

# **TECHNICAL REPORT**



**for the  
Pennsylvania  
System of School Assessment**

**2011 Grade 12 Fall Retest  
Mathematics, Reading, Science, and Writing**

**Provided by  
Data Recognition Corporation**

**April 2012**

# *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: 2011 Grade 12 Fall Mathematics Retest Multiple-Choice Item Statistics	11
Appendix B: 2011 Grade 12 Fall Mathematics Retest Multiple-Choice Rasch Item Statistics	13
Appendix C: 2011 Grade 12 Fall Mathematics Retest Open-Ended Item Statistics	15
Appendix D: 2011 Grade 12 Fall Reading Retest Multiple-Choice Item Statistics	16
Appendix E: 2011 Grade 12 Fall Reading Retest Multiple-Choice Rasch Item Statistics	18
Appendix F: 2011 Grade 12 Fall Reading Retest Open-Ended Item Statistics	19
Appendix G: 2011 Grade 12 Fall Science Retest Multiple-Choice Item Statistics	20
Appendix H: 2011 Grade 12 Fall Science Retest Multiple-Choice Rasch Item Statistics	22
Appendix I: 2011 Grade 12 Fall Science Retest Open-Ended Item Statistics	23
Appendix J: 2011 Grade 12 Fall Writing Retest Multiple-Choice Item Statistics	24
Appendix K: 2011 Grade 12 Fall Writing Retest Multiple-Choice Rasch Item Statistics	25
Appendix L: 2011 Grade 12 Fall Writing Retest Prompt Statistics	26
Appendix M: 2011 Grade 12 Fall Writing Retest Percentage Agreement	27
Appendix N: 2011 Grade 12 Fall Mathematics Retest Raw-to-Scaled Score Conversion Table	28
Appendix O: 2011 Grade 12 Fall Reading Retest Raw-to-Scaled Score Conversion Table	30
Appendix P: 2011 Grade 12 Fall Science Retest Raw-to-Scaled Score Conversion Table	32
Appendix Q: 2011 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**

<b>Term</b>	<b>Common Definition</b>
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.)

<b>Term</b>	<b>Common Definition</b>
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 2011 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 2011 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 2011 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 2011 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 “snap shot” 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 2011 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**

<b>Content</b>	<b>Transformation</b>
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 2011 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
<b>N</b>	127797	25694	127997	19491	125307	20893	128775	6646
<b>Mean</b>	1379.0	1142.9	1381.6	1121.6	1244.8	1181.3	1536.8	1188.2
<b>St. Dev.</b>	264.0	154.6	273.4	208.7	96.5	74.2	298.6	245.8
<b>Min</b>	700	700	700	700	1050	1050	700	700
<b>Max</b>	2425	2157	2511	2511	1696	1519	2364	2364
<b>Reli.</b>	0.94	0.88	0.90	0.85	0.92	0.87	0.82	0.85

Figures 1–4 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 2011 PSSA Technical Report.

