1301	27	356	1.4	23700	96.0	95
49	1.1164	0.2720	1308	28	258	1.0	23958	97.1	97
50	1.1910	0.2744	1316	28	195	0.8	24153	97.9	97

2010 PSSA Grade 12 Retest Technical Report for Mathematics, Reading, Science, and Writing

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
51	1.2670	0.2771	1324	28	141	0.6	24294	98.5	98
52	1.3447	0.2800	1332	29	133	0.5	24427	99.0	99
53	1.4240	0.2833	1340	29	73	0.3	24500	99.3	99
54	1.5053	0.2870	1348	29	57	0.2	24557	99.5	99
55	1.5888	0.2911	1356	30	48	0.2	24605	99.7	99
56	1.6749	0.2957	1365	30	27	0.1	24632	99.8	99
57	1.7638	0.3008	1374	31	9	0.0	24641	99.9	99
58	1.8561	0.3066	1384	31	6	0.0	24647	99.9	99
59	1.9520	0.3131	1393	32	11	0.0	24658	99.9	99
60	2.0524	0.3205	1404	33	10	0.0	24668	100.0	99
61	2.1577	0.3290	1414	33	3	0.0	24671	100.0	99
62	2.2692	0.3388	1426	34	3	0.0	24674	100.0	99
63	2.3877	0.3502	1438	36	1	0.0	24675	100.0	99
64	2.5150	0.3637	1451	37	0	0.0	24675	100.0	99
65	2.6530	0.3798	1465	39	0	0.0	24675	100.0	99
66	2.8046	0.3994	1480	41	1	0.0	24676	100.0	99
67	2.9736	0.4236	1497	43	0	0.0	24676	100.0	100
68	3.1658	0.4544	1517	46	0	0.0	24676	100.0	100
69	3.3902	0.4946	1540	50	0	0.0	24676	100.0	100
70	3.6614	0.5496	1567	56	0	0.0	24676	100.0	100
71	4.0064	0.6298	1603	64	0	0.0	24676	100.0	100
72	4.4833	0.7620	1651	78	0	0.0	24676	100.0	100
73	5.2690	1.0516	1731	107	0	0.0	24676	100.0	100
74	6.5547	1.8605	1862	189	0	0.0	24676	100.0	100

Appendix Q: 2010 Grade 12 Fall Writing Retest Raw-to-Scaled Score Conversion Table

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
22	-7.2177	1.8370	700	184	13	0.1	13	0.1	1
23	-5.9847	1.0207	700	102	57	0.6	70	0.8	1
24	-5.2490	0.7371	719	74	113	1.3	183	2.1	1
25	-4.7995	0.6149	764	61	156	1.8	339	3.8	3
26	-4.4664	0.5443	798	54	212	2.4	551	6.3	5
27	-4.1961	0.4979	825	50	201	2.3	752	8.5	7
28	-3.9650	0.4651	848	47	168	1.9	920	10.4	9
29	-3.7604	0.4408	868	44	162	1.8	1082	12.3	11
30	-3.5744	0.4223	887	42	125	1.4	1207	13.7	13
31	-3.4023	0.4080	904	41	113	1.3	1320	15.0	14
32	-3.2405	0.3968	920	40	81	0.9	1401	15.9	15
33	-3.0866	0.3880	936	39	63	0.7	1464	16.6	16
34	-2.9388	0.3812	950	38	49	0.6	1513	17.2	17
35	-2.7956	0.3758	965	38	72	0.8	1585	18.0	18
36	-2.6560	0.3717	979	37	95	1.1	1680	19.1	19
37	-2.5191	0.3686	992	37	127	1.4	1807	20.5	20
38	-2.3842	0.3662	1006	37	181	2.1	1988	22.5	22
39	-2.2507	0.3644	1019	36	189	2.1	2177	24.7	24
40	-2.1184	0.3631	1032	36	213	2.4	2390	27.1	26
41	-1.9870	0.3621	1046	36	206	2.3	2596	29.4	28
42	-1.8561	0.3612	1059	36	227	2.6	2823	32.0	31
43	-1.7260	0.3603	1072	36	170	1.9	2993	33.9	33
44	-1.5966	0.3593	1085	36	102	1.2	3095	35.1	35
45	-1.4679	0.3582	1098	36	90	1.0	3185	36.1	36
46	-1.3400	0.3568	1110	36	84	1.0	3269	37.1	37
47	-1.2133	0.3552	1123	36	119	1.3	3388	38.4	38
48	-1.0877	0.3535	1136	35	159	1.8	3547	40.2	39
49	-0.9635	0.3515	1148	35	237	2.7	3784	42.9	42
50	-0.8406	0.3495	1160	35	272	3.1	4056	46.0	44
51	-0.7192	0.3474	1172	35	401	4.5	4457	50.6	48
52	-0.5992	0.3454	1184	35	438	5.0	4895	55.5	53
53	-0.4805	0.3436	1196	34	442	5.0	5337	60.5	58
54	-0.3631	0.3419	1208	34	393	4.5	5730	65.0	63
55	-0.2466	0.3406	1220	34	312	3.5	6042	68.5	67
56	-0.1310	0.3396	1231	34	164	1.9	6206	70.4	69
57	-0.0159	0.3391	1243	34	76	0.9	6282	71.3	71
58	0.0990	0.3390	1254	34	40	0.5	6322	71.7	71
59	0.2141	0.3395	1266	34	37	0.4	6359	72.1	72
60	0.3297	0.3405	1277	34	78	0.9	6437	73.0	73
61	0.4462	0.3423	1289	34	106	1.2	6543	74.2	74
62	0.5642	0.3448	1301	34	137	1.6	6680	75.8	75
63	0.6842	0.3482	1313	35	218	2.5	6898	78.2	77
64	0.8068	0.3525	1325	35	215	2.4	7113	80.7	79
65	0.9330	0.3579	1338	36	231	2.6	7344	83.3	82
66	1.0634	0.3646	1351	36	174	2.0	7518	85.3	84
67	1.1993	0.3728	1364	37	118	1.3	7636	86.6	86
68	1.3420	0.3828	1379	38	24	0.3	7660	86.9	87
69	1.4931	0.3950	1394	40	16	0.2	7676	87.1	87
70	1.6549	0.4099	1410	41	13	0.1	7689	87.2	87
71	1.8302	0.4282	1427	43	31	0.4	7720	87.6	87

