TECHNICAL REPORT

**for the
Pennsylvania
System of School Assessment**

**2012 Grade 12 Fall Retest
Mathematics, Reading, Science, and Writing**

**Provided by
Data Recognition Corporation**

April 2012

Table of Contents

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 *P*-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 *p*-value (see also *P*-value). The logit difficulty scale is inversely related to *p*-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.
P-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. *P*-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. *P*-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 *p*-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 2012 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 2012 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 2012 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 2012 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 s*caled 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 2012 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 2012 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	125113	23186	125380	18305	121693	18453	125095	6173
Mean	1375.8	1133.9	1370.0	1129.2	1245.9	1186.5	1523.3	1155.3
St. Dev.	267.0	152.5	258.6	218.6	93.5	75.7	293.9	233.2
Min	700	700	700	700	1050	1050	700	700
Max	2587	1994	2520	2214	1617	1479	2418	2418
Reli.	0.94	0.86	0.88	0.86	0.92	0.87	0.83	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 2012 PSSA Technical Report.

Figure 1: Mathematics Operational and Retest Scaled Score Frequency Distributions
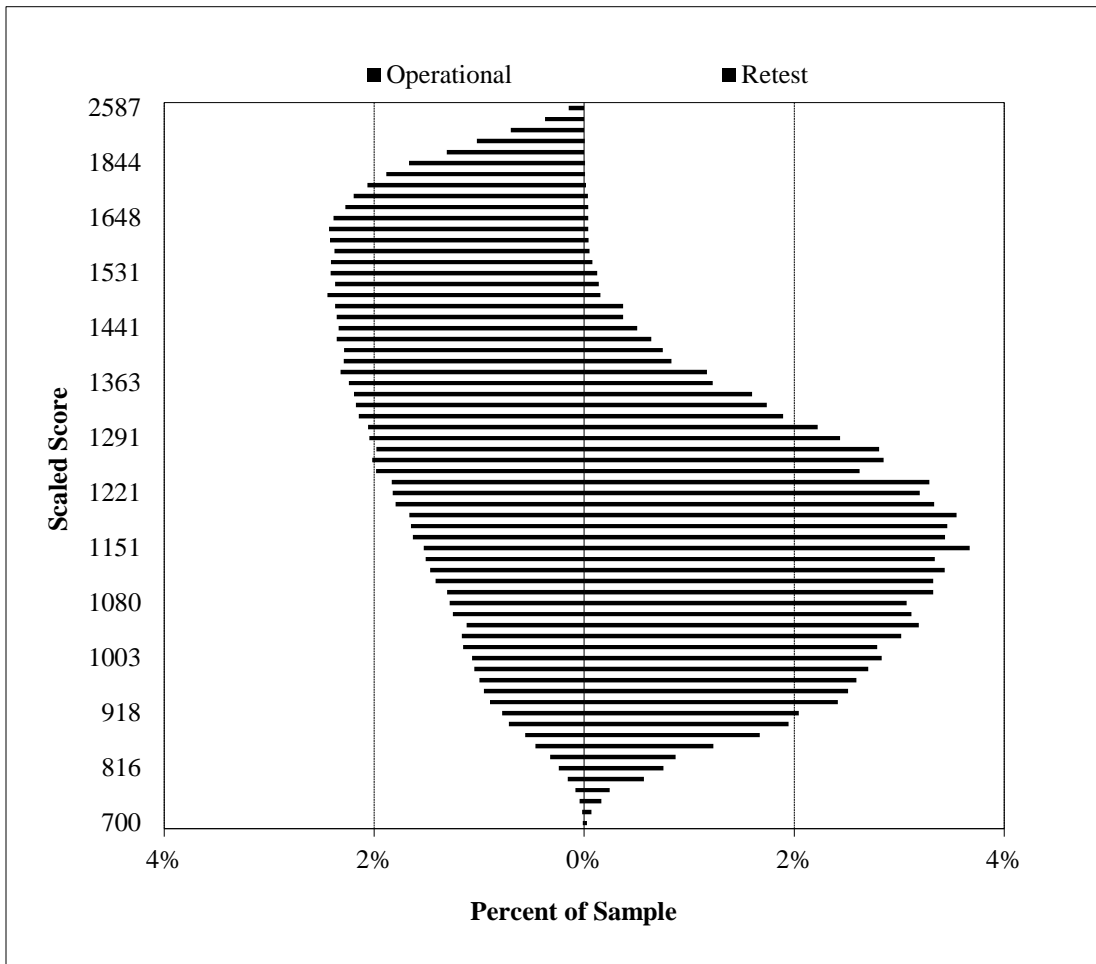
Figure 2: Reading Operational and Retest Scaled Score Frequency Distribution
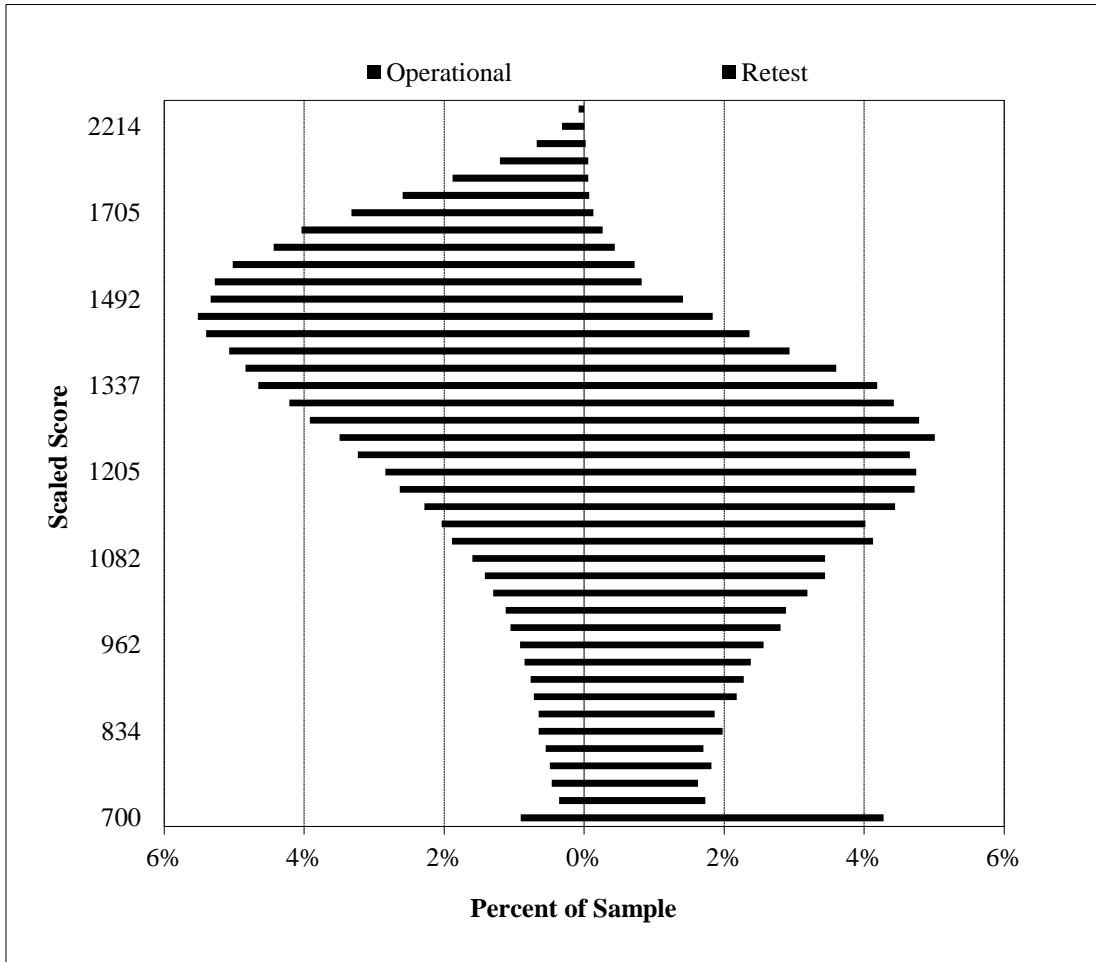
Figure 3: Science Operational and Retest Scaled Score Frequency Distribution
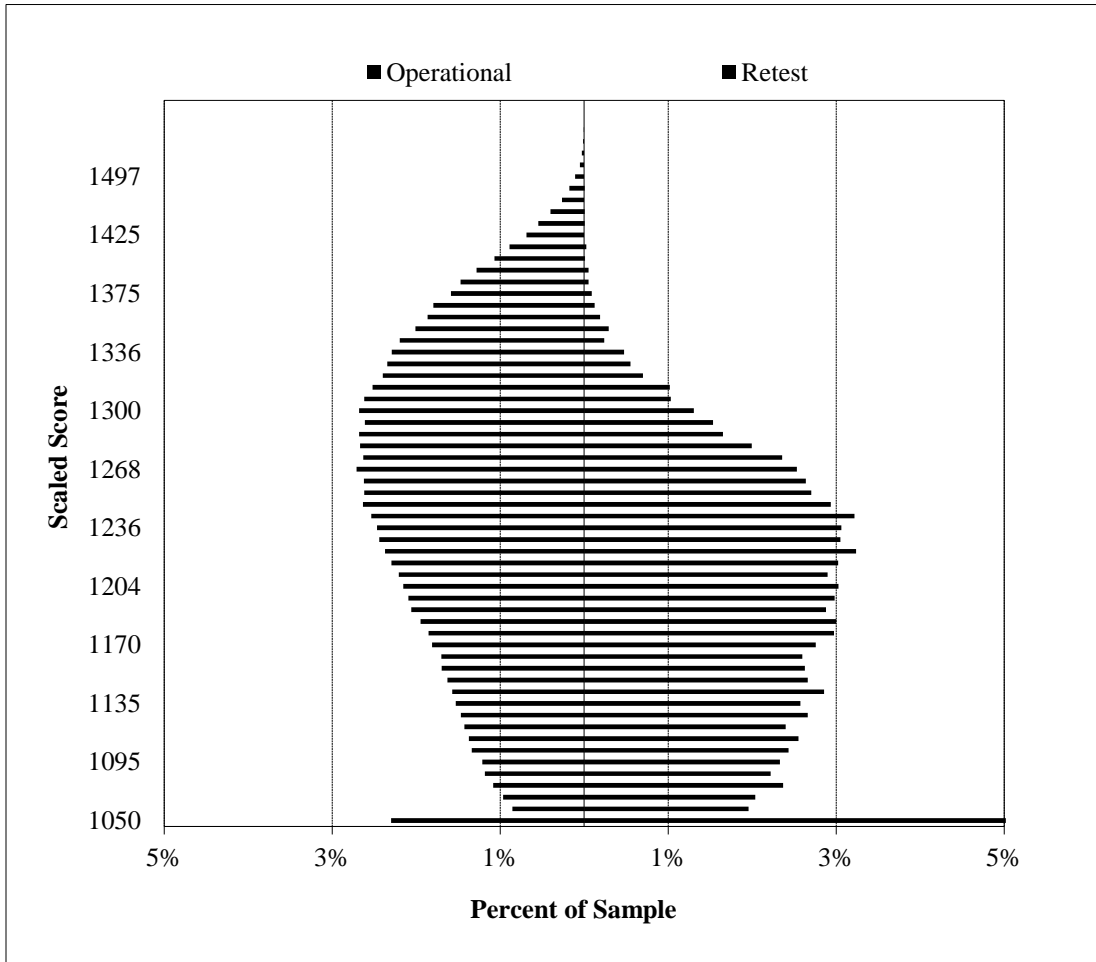
Figure 4: Writing Operational and Retest Scaled Score Frequency Distribution
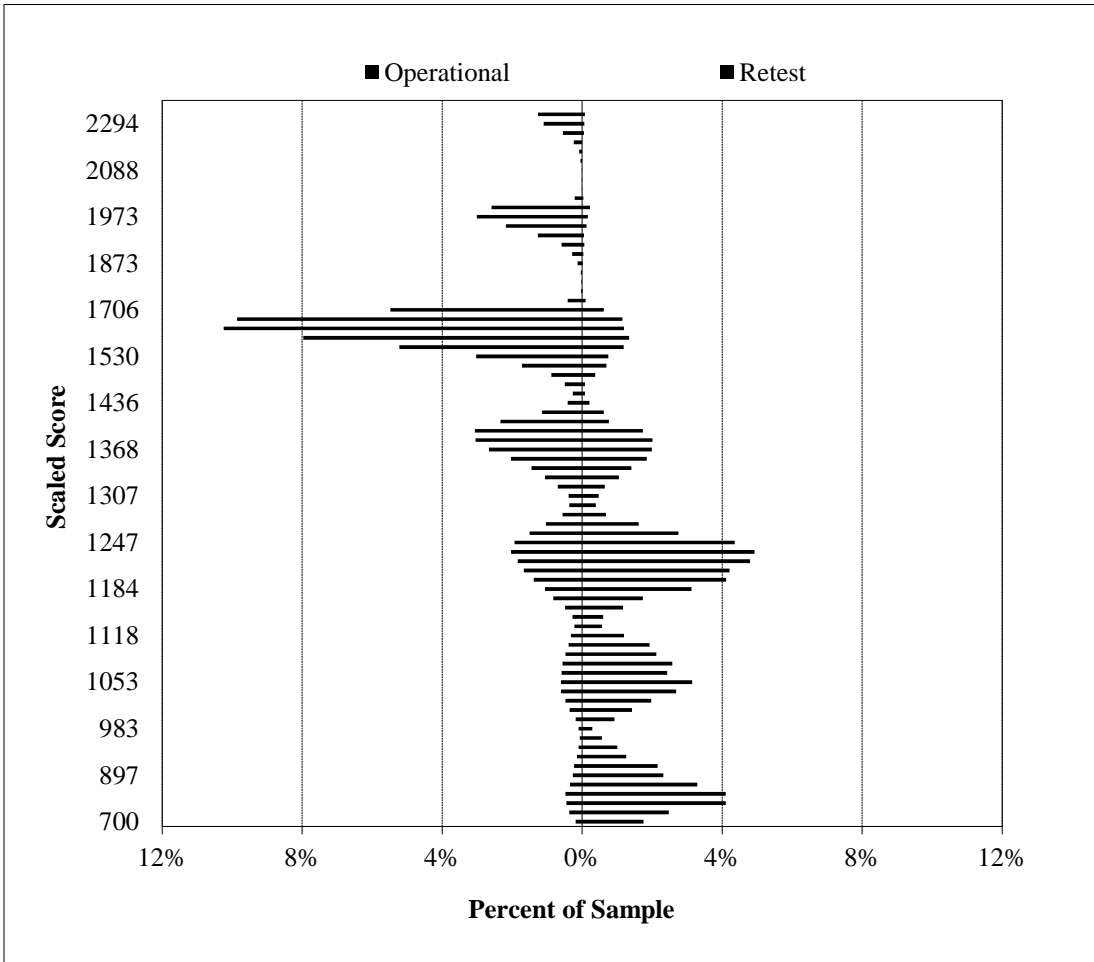
Performance Level Results