**Figure 1: Mathematics Operational and Retest Scaled Score Frequency Distributions**

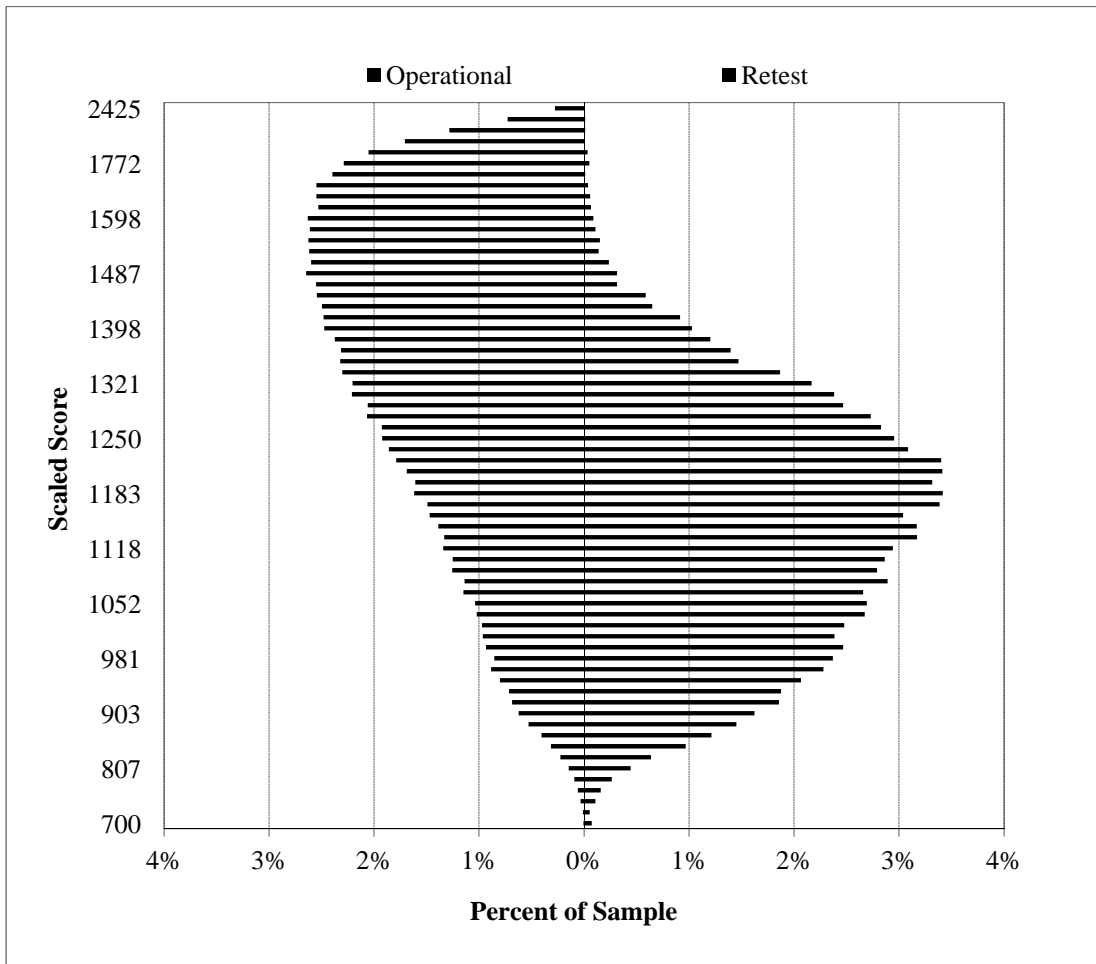


Figure 2: Reading Operational and Retest Scaled Score Frequency Distribution

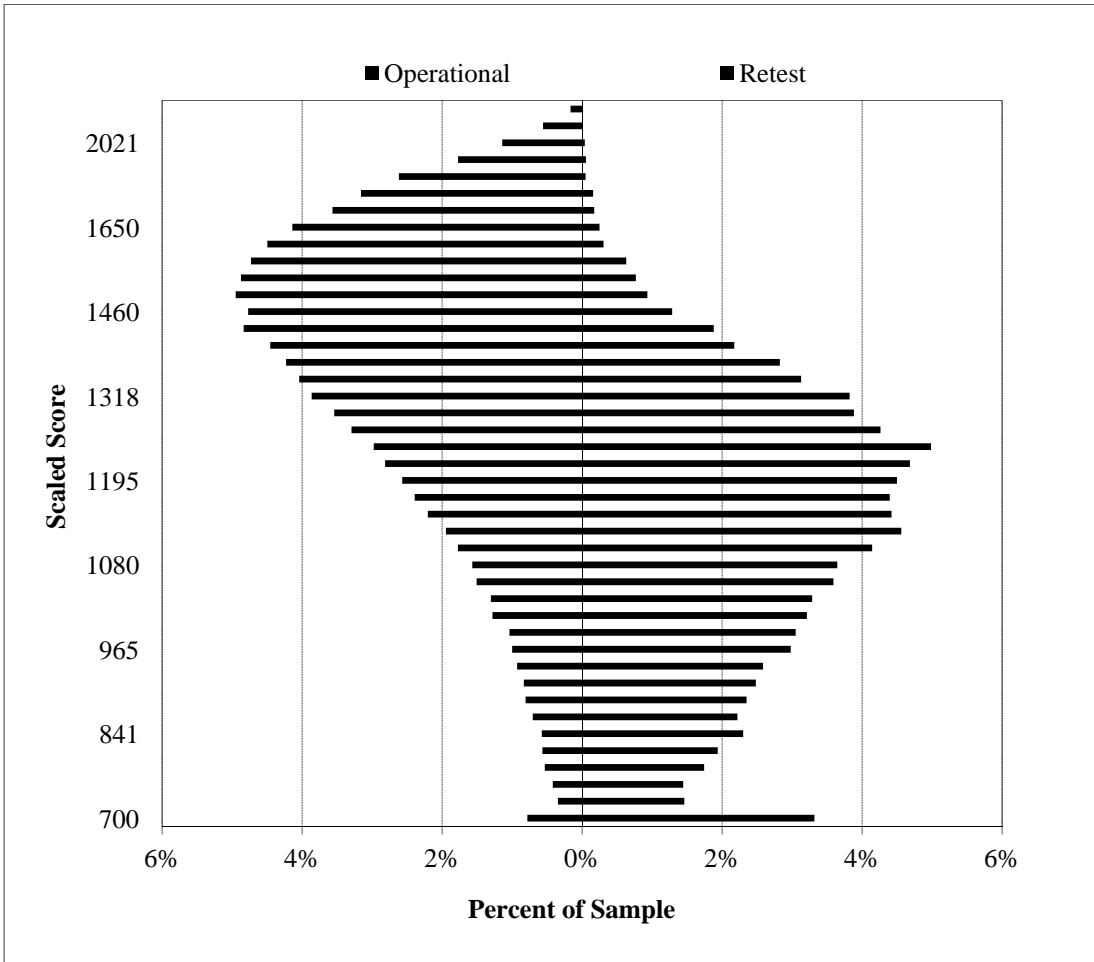


Figure 3: Science Operational and Retest Scaled Score Frequency Distribution

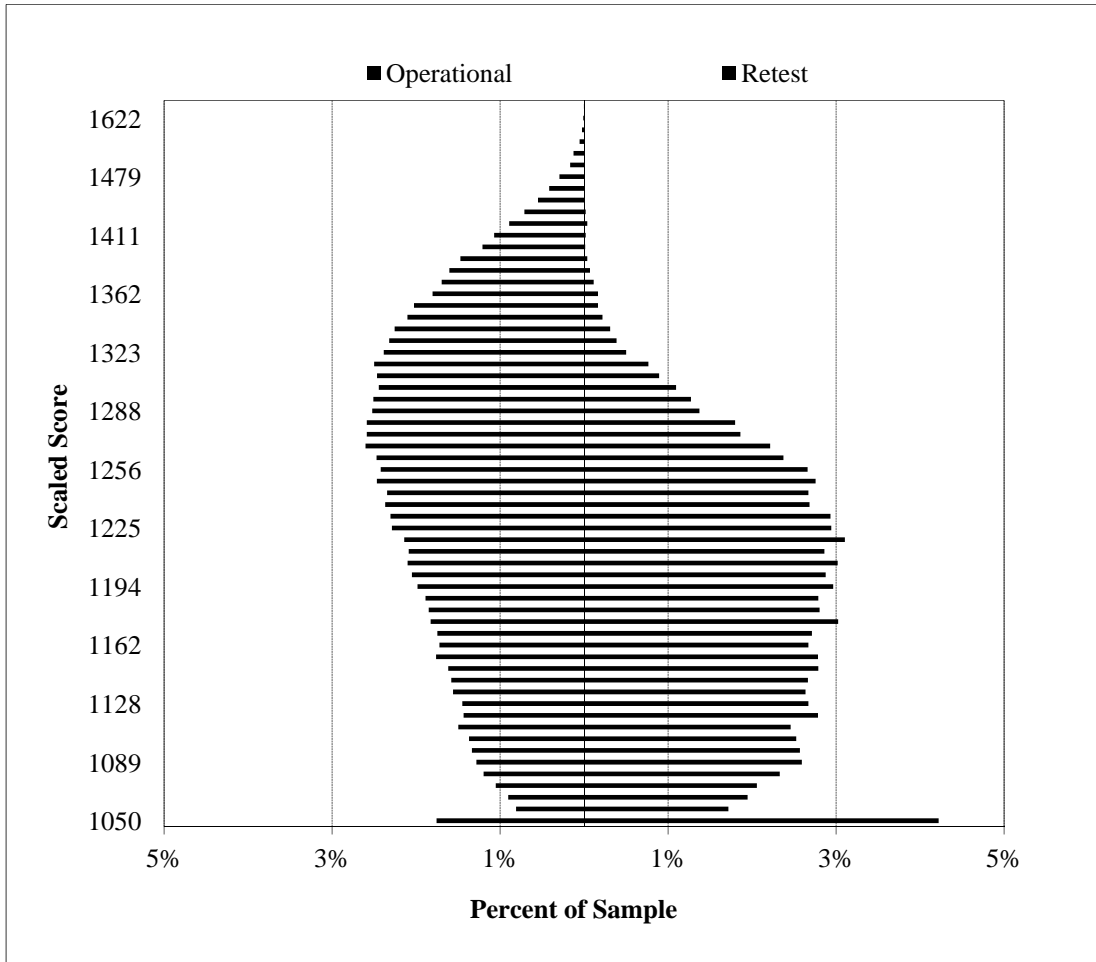
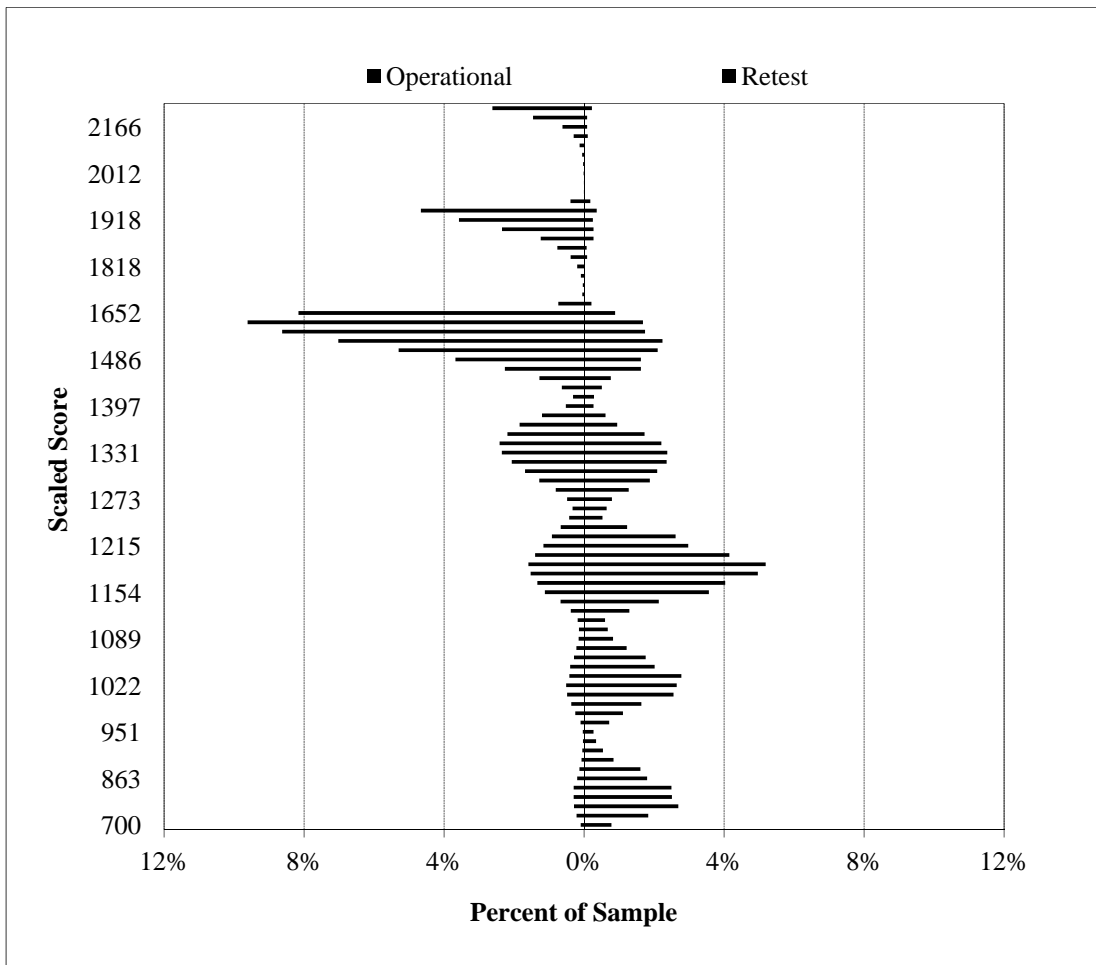




Figure 4: Writing Operational and Retest Scaled Score Frequency Distribution



### ***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 15.3, 26.7, 11.2, and 34.8, respectively. (Numbers may be slightly different than tabled numbers due to rounding.) Approximately 65%–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**

<b>Mathematics</b>			
<b>Performance Level</b>	<b>Scaled Score</b>	<b>Frequency</b>	<b>Percent</b>
<b>Advanced</b>	1509 and up	201	0.8
<b>Proficient</b>	1304-1508	3735	14.5
<b>Basic</b>	1167-1303	7966	31.0
<b>Below Basic</b>	1166 and below	13792	53.7

<b>Reading</b>			
<b>Performance Level</b>	<b>Scaled Score</b>	<b>Frequency</b>	<b>Percent</b>
<b>Advanced</b>	1492 and up	663	3.4
<b>Proficient</b>	1257-1491	4536	23.3
<b>Basic</b>	1112-1256	5370	27.6
<b>Below Basic</b>	1111 and below	8922	45.8

<b>Science</b>			
<b>Performance Level</b>	<b>Scaled Score</b>	<b>Frequency</b>	<b>Percent</b>
<b>Advanced</b>	1347 and up	140	0.7
<b>Proficient</b>	1275-1346	2191	10.5
<b>Basic</b>	1150-1274	11045	52.9
<b>Below Basic</b>	1149 and below	7517	36.0

<b>Writing</b>			
<b>Performance Level</b>	<b>Scaled Score</b>	<b>Frequency</b>	<b>Percent</b>
<b>Advanced</b>	1806 and up	136	2.0
<b>Proficient</b>	1236-1805	2174	32.7
<b>Basic</b>	952-1235	3290	49.5
<b>Below Basic</b>	951 and below	1046	15.7

*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, 62.2% of the students remained at the same performance level in mathematics, while 23.3% transitioned to a higher level and 14.5% regressed to a lower level. For reading, 50.6% of the students stayed at the same level, 35.5% improved and 13.9% regressed. For science, 69.4% of the students stayed at the same level, 16.0% improved and 14.6% regressed. For writing, 55.6% of the students remained at the same level, 31.1% improved and 13.3% regressed.