2010 PSSA Grade 12 Retest Technical Report for Mathematics, Reading, Science, and Writing

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
72	2.0232	0.4511	1447	45	45	0.5	7765	88.1	88
73	2.2394	0.4798	1468	48	89	1.0	7854	89.1	89
74	2.4870	0.5166	1493	52	158	1.8	8012	90.9	90
75	2.7779	0.5638	1522	56	150	1.7	8162	92.6	92
76	3.1291	0.6225	1557	62	198	2.2	8360	94.8	94
77	3.5571	0.6840	1600	68	215	2.4	8575	97.3	96
78	4.0522	0.7141	1650	71	110	1.2	8685	98.5	98
79	4.5474	0.6842	1699	68	18	0.2	8703	98.7	99
80	4.9767	0.6247	1742	62	1	0.0	8704	98.7	99
81	5.3322	0.5692	1778	57	1	0.0	8705	98.7	99
82	5.6311	0.5259	1807	53	3	0.0	8708	98.8	99
83	5.8901	0.4935	1833	49	3	0.0	8711	98.8	99
84	6.1214	0.4696	1856	47	5	0.1	8716	98.9	99
85	6.3334	0.4521	1878	45	11	0.1	8727	99.0	99
86	6.5319	0.4397	1897	44	12	0.1	8739	99.1	99
87	6.7214	0.4315	1916	43	18	0.2	8757	99.3	99
88	6.9054	0.4270	1935	43	12	0.1	8769	99.5	99
89	7.0870	0.4259	1953	43	19	0.2	8788	99.7	99
90	7.2691	0.4281	1971	43	5	0.1	8793	99.7	99
91	7.4546	0.4339	1990	43	0	0.0	8793	99.7	99
92	7.6468	0.4437	2009	44	0	0.0	8793	99.7	99
93	7.8498	0.4584	2029	46	0	0.0	8793	99.7	99
94	8.0692	0.4795	2051	48	0	0.0	8793	99.7	99
95	8.3129	0.5095	2076	51	0	0.0	8793	99.7	99
96	8.5940	0.5535	2104	55	1	0.0	8794	99.8	99
97	8.9367	0.6220	2138	62	2	0.0	8796	99.8	99
98	9.3943	0.7423	2184	74	4	0.0	8800	99.8	99
99	10.1373	1.0240	2258	102	10	0.1	8810	99.9	99
100	11.3748	1.8386	2382	184	6	0.1	8816	100.0	99