## Appendix A: 2011 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	B	25694	0.414	0.082	0.414	0.105	0.399	0.001	0.000	0.103	-0.182	0.103	-0.167	0.105
2	C	25694	0.725	0.035	0.055	0.725	0.183	0.001	0.000	0.262	-0.114	-0.146	0.262	-0.159
3	B	25694	0.686	0.090	0.686	0.157	0.065	0.001	0.000	0.260	-0.089	0.260	-0.158	-0.147
4	B	25694	0.306	0.134	0.306	0.394	0.165	0.002	0.000	0.147	-0.082	0.147	-0.024	-0.073
5	C	25694	0.628	0.050	0.176	0.628	0.145	0.001	0.000	0.262	-0.061	-0.210	0.262	-0.091
6	C	25694	0.382	0.174	0.288	0.382	0.149	0.005	0.000	0.177	-0.091	-0.032	0.177	-0.102
7	A	25694	0.643	0.643	0.166	0.102	0.084	0.004	0.000	0.181	0.181	-0.064	-0.107	-0.108
8	C	25694	0.600	0.233	0.104	0.600	0.059	0.003	0.000	0.411	-0.265	-0.174	0.411	-0.147
9	C	25694	0.653	0.165	0.105	0.653	0.074	0.002	0.000	0.229	-0.154	-0.161	0.229	-0.003
10	C	25694	0.648	0.053	0.132	0.648	0.164	0.002	0.000	0.292	-0.129	-0.161	0.292	-0.147
11	D	25694	0.552	0.149	0.171	0.124	0.552	0.003	0.000	0.391	-0.196	-0.187	-0.160	0.391
12	C	25694	0.794	0.060	0.081	0.794	0.063	0.001	0.000	0.282	-0.229	-0.111	0.282	-0.113
13	B	25694	0.703	0.088	0.703	0.103	0.104	0.001	0.000	0.360	-0.243	0.360	-0.149	-0.158
14	C	25694	0.495	0.153	0.220	0.495	0.123	0.008	0.000	0.261	-0.091	-0.111	0.261	-0.152
15	D	25694	0.443	0.096	0.343	0.116	0.443	0.002	0.000	0.302	-0.213	-0.036	-0.211	0.302
16	B	25694	0.531	0.162	0.531	0.171	0.132	0.004	0.001	0.331	-0.050	0.331	-0.246	-0.154
17	B	25694	0.437	0.268	0.437	0.115	0.177	0.002	0.001	0.357	-0.090	0.357	-0.128	-0.247
18	D	25694	0.555	0.085	0.198	0.158	0.555	0.004	0.001	0.371	-0.183	-0.173	-0.169	0.371
19	A	25694	0.623	0.623	0.070	0.078	0.226	0.003	0.001	0.165	0.165	-0.215	-0.219	0.086
20	B	25694	0.579	0.111	0.579	0.169	0.137	0.003	0.000	0.331	-0.054	0.331	-0.194	-0.207
21	D	25694	0.385	0.135	0.176	0.298	0.385	0.005	0.000	0.279	-0.114	-0.126	-0.100	0.279
22	C	25694	0.459	0.199	0.151	0.459	0.187	0.003	0.001	0.362	-0.067	-0.159	0.362	-0.240
23	B	25694	0.747	0.075	0.747	0.079	0.094	0.004	0.000	0.400	-0.181	0.400	-0.185	-0.253
26	B	25694	0.641	0.183	0.641	0.144	0.029	0.003	0.000	0.313	-0.126	0.313	-0.249	-0.078
27	A	25694	0.591	0.591	0.090	0.272	0.043	0.002	0.000	0.344	0.344	-0.218	-0.175	-0.136
28	C	25694	0.528	0.156	0.261	0.528	0.051	0.003	0.000	0.314	-0.151	-0.146	0.314	-0.161
29	D	25694	0.691	0.087	0.070	0.149	0.691	0.002	0.001	0.406	-0.190	-0.232	-0.204	0.406
30	A	25694	0.539	0.539	0.207	0.122	0.129	0.003	0.001	0.329	0.329	-0.261	-0.158	-0.011
31	C	25694	0.496	0.107	0.217	0.496	0.175	0.004	0.000	0.232	-0.164	-0.072	0.232	-0.088
32	A	25694	0.480	0.480	0.198	0.168	0.150	0.003	0.001	0.368	0.368	-0.147	-0.201	-0.130
33	B	25694	0.428	0.335	0.428	0.158	0.073	0.006	0.000	0.179	-0.030	0.179	-0.115	-0.116
34	A	25694	0.530	0.530	0.136	0.190	0.140	0.003	0.001	0.374	0.374	-0.223	-0.161	-0.126
35	B	25694	0.536	0.188	0.536	0.182	0.089	0.004	0.001	0.304	-0.156	0.304	-0.125	-0.137
36	A	25694	0.581	0.581	0.203	0.116	0.097	0.002	0.001	0.372	0.372	-0.136	-0.215	-0.189

**2011 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	25694	0.537	0.059	0.345	0.537	0.056	0.003	0.000	0.323	-0.198	-0.167	0.323	-0.136
38	B	25694	0.409	0.286	0.409	0.192	0.105	0.007	0.001	0.183	-0.014	0.183	-0.117	-0.110
39	B	25694	0.455	0.306	0.455	0.130	0.104	0.004	0.000	0.225	-0.036	0.225	-0.152	-0.129
40	B	25694	0.427	0.204	0.427	0.136	0.227	0.006	0.001	0.133	0.021	0.133	-0.169	-0.024
41	C	25694	0.533	0.135	0.153	0.533	0.172	0.006	0.001	0.337	-0.153	-0.223	0.337	-0.082
42	B	25694	0.669	0.157	0.669	0.100	0.067	0.005	0.001	0.407	-0.180	0.407	-0.237	-0.194
43	C	25694	0.443	0.168	0.280	0.443	0.105	0.005	0.001	0.159	-0.073	-0.064	0.159	-0.055
44	D	25694	0.660	0.089	0.120	0.126	0.660	0.004	0.001	0.441	-0.212	-0.206	-0.226	0.441
45	B	25694	0.451	0.137	0.451	0.335	0.074	0.004	0.000	0.152	-0.204	0.152	0.028	-0.050
46	A	25694	0.442	0.442	0.214	0.240	0.096	0.006	0.001	0.274	0.274	-0.074	-0.175	-0.082
47	D	25694	0.396	0.123	0.254	0.217	0.396	0.009	0.001	0.273	-0.105	-0.119	-0.097	0.273
48	A	25694	0.520	0.520	0.178	0.131	0.163	0.007	0.001	0.360	0.360	-0.107	-0.182	-0.188
49	C	25694	0.528	0.140	0.266	0.528	0.060	0.006	0.001	0.282	-0.125	-0.104	0.282	-0.184
50	A	25694	0.438	0.438	0.189	0.214	0.150	0.009	0.001	0.340	0.340	-0.140	-0.146	-0.128
51	A	25694	0.462	0.462	0.141	0.102	0.288	0.006	0.001	0.306	0.306	-0.156	-0.235	-0.042
52	C	25694	0.533	0.089	0.275	0.533	0.096	0.007	0.001	0.288	-0.180	-0.101	0.288	-0.135
53	C	25694	0.438	0.177	0.235	0.438	0.140	0.009	0.001	0.173	-0.099	-0.054	0.173	-0.048
54	B	25694	0.441	0.123	0.441	0.263	0.164	0.008	0.001	0.079	-0.167	0.079	0.070	-0.017
55	B	25694	0.469	0.177	0.469	0.199	0.146	0.007	0.001	0.311	-0.078	0.311	-0.164	-0.144
56	C	25694	0.387	0.376	0.117	0.387	0.110	0.008	0.001	0.254	0.015	-0.260	0.254	-0.122
57	B	25694	0.486	0.082	0.486	0.192	0.232	0.007	0.001	0.299	-0.180	0.299	-0.100	-0.123
58	D	25694	0.436	0.176	0.112	0.267	0.436	0.008	0.001	0.331	-0.105	-0.244	-0.085	0.331
59	D	25694	0.309	0.141	0.348	0.193	0.309	0.009	0.001	0.200	-0.138	0.012	-0.102	0.200
60	A	25694	0.498	0.498	0.120	0.183	0.189	0.008	0.001	0.306	0.306	-0.192	-0.182	-0.027
61	D	25694	0.471	0.140	0.255	0.125	0.471	0.008	0.001	0.300	-0.155	-0.094	-0.139	0.300
62	A	25694	0.448	0.448	0.182	0.263	0.097	0.009	0.001	0.295	0.295	-0.117	-0.087	-0.182

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

## Appendix B: 2011 Grade 12 Fall Mathematics Retest Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit		OutFit	
			MS	ZSTD	MS	ZSTD
1	0.5376	0.0141	1.28	9.9	1.41	9.9
2	-1.0568	0.0139	0.92	-9.9	0.91	-9.9
3	-1.0192	0.0139	0.97	-5.1	0.96	-5.1
4	0.8174	0.0148	1.16	9.9	1.27	9.9
5	-0.6290	0.0134	0.99	-2.5	0.98	-2.8
6	0.3125	0.0137	1.09	9.9	1.13	9.9
7	-0.4286	0.0133	1.06	9.9	1.07	9.9
8	-0.7163	0.0135	0.91	-9.9	0.88	-9.9
9	-0.7412	0.0135	1.00	-0.8	1.00	0.4
10	-0.6938	0.0134	0.95	-9.9	0.96	-7.3
11	-0.5096	0.0133	0.93	-9.9	0.91	-9.9
12	-1.4832	0.0150	0.85	-9.9	0.84	-9.9
13	-1.0047	0.0138	0.87	-9.9	0.86	-9.9
14	-0.2823	0.0132	1.02	5.7	1.03	5.3
15	0.1707	0.0135	1.03	6.0	1.04	5.9
16	-0.3215	0.0132	0.97	-7.2	0.97	-6.3
17	-0.0600	0.0133	0.95	-9.9	0.94	-9.9
18	-0.5146	0.0133	0.94	-9.9	0.93	-9.9
19	-0.7486	0.0135	1.07	9.9	1.08	9.9
20	-0.4629	0.0133	0.96	-9.8	0.95	-9.4
21	-0.0209	0.0133	0.98	-4.2	0.98	-4.0
22	-0.2506	0.0132	0.95	-9.9	0.94	-9.9
23	-1.4873	0.0150	0.89	-9.9	0.84	-9.9
26	-0.9606	0.0138	0.98	-4.0	0.98	-2.5
27	-1.0772	0.0140	1.07	9.9	1.07	9.2
28	-0.2737	0.0132	0.99	-3.6	0.98	-3.0
29	-1.2931	0.0145	0.94	-9.9	0.90	-9.9
30	-0.4532	0.0133	0.97	-6.9	0.97	-6.3
31	-0.1848	0.0133	1.05	9.9	1.06	9.9
32	-0.5348	0.0133	0.98	-4.7	0.97	-5.4
33	0.2111	0.0135	1.11	9.9	1.16	9.9
34	-0.6861	0.0134	0.98	-4.4	0.97	-5.2
35	-0.4082	0.0133	0.99	-2.8	0.99	-2.0
36	-0.6624	0.0134	0.94	-9.9	0.93	-9.9
37	-0.4484	0.0133	0.98	-5.8	0.97	-5.2
38	0.1237	0.0134	1.08	9.9	1.10	9.9
39	0.1710	0.0135	1.10	9.9	1.13	9.9
40	-0.1055	0.0133	1.11	9.9	1.12	9.9
41	-0.7492	0.0135	1.02	5.1	1.02	3.1
42	-1.3961	0.0147	1.04	5.1	1.00	-0.5
43	0.5713	0.0142	1.31	9.9	1.42	9.9
44	-1.4293	0.0148	1.05	7.0	1.01	1.2
45	0.0773	0.0134	1.12	9.9	1.15	9.9
46	-0.2861	0.0132	1.01	3.7	1.01	2.7
47	-0.0270	0.0133	0.99	-2.1	0.99	-1.5
48	-0.4128	0.0133	0.95	-9.9	0.95	-9.9
49	-0.4220	0.0133	1.01	1.8	1.01	1.6
50	-0.4071	0.0133	0.99	-3.0	0.98	-3.4
51	-0.3100	0.0132	0.99	-2.0	0.99	-1.7
52	-0.3761	0.0132	1.00	0.2	1.00	0.1
53	0.0874	0.0134	1.10	9.9	1.13	9.9

Seq.	Anchored Measure	Measure SE	InFit		OutFit	
			MS	ZSTD	MS	ZSTD
54	0.5351	0.0141	1.35	9.9	1.47	9.9
55	-0.4611	0.0133	1.01	2.4	1.01	1.7
56	-0.1397	0.0133	1.01	2.0	1.01	2.8
57	-0.5366	0.0133	1.03	6.6	1.02	4.5
58	-0.2687	0.0132	0.97	-7.1	0.96	-7.0
59	0.7905	0.0147	1.12	9.9	1.18	9.9
60	-0.3775	0.0132	0.99	-1.7	0.99	-1.2
61	-0.0286	0.0133	1.01	1.9	1.01	2.4
62	-0.3476	0.0132	1.01	1.6	1.01	1.7

## Appendix C: 2011 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
24	4	25694	0.534	0.629	0.258	0.073	0.030	0.010	0.042	0.000	0.000	0.408	-0.372	0.183	0.209	0.179	0.149
25	4	25694	1.544	0.210	0.201	0.454	0.107	0.028	0.043	0.001	0.000	0.517	-0.441	-0.049	0.165	0.268	0.207
63	4	25694	0.544	0.543	0.380	0.067	0.009	0.001	0.068	0.000	0.000	0.505	-0.488	0.341	0.256	0.132	0.055

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



## Appendix D: 2011 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	19491	0.259	0.259	0.663	0.043	0.034	0.001	0.000	0.139	0.139	0.008	-0.195	-0.131
2	B	19491	0.899	0.030	0.899	0.057	0.014	0.001	0.000	0.357	-0.181	0.357	-0.243	-0.169
3	C	19491	0.496	0.113	0.225	0.496	0.163	0.003	0.000	0.172	-0.082	0.035	0.172	-0.198
4	A	19491	0.448	0.448	0.391	0.089	0.070	0.001	0.001	0.143	0.143	0.032	-0.155	-0.158
5	A	19491	0.556	0.556	0.124	0.175	0.142	0.002	0.000	0.328	0.328	-0.138	-0.179	-0.134
6	D	19491	0.690	0.066	0.151	0.090	0.690	0.002	0.001	0.380	-0.215	-0.134	-0.252	0.380
7	C	19491	0.713	0.174	0.068	0.713	0.042	0.003	0.000	0.239	-0.037	-0.233	0.239	-0.168
9	B	19491	0.618	0.313	0.618	0.026	0.043	0.001	0.000	0.274	-0.151	0.274	-0.188	-0.159
10	C	19491	0.541	0.040	0.303	0.541	0.114	0.001	0.000	0.295	-0.224	-0.147	0.295	-0.108
11	C	19491	0.526	0.149	0.040	0.526	0.284	0.001	0.000	0.159	-0.174	-0.256	0.159	0.076
12	A	19491	0.736	0.736	0.075	0.107	0.080	0.001	0.001	0.387	0.387	-0.258	-0.200	-0.142
13	D	19491	0.413	0.391	0.101	0.093	0.413	0.002	0.000	0.257	0.039	-0.226	-0.259	0.257
14	A	19491	0.628	0.628	0.204	0.061	0.105	0.001	0.001	0.352	0.352	-0.147	-0.238	-0.169
15	B	19491	0.605	0.119	0.605	0.209	0.064	0.002	0.000	0.404	-0.200	0.404	-0.204	-0.194
16	C	19491	0.440	0.206	0.259	0.440	0.092	0.003	0.001	0.201	-0.070	-0.061	0.201	-0.147
17	C	19491	0.596	0.162	0.124	0.596	0.115	0.003	0.000	0.313	-0.190	-0.199	0.313	-0.048
18	C	19491	0.439	0.358	0.156	0.439	0.041	0.004	0.000	0.156	-0.041	-0.053	0.156	-0.179
19	A	19491	0.482	0.482	0.263	0.104	0.148	0.003	0.001	0.330	0.330	-0.172	-0.183	-0.083
20	A	19491	0.468	0.468	0.324	0.066	0.137	0.004	0.001	0.326	0.326	-0.071	-0.278	-0.162
21	D	19491	0.384	0.133	0.158	0.319	0.384	0.006	0.001	0.260	-0.170	-0.163	-0.011	0.260
22	A	19491	0.633	0.633	0.104	0.162	0.096	0.004	0.001	0.419	0.419	-0.257	-0.231	-0.116
23	B	19491	0.625	0.132	0.625	0.114	0.123	0.005	0.001	0.376	-0.140	0.376	-0.234	-0.170
24	B	19491	0.647	0.151	0.647	0.107	0.089	0.005	0.001	0.451	-0.182	0.451	-0.280	-0.209
26	C	19491	0.458	0.122	0.339	0.458	0.079	0.001	0.000	0.188	-0.219	0.039	0.188	-0.143
27	D	19491	0.545	0.073	0.252	0.128	0.545	0.002	0.000	0.394	-0.187	-0.142	-0.250	0.394
28	D	19491	0.633	0.115	0.106	0.143	0.633	0.002	0.001	0.428	-0.179	-0.229	-0.216	0.428
29	B	19491	0.330	0.047	0.330	0.363	0.257	0.002	0.001	0.179	-0.235	0.179	-0.106	0.045
30	D	19491	0.371	0.308	0.180	0.137	0.371	0.003	0.001	0.175	-0.018	-0.081	-0.122	0.175
31	A	19491	0.415	0.415	0.220	0.222	0.139	0.003	0.001	0.248	0.248	-0.134	-0.047	-0.125
32	C	19491	0.550	0.186	0.078	0.550	0.182	0.003	0.001	0.341	-0.159	-0.261	0.341	-0.090
33	B	19491	0.365	0.160	0.365	0.235	0.237	0.003	0.000	0.237	-0.128	0.237	-0.153	0.001
35	C	19491	0.733	0.063	0.134	0.733	0.063	0.006	0.000	0.379	-0.283	-0.121	0.379	-0.206

**2011 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	19491	0.387	0.156	0.125	0.325	0.387	0.007	0.000	0.319	-0.060	-0.245	-0.095	0.319
37	A	19491	0.608	0.608	0.095	0.234	0.053	0.008	0.001	0.284	0.284	-0.237	-0.030	-0.205
38	D	19491	0.676	0.094	0.105	0.117	0.676	0.007	0.001	0.434	-0.221	-0.216	-0.196	0.434
39	A	19491	0.354	0.354	0.344	0.246	0.047	0.008	0.001	0.241	0.241	0.018	-0.154	-0.227
40	D	19491	0.437	0.085	0.290	0.179	0.437	0.008	0.001	0.205	-0.202	-0.031	-0.058	0.205
41	B	19491	0.364	0.238	0.364	0.089	0.299	0.009	0.001	0.212	-0.071	0.212	-0.271	0.033
42	B	19491	0.550	0.164	0.550	0.173	0.104	0.009	0.001	0.324	-0.098	0.324	-0.208	-0.124
43	C	19491	0.405	0.291	0.180	0.405	0.114	0.009	0.001	0.206	-0.009	-0.076	0.206	-0.187

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

## Appendix E: 2011 Grade 12 Fall Reading Retest Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit		OutFit	
			MS	ZSTD	MS	ZSTD
1	1.3812	0.0179	1.16	9.9	1.46	9.9
2	-2.5113	0.0252	0.93	-4.1	0.75	-9.4
3	0.4338	0.0156	1.19	9.9	1.26	9.9
4	0.5191	0.0157	1.19	9.9	1.28	9.9
5	-0.0197	0.0155	0.99	-1.3	1.00	-0.7
6	-1.0022	0.0169	0.97	-3.6	0.96	-3.9
7	-0.5731	0.0159	0.97	-4.4	0.99	-1.6
9	-0.4120	0.0157	1.02	3.9	1.02	3.1
10	-0.0616	0.0155	1.02	3.8	1.03	3.8
11	0.2289	0.0155	1.16	9.9	1.22	9.9
12	-1.0664	0.0171	0.89	-9.9	0.87	-9.9
13	0.4257	0.0156	1.04	7.9	1.06	8.0
14	-0.5698	0.0159	0.97	-5.1	0.96	-4.7
15	-0.5685	0.0159	0.95	-7.6	0.93	-8.7
16	0.5036	0.0157	1.13	9.9	1.20	9.9
17	-0.2779	0.0156	0.99	-1.5	1.00	0.1
18	0.3908	0.0156	1.14	9.9	1.21	9.9
19	0.0694	0.0154	0.99	-2.0	1.00	-0.5
20	-0.0715	0.0155	1.01	1.4	1.00	0.7
21	0.4730	0.0157	1.02	4.1	1.04	5.5
22	-0.7510	0.0163	0.96	-6.6	0.92	-8.0
23	-0.6014	0.0160	0.96	-6.5	0.95	-6.3
24	-0.8286	0.0165	0.93	-9.4	0.90	-9.9
26	0.6013	0.0158	1.19	9.9	1.28	9.9
27	-0.2987	0.0156	0.95	-8.8	0.96	-5.2
28	-0.9536	0.0168	1.03	4.1	1.00	0.3
29	1.1644	0.0171	1.21	9.9	1.42	9.9
30	0.6919	0.0160	1.11	9.9	1.21	9.9
31	0.4983	0.0157	1.06	9.9	1.10	9.9
32	-0.4524	0.0158	1.03	5.4	1.02	2.8
33	0.4944	0.0157	1.02	4.3	1.07	8.0
35	-1.2005	0.0175	0.95	-5.9	0.94	-4.5
36	0.2012	0.0155	0.98	-4.2	0.97	-3.9
37	-0.2321	0.0155	1.01	1.9	1.02	3.2
38	-1.1510	0.0173	1.03	3.6	0.99	-0.6
39	0.4078	0.0156	1.01	2.0	1.04	5.2
40	0.8014	0.0162	1.23	9.9	1.33	9.9
41	0.6828	0.0160	1.07	9.9	1.15	9.9
42	-0.5041	0.0158	1.06	9.9	1.07	7.8
43	0.6928	0.0160	1.13	9.9	1.23	9.9

## Appendix F: 2011 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	19491	1.592	0.092	0.366	0.400	0.142	0.033	0.000	0.001	0.000	0.527	-0.350	-0.275	0.254	0.313
25	3	19491	1.099	0.211	0.508	0.253	0.029	0.067	0.000	0.010	0.000	0.548	-0.499	0.051	0.335	0.192
34	3	19491	1.367	0.158	0.381	0.398	0.064	0.061	0.000	0.002	0.000	0.580	-0.451	-0.178	0.387	0.251
44	3	19491	1.543	0.124	0.317	0.452	0.108	0.052	0.000	0.002	0.000	0.533	-0.408	-0.203	0.287	0.277

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

## Appendix G: 2011 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	20893	0.589	0.589	0.065	0.283	0.061	0.002	0.000	0.270	0.270	-0.102	-0.196	-0.073
2	A	20893	0.336	0.336	0.419	0.140	0.101	0.003	0.000	0.161	0.161	-0.045	-0.171	0.024
3	D	20893	0.726	0.077	0.093	0.085	0.726	0.018	0.000	0.362	-0.148	-0.237	-0.167	0.362
4	D	20893	0.773	0.040	0.128	0.059	0.773	0.001	0.000	0.373	-0.188	-0.196	-0.226	0.373
5	B	20893	0.578	0.131	0.578	0.192	0.097	0.002	0.000	0.332	-0.131	0.332	-0.156	-0.191
6	C	20893	0.415	0.180	0.251	0.415	0.150	0.002	0.000	0.154	-0.113	-0.073	0.154	0.002
7	B	20893	0.467	0.077	0.467	0.277	0.157	0.022	0.000	0.219	-0.188	0.219	-0.056	-0.077
8	A	20893	0.787	0.787	0.096	0.052	0.064	0.001	0.000	0.290	0.290	-0.176	-0.157	-0.130
9	B	20893	0.512	0.119	0.512	0.237	0.129	0.003	0.000	0.267	-0.131	0.267	-0.133	-0.095
10	C	20893	0.562	0.100	0.046	0.562	0.290	0.001	0.000	0.307	-0.212	-0.227	0.307	-0.089
11	C	20893	0.474	0.255	0.185	0.474	0.082	0.004	0.000	0.298	-0.081	-0.192	0.298	-0.134
12	A	20893	0.567	0.567	0.249	0.116	0.060	0.008	0.000	0.314	0.314	-0.099	-0.200	-0.197
13	C	20893	0.599	0.281	0.064	0.599	0.054	0.001	0.000	0.266	-0.058	-0.242	0.266	-0.195
14	D	20893	0.471	0.169	0.176	0.178	0.471	0.005	0.000	0.331	-0.118	-0.112	-0.201	0.331
16	D	20893	0.280	0.233	0.156	0.328	0.280	0.003	0.000	0.156	-0.011	-0.217	0.033	0.156
17	C	20893	0.457	0.200	0.120	0.457	0.218	0.004	0.000	0.326	-0.191	-0.226	0.326	-0.024
18	C	20893	0.635	0.123	0.085	0.635	0.141	0.015	0.000	0.384	-0.172	-0.244	0.384	-0.159
19	A	20893	0.592	0.592	0.085	0.065	0.254	0.004	0.001	0.247	0.247	-0.069	-0.165	-0.134
21	B	20893	0.466	0.133	0.466	0.360	0.038	0.002	0.000	0.195	-0.121	0.195	-0.056	-0.143
22	C	20893	0.632	0.138	0.082	0.632	0.144	0.003	0.000	0.433	-0.290	-0.242	0.433	-0.114
23	A	20893	0.380	0.380	0.199	0.321	0.098	0.002	0.000	0.158	0.158	-0.056	-0.035	-0.123
24	B	20893	0.412	0.154	0.412	0.222	0.208	0.004	0.000	0.218	-0.156	0.218	-0.082	-0.037
26	C	20893	0.591	0.192	0.145	0.591	0.064	0.007	0.000	0.295	-0.085	-0.210	0.295	-0.121
27	A	20893	0.578	0.578	0.129	0.237	0.055	0.002	0.000	0.280	0.280	-0.191	-0.086	-0.162
28	B	20893	0.467	0.250	0.467	0.154	0.117	0.012	0.000	0.343	-0.050	0.343	-0.218	-0.206
29	A	20893	0.617	0.617	0.239	0.080	0.062	0.001	0.000	0.281	0.281	-0.054	-0.227	-0.213
30	C	20893	0.466	0.234	0.206	0.466	0.091	0.002	0.000	0.365	-0.098	-0.174	0.365	-0.237
31	D	20893	0.537	0.153	0.120	0.188	0.537	0.002	0.000	0.360	-0.134	-0.241	-0.133	0.360
32	C	20893	0.417	0.159	0.084	0.417	0.338	0.002	0.000	0.264	-0.124	-0.222	0.264	-0.047
33	B	20893	0.507	0.170	0.507	0.238	0.082	0.003	0.000	0.344	-0.170	0.344	-0.108	-0.222
34	B	20893	0.527	0.102	0.527	0.298	0.072	0.001	0.000	0.288	-0.251	0.288	-0.061	-0.152
35	D	20893	0.416	0.217	0.220	0.139	0.416	0.008	0.000	0.337	-0.112	-0.095	-0.223	0.337
36	A	20893	0.519	0.519	0.248	0.180	0.052	0.002	0.000	0.304	0.304	-0.112	-0.158	-0.186

**2011 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	20893	0.284	0.284	0.265	0.293	0.156	0.002	0.000	0.124	0.124	-0.113	0.005	-0.019
41	A	20893	0.271	0.271	0.343	0.282	0.101	0.003	0.000	0.157	0.157	-0.176	0.041	-0.011
42	B	20893	0.316	0.345	0.316	0.238	0.098	0.003	0.000	0.261	-0.015	0.261	-0.144	-0.169
43	A	20893	0.469	0.469	0.156	0.216	0.145	0.012	0.000	0.273	0.273	-0.161	-0.059	-0.159
44	B	20893	0.546	0.156	0.546	0.158	0.130	0.010	0.001	0.318	-0.004	0.318	-0.143	-0.290
46	A	20893	0.255	0.255	0.227	0.225	0.290	0.003	0.000	0.209	0.209	-0.175	-0.230	0.178
47	D	20893	0.375	0.232	0.191	0.196	0.375	0.006	0.000	0.204	0.034	-0.109	-0.168	0.204
48	C	20893	0.474	0.157	0.194	0.474	0.162	0.013	0.000	0.342	-0.244	-0.100	0.342	-0.099
49	D	20893	0.428	0.179	0.248	0.144	0.428	0.001	0.000	0.179	-0.113	-0.059	-0.054	0.179
50	B	20893	0.362	0.263	0.362	0.194	0.177	0.004	0.000	0.187	0.053	0.187	-0.189	-0.097
51	C	20893	0.399	0.169	0.301	0.399	0.127	0.004	0.000	0.266	-0.211	0.013	0.266	-0.167
52	B	20893	0.431	0.149	0.431	0.341	0.075	0.004	0.000	0.183	-0.046	0.183	-0.083	-0.123
53	D	20893	0.370	0.183	0.292	0.151	0.370	0.004	0.000	0.306	-0.193	-0.058	-0.124	0.306
54	B	20893	0.383	0.146	0.383	0.328	0.134	0.008	0.000	0.270	-0.094	0.270	-0.052	-0.201
55	C	20893	0.478	0.246	0.154	0.478	0.119	0.003	0.000	0.262	-0.023	-0.130	0.262	-0.221
56	D	20893	0.525	0.230	0.109	0.132	0.525	0.004	0.000	0.401	-0.140	-0.246	-0.185	0.401
57	C	20893	0.469	0.124	0.160	0.469	0.244	0.004	0.000	0.282	-0.264	-0.189	0.282	0.043

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

## Appendix H: 2011 Grade 12 Fall Science Retest Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit		OutFit	
			MS	ZSTD	MS	ZSTD
1	-0.3778	0.0148	1.01	1.1	1.01	1.2
2	0.8870	0.0161	1.18	9.9	1.28	9.9
3	-1.2062	0.0163	0.91	-9.9	0.90	-9.5
4	-1.5699	0.0176	0.94	-6.7	0.88	-9.4
5	-0.4596	0.0149	0.97	-5.8	0.96	-6.1
6	0.5847	0.0154	1.19	9.9	1.27	9.9
7	0.2618	0.0149	1.10	9.9	1.13	9.9
8	-1.5296	0.0174	0.93	-8.2	0.93	-5.7
9	-0.1975	0.0147	1.03	5.5	1.03	5.2
10	-0.4262	0.0148	1.00	-1.0	0.99	-1.8
11	0.1459	0.0148	1.02	3.7	1.03	5.5
12	-0.4562	0.0149	0.99	-2.3	0.99	-2.1
13	-0.5595	0.0150	1.02	3.3	1.02	2.7
14	-0.0843	0.0147	0.98	-5.4	0.97	-5.3
16	1.1180	0.0168	1.17	9.9	1.25	9.9
17	0.0504	0.0148	0.98	-4.4	0.98	-2.9
18	-0.8459	0.0154	0.95	-8.3	0.93	-8.9
19	-0.5332	0.0149	1.03	5.7	1.04	6.2
21	0.2617	0.0149	1.11	9.9	1.15	9.9
22	-0.8679	0.0154	0.92	-9.9	0.90	-9.9
23	0.4878	0.0152	1.11	9.9	1.15	9.9
24	0.4180	0.0151	1.09	9.9	1.12	9.9
26	-0.4071	0.0148	0.98	-3.4	0.99	-2.2
27	-0.5422	0.0149	1.02	3.6	1.03	4.0
28	-0.1578	0.0147	0.97	-6.3	0.97	-5.7
29	-0.7516	0.0152	1.02	3.4	1.03	4.4
30	-0.2146	0.0147	0.96	-7.7	0.96	-7.9
31	-0.4371	0.0148	0.97	-7.0	0.96	-6.2
32	0.1984	0.0149	1.02	3.6	1.03	5.4
33	-0.0428	0.0147	0.97	-6.3	0.97	-5.6
34	-0.2522	0.0147	1.01	1.8	1.01	1.8
35	0.0664	0.0148	0.96	-9.0	0.95	-8.2
36	-0.0378	0.0147	1.00	0.9	1.01	1.4
37	0.8987	0.0161	1.09	9.9	1.22	9.9
41	1.0527	0.0166	1.10	9.9	1.21	9.9
42	0.3513	0.0150	0.95	-9.0	0.96	-5.4
43	-0.3327	0.0148	1.06	9.9	1.06	9.5
44	-0.4491	0.0148	1.00	-0.5	0.99	-1.0
46	0.9715	0.0164	1.00	-0.5	1.05	4.3
47	0.3750	0.0150	1.05	9.4	1.08	9.9
48	0.0462	0.0148	0.97	-6.2	0.97	-5.6
49	0.2994	0.0150	1.10	9.9	1.13	9.9
50	0.4874	0.0152	1.07	9.9	1.11	9.9
51	0.2725	0.0149	1.01	2.2	1.02	3.2
52	0.2127	0.0149	1.09	9.9	1.11	9.9
53	0.2701	0.0149	0.96	-7.6	0.95	-7.1
54	0.3800	0.0150	1.01	1.9	1.02	3.4
55	-0.1168	0.0147	1.03	6.1	1.03	5.7
56	-0.4084	0.0148	0.94	-9.9	0.93	-9.9
57	-0.0398	0.0147	1.01	2.4	1.02	2.7

## Appendix I: 2011 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	20893	0.459	0.578	0.385	0.037			0.056	0.000	0.001	0.000	0.274	-0.258	0.204	0.148		
20	4	20893	0.984	0.394	0.320	0.207	0.065	0.014	0.099	0.000	0.001	0.000	0.510	-0.470	0.091	0.283	0.233	0.134
25	4	20893	1.581	0.233	0.244	0.286	0.185	0.053	0.082	0.000	0.001	0.000	0.577	-0.475	-0.113	0.155	0.326	0.235
38	2	20893	0.716	0.477	0.330	0.193			0.108	0.000	0.005	0.000	0.450	-0.443	0.207	0.314		
39	2	20893	0.239	0.788	0.185	0.027			0.186	0.000	0.001	0.000	0.335	-0.333	0.283	0.163		
40	2	20893	0.286	0.780	0.154	0.066			0.139	0.000	0.003	0.000	0.356	-0.364	0.265	0.224		
45	4	20893	0.786	0.481	0.310	0.158	0.044	0.007	0.145	0.000	0.003	0.000	0.468	-0.436	0.149	0.268	0.206	0.105
58	2	20893	0.414	0.663	0.261	0.077			0.146	0.000	0.002	0.000	0.428	-0.433	0.318	0.243		
59	2	20893	0.801	0.393	0.414	0.193			0.160	0.000	0.002	0.000	0.470	-0.468	0.221	0.302		

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



## Appendix J: 2011 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	C	6646	0.385	0.495	0.088	0.385	0.029	0.002	0.000	0.298	-0.108	-0.259	0.298	-0.094
2	B	6646	0.702	0.144	0.702	0.083	0.069	0.002	0.000	0.241	-0.187	0.241	-0.060	-0.105
3	A	6646	0.593	0.593	0.166	0.125	0.113	0.003	0.000	0.338	0.338	-0.237	-0.167	-0.067
4	C	6646	0.556	0.082	0.044	0.556	0.316	0.002	0.000	0.370	-0.186	-0.194	0.370	-0.196
5	C	6646	0.532	0.147	0.193	0.532	0.125	0.003	0.000	0.278	-0.117	-0.141	0.278	-0.123
6	B	6646	0.700	0.056	0.700	0.108	0.133	0.002	0.000	0.353	-0.181	0.353	-0.266	-0.107
7	D	6646	0.263	0.246	0.206	0.281	0.263	0.004	0.000	0.249	-0.006	-0.174	-0.079	0.249
8	A	6646	0.527	0.527	0.209	0.221	0.041	0.002	0.000	0.343	0.343	-0.096	-0.248	-0.140
9	B	6646	0.790	0.109	0.790	0.054	0.045	0.002	0.000	0.376	-0.243	0.376	-0.216	-0.130
10	C	6646	0.701	0.103	0.132	0.701	0.062	0.002	0.000	0.327	-0.120	-0.188	0.327	-0.199
11	D	6646	0.423	0.060	0.283	0.232	0.423	0.002	0.000	0.355	-0.220	-0.165	-0.112	0.355
12	B	6646	0.354	0.132	0.354	0.088	0.423	0.002	0.000	0.217	-0.185	0.217	-0.276	0.079

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

## Appendix K: 2011 Grade 12 Fall Writing Retest Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit		OutFit	
			MS	ZSTD	MS	ZSTD
1	1.1315	0.0370	2.34	9.9	9.90	9.9
2	-0.2584	0.0324	2.21	9.9	9.90	9.9
3	0.3061	0.0336	2.27	9.9	9.90	9.9
4	0.1735	0.0332	2.04	9.9	6.53	9.9
5	0.9974	0.0363	2.90	9.9	9.90	9.9
6	-0.5415	0.0322	1.87	9.9	4.62	9.9
7	2.6242	0.0476	3.18	9.9	9.90	9.9
8	0.9157	0.0359	2.65	9.9	9.90	9.9
9	-1.6231	0.0338	1.52	9.9	2.71	6.0
10	-0.4201	0.0323	1.92	9.9	8.22	9.9
11	1.0395	0.0365	2.32	9.9	9.90	9.9
12	2.0469	0.0428	3.28	9.9	9.90	9.9

## Appendix L: 2011 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	6646	2.073	0.221	0.499	0.264	0.016	0.677	-0.512	-0.084	0.505	0.258
13	4	6646	2.093	0.210	0.509	0.261	0.021	0.708	-0.524	-0.110	0.516	0.290
14	4	6646	1.946	0.288	0.491	0.208	0.013	0.657	-0.525	0.036	0.476	0.239
14	4	6646	1.960	0.290	0.479	0.213	0.018	0.702	-0.559	0.024	0.500	0.275

## Appendix M: 2011 Grade 12 Fall Writing Retest Percentage Agreement

Prompt	Composition			Revising & Editing		
	% Exact Agreement	% Adjacent Agreement	% Exact + Adjacent	% Exact Agreement	% Adjacent Agreement	% Exact + Adjacent
1	87	13	100	85	15	100
2	87	13	100	87	13	100

## Appendix N: 2011 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	-5.8526	1.8312	700	378	1	0.0	1	0.0	1
1	-4.6346	1.0100	700	208	1	0.0	2	0.0	1
2	-3.9213	0.7213	700	149	3	0.0	5	0.0	1
3	-3.4956	0.5947	700	123	1	0.0	6	0.0	1
4	-3.1874	0.5201	700	107	0	0.0	6	0.0	1
5	-2.9437	0.4698	700	97	2	0.0	8	0.0	1
6	-2.7406	0.4331	700	89	3	0.0	11	0.0	1
7	-2.5655	0.4049	700	84	8	0.0	19	0.1	1
8	-2.4108	0.3824	705	79	14	0.1	33	0.1	1
9	-2.2716	0.3641	734	75	28	0.1	61	0.2	1
10	-2.1446	0.3489	760	72	41	0.2	102	0.4	1
11	-2.0276	0.3359	785	69	68	0.3	170	0.7	1
12	-1.9185	0.3249	807	67	114	0.4	284	1.1	1
13	-1.8161	0.3153	828	65	164	0.6	448	1.7	1
14	-1.7194	0.3069	848	63	249	1.0	697	2.7	2
15	-1.6274	0.2996	867	62	312	1.2	1009	3.9	3
16	-1.5397	0.2931	885	61	373	1.5	1382	5.4	5
17	-1.4555	0.2874	903	59	417	1.6	1799	7.0	6
18	-1.3744	0.2823	919	58	477	1.9	2276	8.9	8
19	-1.2960	0.2777	936	57	482	1.9	2758	10.7	10
20	-1.2200	0.2737	951	56	531	2.1	3289	12.8	12
21	-1.1461	0.2701	967	56	586	2.3	3875	15.1	14
22	-1.0740	0.2670	981	55	609	2.4	4484	17.5	16
23	-1.0034	0.2642	996	55	634	2.5	5118	19.9	19
24	-0.9343	0.2617	1010	54	613	2.4	5731	22.3	21
25	-0.8664	0.2595	1024	54	637	2.5	6368	24.8	24
26	-0.7996	0.2577	1038	53	687	2.7	7055	27.5	26
27	-0.7336	0.2560	1052	53	692	2.7	7747	30.2	29
28	-0.6684	0.2547	1065	53	683	2.7	8430	32.8	31
29	-0.6039	0.2536	1078	52	743	2.9	9173	35.7	34
30	-0.5398	0.2527	1092	52	717	2.8	9890	38.5	37
31	-0.4761	0.2520	1105	52	736	2.9	10626	41.4	40
32	-0.4127	0.2516	1118	52	756	2.9	11382	44.3	43
33	-0.3495	0.2513	1131	52	815	3.2	12197	47.5	46
34	-0.2864	0.2512	1144	52	814	3.2	13011	50.6	49
35	-0.2233	0.2513	1157	52	781	3.0	13792	53.7	52
36	-0.1601	0.2516	1170	52	870	3.4	14662	57.1	55
37	-0.0966	0.2521	1183	52	878	3.4	15540	60.5	59
38	-0.0329	0.2528	1196	52	852	3.3	16392	63.8	62
39	0.0312	0.2536	1210	52	877	3.4	17269	67.2	66
40	0.0957	0.2546	1223	53	874	3.4	18143	70.6	69
41	0.1609	0.2558	1236	53	793	3.1	18936	73.7	72
42	0.2267	0.2572	1250	53	759	3.0	19695	76.7	75
43	0.2932	0.2587	1264	53	727	2.8	20422	79.5	78
44	0.3606	0.2605	1278	54	702	2.7	21124	82.2	81
45	0.4290	0.2624	1292	54	634	2.5	21758	84.7	83
46	0.4983	0.2645	1306	55	612	2.4	22370	87.1	86
47	0.5689	0.2669	1321	55	557	2.2	22927	89.2	88
48	0.6408	0.2694	1335	56	480	1.9	23407	91.1	90
49	0.7141	0.2722	1351	56	378	1.5	23785	92.6	92
50	0.7890	0.2752	1366	57	359	1.4	24144	94.0	93

**2011 PSSA Grade 12 Retest Technical Report for Mathematics, Reading, Science, and Writing**

---

<b>Raw Score</b>	<b>Measure</b>	<b>Measure SE</b>	<b>Scaled Score</b>	<b>Scaled Score SE</b>	<b>Freq.</b>	<b>Freq. %</b>	<b>Cum. Freq.</b>	<b>Cum. Freq. %</b>	<b>Percentile</b>
51	0.8656	0.2785	1382	57	309	1.2	24453	95.2	95
52	0.9442	0.2821	1398	58	264	1.0	24717	96.2	96
53	1.0249	0.2861	1415	59	235	0.9	24952	97.1	97
54	1.1080	0.2905	1432	60	167	0.6	25119	97.8	97
55	1.1938	0.2953	1450	61	151	0.6	25270	98.3	98
56	1.2825	0.3008	1468	62	81	0.3	25351	98.7	99
57	1.3749	0.3069	1487	63	81	0.3	25432	99.0	99
58	1.4712	0.3139	1507	65	61	0.2	25493	99.2	99
59	1.5722	0.3220	1528	66	36	0.1	25529	99.4	99
60	1.6789	0.3314	1550	68	39	0.2	25568	99.5	99
61	1.7923	0.3425	1573	71	28	0.1	25596	99.6	99
62	1.9141	0.3557	1598	73	23	0.1	25619	99.7	99
63	2.0462	0.3717	1625	77	17	0.1	25636	99.8	99
64	2.1915	0.3912	1655	81	15	0.1	25651	99.8	99
65	2.3539	0.4155	1689	86	10	0.0	25661	99.9	99
66	2.5392	0.4466	1727	92	3	0.0	25664	99.9	99
67	2.7566	0.4875	1772	101	13	0.1	25677	99.9	99
68	3.0209	0.5436	1827	112	9	0.0	25686	100.0	99
69	3.3600	0.6259	1897	129	2	0.0	25688	100.0	99
70	3.8338	0.7617	1994	157	3	0.0	25691	100.0	99
71	4.6234	1.0560	2157	218	3	0.0	25694	100.0	99
72	5.9191	1.8657	2425	385	0	0.0	25694	100.0	100

## Appendix O: 2011 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	-5.6121	1.8376	700	451	1	0.0	1	0.0	1
1	-4.3782	1.0209	700	251	1	0.0	2	0.0	1
2	-3.6435	0.7354	700	181	1	0.0	3	0.0	1
3	-3.1978	0.6108	700	150	2	0.0	5	0.0	1
4	-2.8708	0.5376	700	132	9	0.0	14	0.1	1
5	-2.6089	0.4884	700	120	19	0.1	33	0.2	1
6	-2.3881	0.4527	700	111	40	0.2	73	0.4	1
7	-2.1958	0.4255	700	104	73	0.4	146	0.7	1
8	-2.0240	0.4041	700	99	140	0.7	286	1.5	1
9	-1.8678	0.3868	700	95	167	0.9	453	2.3	2
10	-1.7238	0.3726	700	91	194	1.0	647	3.3	3
11	-1.5895	0.3607	725	89	285	1.5	932	4.8	4
12	-1.4631	0.3508	756	86	282	1.4	1214	6.2	6
13	-1.3431	0.3424	786	84	340	1.7	1554	8.0	7
14	-1.2283	0.3352	814	82	378	1.9	1932	9.9	9
15	-1.1180	0.3291	841	81	449	2.3	2381	12.2	11
16	-1.0115	0.3239	867	80	433	2.2	2814	14.4	13
17	-0.9080	0.3195	892	78	458	2.3	3272	16.8	16
18	-0.8071	0.3158	917	78	484	2.5	3756	19.3	18
19	-0.7084	0.3127	941	77	504	2.6	4260	21.9	21
20	-0.6114	0.3102	965	76	581	3.0	4841	24.8	23
21	-0.5158	0.3082	989	76	595	3.1	5436	27.9	26
22	-0.4214	0.3066	1012	75	626	3.2	6062	31.1	29
23	-0.3278	0.3055	1035	75	641	3.3	6703	34.4	33
24	-0.2347	0.3048	1058	75	700	3.6	7403	38.0	36
25	-0.1419	0.3045	1080	75	711	3.6	8114	41.6	40
26	-0.0492	0.3046	1103	75	808	4.1	8922	45.8	44
27	0.0437	0.3051	1126	75	889	4.6	9811	50.3	48
28	0.1370	0.3060	1149	75	862	4.4	10673	54.8	53
29	0.2310	0.3073	1172	75	857	4.4	11530	59.2	57
30	0.3260	0.3091	1195	76	877	4.5	12407	63.7	61
31	0.4221	0.3112	1219	76	913	4.7	13320	68.3	66
32	0.5198	0.3139	1243	77	972	5.0	14292	73.3	71
33	0.6194	0.3171	1267	78	831	4.3	15123	77.6	75
34	0.7211	0.3209	1292	79	757	3.9	15880	81.5	80
35	0.8254	0.3253	1318	80	745	3.8	16625	85.3	83
36	0.9329	0.3304	1344	81	610	3.1	17235	88.4	87
37	1.0439	0.3362	1371	83	551	2.8	17786	91.3	90
38	1.1592	0.3430	1400	84	424	2.2	18210	93.4	92
39	1.2795	0.3509	1429	86	367	1.9	18577	95.3	94
40	1.4058	0.3600	1460	88	251	1.3	18828	96.6	96
41	1.5392	0.3706	1493	91	182	0.9	19010	97.5	97
42	1.6811	0.3831	1528	94	150	0.8	19160	98.3	98
43	1.8334	0.3979	1565	98	123	0.6	19283	98.9	99
44	1.9988	0.4158	1606	102	60	0.3	19343	99.2	99
45	2.1805	0.4376	1650	107	49	0.3	19392	99.5	99
46	2.3838	0.4651	1700	114	34	0.2	19426	99.7	99
47	2.6162	0.5008	1757	123	31	0.2	19457	99.8	99
48	2.8908	0.5497	1825	135	10	0.1	19467	99.9	99
49	3.2314	0.6221	1908	153	11	0.1	19478	99.9	99
50	3.6914	0.7453	2021	183	8	0.0	19486	100.0	99

***2011 PSSA Grade 12 Retest Technical Report for Mathematics, Reading, Science, and Writing***

---

<b>Raw Score</b>	<b>Measure</b>	<b>Measure SE</b>	<b>Scaled Score</b>	<b>Scaled Score SE</b>	<b>Freq.</b>	<b>Freq. %</b>	<b>Cum. Freq.</b>	<b>Cum. Freq. %</b>	<b>Percentile</b>
51	4.4410	1.0283	2205	252	2	0.0	19488	100.0	99
52	5.6857	1.8418	2511	452	3	0.0	19491	100.0	99



## Appendix P: 2011 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	-5.4748	1.8323	1050	187	1	0.0	1	0.0	1
1	-4.2541	1.0118	1050	103	0	0.0	1	0.0	1
2	-3.5372	0.7236	1050	74	0	0.0	1	0.0	1
3	-3.1082	0.5974	1050	61	0	0.0	1	0.0	1
4	-2.7969	0.5230	1050	53	0	0.0	1	0.0	1
5	-2.5503	0.4727	1050	48	6	0.0	7	0.0	1
6	-2.3445	0.4360	1050	44	13	0.1	20	0.1	1
7	-2.1670	0.4078	1050	42	30	0.1	50	0.2	1
8	-2.0099	0.3853	1050	39	54	0.3	104	0.5	1
9	-1.8687	0.3669	1050	37	97	0.5	201	1.0	1
10	-1.7398	0.3515	1050	36	179	0.9	380	1.8	1
11	-1.6209	0.3385	1050	34	212	1.0	592	2.8	2
12	-1.5102	0.3272	1050	33	290	1.4	882	4.2	4
13	-1.4064	0.3174	1052	32	359	1.7	1241	5.9	5
14	-1.3083	0.3089	1061	31	407	1.9	1648	7.9	7
15	-1.2153	0.3013	1071	31	430	2.1	2078	9.9	9
16	-1.1265	0.2946	1080	30	487	2.3	2565	12.3	11
17	-1.0415	0.2887	1089	29	542	2.6	3107	14.9	14
18	-0.9597	0.2833	1097	29	537	2.6	3644	17.4	16
19	-0.8808	0.2785	1105	28	528	2.5	4172	20.0	19
20	-0.8044	0.2742	1113	28	514	2.5	4686	22.4	21
21	-0.7303	0.2704	1120	28	582	2.8	5268	25.2	24
22	-0.6582	0.2669	1128	27	558	2.7	5826	27.9	27
23	-0.5878	0.2638	1135	27	551	2.6	6377	30.5	29
24	-0.5190	0.2610	1142	27	557	2.7	6934	33.2	32
25	-0.4515	0.2585	1149	26	583	2.8	7517	36.0	35
26	-0.3853	0.2562	1155	26	582	2.8	8099	38.8	37
27	-0.3202	0.2543	1162	26	558	2.7	8657	41.4	40
28	-0.2559	0.2525	1169	26	567	2.7	9224	44.1	43
29	-0.1925	0.2510	1175	26	632	3.0	9856	47.2	46
30	-0.1299	0.2497	1181	25	586	2.8	10442	50.0	49
31	-0.0678	0.2486	1188	25	583	2.8	11025	52.8	51
32	-0.0062	0.2477	1194	25	620	3.0	11645	55.7	54
33	0.0549	0.2470	1200	25	601	2.9	12246	58.6	57
34	0.1158	0.2464	1206	25	631	3.0	12877	61.6	60
35	0.1764	0.2461	1213	25	598	2.9	13475	64.5	63
36	0.2369	0.2459	1219	25	649	3.1	14124	67.6	66
37	0.2973	0.2458	1225	25	615	2.9	14739	70.5	69
38	0.3578	0.2460	1231	25	613	2.9	15352	73.5	72
39	0.4184	0.2463	1237	25	561	2.7	15913	76.2	75
40	0.4792	0.2468	1243	25	558	2.7	16471	78.8	77
41	0.5402	0.2474	1250	25	576	2.8	17047	81.6	80
42	0.6017	0.2483	1256	25	556	2.7	17603	84.3	83
43	0.6635	0.2493	1262	25	496	2.4	18099	86.6	85
44	0.7260	0.2505	1269	26	463	2.2	18562	88.8	88
45	0.7891	0.2519	1275	26	389	1.9	18951	90.7	90
46	0.8529	0.2535	1282	26	376	1.8	19327	92.5	92
47	0.9176	0.2553	1288	26	287	1.4	19614	93.9	93
48	0.9833	0.2573	1295	26	266	1.3	19880	95.2	95
49	1.0501	0.2596	1302	26	229	1.1	20109	96.2	96
50	1.1181	0.2622	1309	27	187	0.9	20296	97.1	97

**2011 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	1.1876	0.2650	1316	27	160	0.8	20456	97.9	98
52	1.2587	0.2682	1323	27	105	0.5	20561	98.4	98
53	1.3315	0.2717	1330	28	81	0.4	20642	98.8	99
54	1.4064	0.2756	1338	28	65	0.3	20707	99.1	99
55	1.4835	0.2800	1346	29	46	0.2	20753	99.3	99
56	1.5633	0.2848	1354	29	35	0.2	20788	99.5	99
57	1.6459	0.2902	1362	30	35	0.2	20823	99.7	99
58	1.7319	0.2963	1371	30	24	0.1	20847	99.8	99
59	1.8216	0.3031	1380	31	15	0.1	20862	99.9	99
60	1.9159	0.3108	1390	32	8	0.0	20870	99.9	99
61	2.0151	0.3195	1400	33	3	0.0	20873	99.9	99
62	2.1204	0.3296	1411	34	4	0.0	20877	99.9	99
63	2.2327	0.3411	1422	35	8	0.0	20885	100.0	99
64	2.3536	0.3545	1434	36	4	0.0	20889	100.0	99
65	2.4848	0.3704	1448	38	2	0.0	20891	100.0	99
66	2.6289	0.3893	1462	40	1	0.0	20892	100.0	99
67	2.7894	0.4125	1479	42	0	0.0	20892	100.0	99
68	2.9713	0.4415	1497	45	0	0.0	20892	100.0	99
69	3.1824	0.4790	1519	49	1	0.0	20893	100.0	99
70	3.4358	0.5302	1544	54	0	0.0	20893	100.0	100
71	3.7556	0.6054	1577	62	0	0.0	20893	100.0	100
72	4.1956	0.7321	1622	75	0	0.0	20893	100.0	100
73	4.9266	1.0196	1696	104	0	0.0	20893	100.0	100
74	6.1592	1.8374	1822	187	0	0.0	20893	100.0	100

## Appendix Q: 2011 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	-7.0630	1.8366	700	184	9	0.1	9	0.1	1
23	-5.8307	1.0203	700	102	43	0.6	52	0.8	1
24	-5.0959	0.7365	735	74	122	1.8	174	2.6	2
25	-4.6472	0.6143	780	61	179	2.7	353	5.3	4
26	-4.3147	0.5438	813	54	167	2.5	520	7.8	7
27	-4.0450	0.4974	840	50	166	2.5	686	10.3	9
28	-3.8142	0.4648	863	46	120	1.8	806	12.1	11
29	-3.6098	0.4407	883	44	107	1.6	913	13.7	13
30	-3.4239	0.4224	902	42	56	0.8	969	14.6	14
31	-3.2515	0.4084	919	41	36	0.5	1005	15.1	15
32	-3.0893	0.3975	935	40	23	0.3	1028	15.5	15
33	-2.9348	0.3891	951	39	18	0.3	1046	15.7	16
34	-2.7861	0.3826	966	38	48	0.7	1094	16.5	16
35	-2.6416	0.3776	980	38	74	1.1	1168	17.6	17
36	-2.5005	0.3739	994	37	109	1.6	1277	19.2	18
37	-2.3617	0.3712	1008	37	170	2.6	1447	21.8	20
38	-2.2246	0.3693	1022	37	176	2.6	1623	24.4	23
39	-2.0888	0.3679	1035	37	185	2.8	1808	27.2	26
40	-1.9538	0.3669	1049	37	134	2.0	1942	29.2	28
41	-1.8195	0.3661	1062	37	117	1.8	2059	31.0	30
42	-1.6858	0.3653	1076	37	81	1.2	2140	32.2	32
43	-1.5526	0.3645	1089	36	55	0.8	2195	33.0	33
44	-1.4201	0.3634	1102	36	45	0.7	2240	33.7	33
45	-1.2886	0.3620	1115	36	40	0.6	2280	34.3	34
46	-1.1582	0.3603	1128	36	86	1.3	2366	35.6	35
47	-1.0291	0.3583	1141	36	142	2.1	2508	37.7	37
48	-0.9015	0.3560	1154	36	237	3.6	2745	41.3	40
49	-0.7757	0.3535	1167	35	268	4.0	3013	45.3	43
50	-0.6516	0.3509	1179	35	330	5.0	3343	50.3	48
51	-0.5294	0.3484	1191	35	345	5.2	3688	55.5	53
52	-0.4088	0.3460	1203	35	276	4.2	3964	59.6	58
53	-0.2899	0.3437	1215	34	198	3.0	4162	62.6	61
54	-0.1725	0.3418	1227	34	174	2.6	4336	65.2	64
55	-0.0562	0.3402	1239	34	82	1.2	4418	66.5	66
56	0.0591	0.3390	1250	34	35	0.5	4453	67.0	67
57	0.1737	0.3383	1262	34	43	0.6	4496	67.6	67
58	0.2881	0.3382	1273	34	53	0.8	4549	68.4	68
59	0.4026	0.3386	1285	34	85	1.3	4634	69.7	69
60	0.5176	0.3397	1296	34	125	1.9	4759	71.6	71
61	0.6336	0.3415	1308	34	139	2.1	4898	73.7	73
62	0.7510	0.3441	1319	34	157	2.4	5055	76.1	75
63	0.8706	0.3476	1331	35	158	2.4	5213	78.4	77
64	0.9929	0.3520	1344	35	147	2.2	5360	80.7	80
65	1.1187	0.3576	1356	36	115	1.7	5475	82.4	82
66	1.2489	0.3643	1369	36	63	0.9	5538	83.3	83
67	1.3846	0.3726	1383	37	41	0.6	5579	83.9	84
68	1.5271	0.3826	1397	38	18	0.3	5597	84.2	84
69	1.6780	0.3947	1412	39	19	0.3	5616	84.5	84
70	1.8395	0.4094	1428	41	34	0.5	5650	85.0	85
71	2.0142	0.4272	1446	43	51	0.8	5701	85.8	85

**2011 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	2.2058	0.4489	1465	45	108	1.6	5809	87.4	87
73	2.4192	0.4758	1486	48	108	1.6	5917	89.0	88
74	2.6612	0.5089	1510	51	140	2.1	6057	91.1	90
75	2.9406	0.5493	1538	55	149	2.2	6206	93.4	92
76	3.2681	0.5955	1571	60	116	1.7	6322	95.1	94
77	3.6500	0.6383	1609	64	112	1.7	6434	96.8	96
78	4.0732	0.6567	1652	66	59	0.9	6493	97.7	97
79	4.4954	0.6370	1694	64	14	0.2	6507	97.9	98
80	4.8759	0.5950	1732	60	0	0.0	6507	97.9	98
81	5.2043	0.5515	1765	55	1	0.0	6508	97.9	98
82	5.4881	0.5150	1793	52	2	0.0	6510	98.0	98
83	5.7382	0.4864	1818	49	1	0.0	6511	98.0	98
84	5.9638	0.4646	1841	46	6	0.1	6517	98.1	98
85	6.1719	0.4485	1861	45	5	0.1	6522	98.1	98
86	6.3676	0.4369	1881	44	18	0.3	6540	98.4	98
87	6.5549	0.4293	1900	43	18	0.3	6558	98.7	99
88	6.7372	0.4252	1918	43	17	0.3	6575	98.9	99
89	6.9174	0.4243	1936	42	24	0.4	6599	99.3	99
90	7.0983	0.4268	1954	43	12	0.2	6611	99.5	99
91	7.2827	0.4327	1973	43	0	0.0	6611	99.5	99
92	7.4740	0.4427	1992	44	0	0.0	6611	99.5	99
93	7.6761	0.4575	2012	46	0	0.0	6611	99.5	99
94	7.8947	0.4786	2034	48	0	0.0	6611	99.5	99
95	8.1376	0.5088	2058	51	0	0.0	6611	99.5	99
96	8.4180	0.5529	2086	55	1	0.0	6612	99.5	99
97	8.7599	0.6215	2120	62	7	0.1	6619	99.6	99
98	9.2170	0.7418	2166	74	6	0.1	6625	99.7	99
99	9.9593	1.0237	2240	102	6	0.1	6631	99.8	99
100	11.1964	1.8384	2364	184	15	0.2	6646	100.0	99