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 14.1, 28.1, 13.7, and 31.0, respectively. (Numbers may be slightly different than tabled numbers due to rounding.) Approximately 69%–86% 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

Performance Level	Mathematics		
	Scaled Score	Frequency	Percent
Advanced	1509 and up	145	0.6
Proficient	1304-1508	3124	13.5
Basic	1167-1303	6385	27.5
Below Basic	1166 and below	13532	58.4

Performance Level	Reading		
	Scaled Score	Frequency	Percent
Advanced	1492 and up	732	4.0
Proficient	1257-1491	4416	24.1
Basic	1112-1256	5049	27.6
Below Basic	1111 and below	8108	44.3

Performance Level	Science		
	Scaled Score	Frequency	Percent
Advanced	1347 and up	159	0.9
Proficient	1275-1346	2373	12.9
Basic	1150-1274	10114	54.8
Below Basic	1149 and below	5807	31.5

Performance Level	Writing		
	Scaled Score	Frequency	Percent
Advanced	1806 and up	56	0.9
Proficient	1236-1805	1860	30.1
Basic	952-1235	2933	47.5
Below Basic	951 and below	1324	21.4

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, 62.1% of the students remained at the same performance level in mathematics, while 21.8% transitioned to a higher level and 16.1% regressed to a lower level. For reading, 48.5% of the students stayed at the same level, 36.4% improved and 15.1% regressed. For science, 67.6% of the students stayed at the same level, 18.6% improved and 13.8% regressed. For writing, 53.2% of the students remained at the same level, 29.3% improved and 17.6% regressed.

Appendix A: 2012 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	23186	0.414	0.086	0.414	0.109	0.390	0.001	0.000	0.122	-0.182	0.122	-0.162	0.087
2	D	23186	0.565	0.074	0.232	0.128	0.565	0.001	0.000	0.347	-0.208	-0.215	-0.076	0.347
3	B	23186	0.727	0.065	0.727	0.164	0.041	0.002	0.000	0.287	-0.161	0.287	-0.171	-0.118
4	A	23186	0.553	0.553	0.223	0.155	0.067	0.002	0.000	0.232	0.232	-0.038	-0.173	-0.141
5	A	23186	0.779	0.779	0.135	0.055	0.028	0.001	0.001	0.396	0.396	-0.263	-0.216	-0.143
6	D	23186	0.378	0.173	0.166	0.276	0.378	0.006	0.000	0.279	-0.116	-0.096	-0.122	0.279
7	D	23186	0.364	0.172	0.247	0.215	0.364	0.001	0.000	0.198	-0.111	-0.126	0.004	0.198
8	C	23186	0.350	0.192	0.185	0.350	0.268	0.005	0.000	0.161	-0.124	-0.074	0.161	0.002
9	D	23186	0.350	0.111	0.440	0.095	0.350	0.003	0.000	0.137	-0.195	0.100	-0.178	0.137
10	C	23186	0.663	0.055	0.123	0.663	0.156	0.002	0.000	0.277	-0.123	-0.163	0.277	-0.133
11	D	23186	0.537	0.161	0.156	0.142	0.537	0.004	0.000	0.375	-0.179	-0.187	-0.150	0.375
12	C	23186	0.795	0.047	0.092	0.795	0.064	0.001	0.000	0.237	-0.194	-0.089	0.237	-0.112
13	C	23186	0.461	0.159	0.252	0.461	0.123	0.005	0.000	0.254	-0.144	-0.138	0.254	-0.041
14	B	23186	0.345	0.160	0.345	0.297	0.195	0.003	0.000	0.146	-0.095	0.146	-0.037	-0.042
15	C	23186	0.343	0.381	0.194	0.343	0.078	0.003	0.000	0.118	0.111	-0.166	0.118	-0.157
16	D	23186	0.219	0.104	0.065	0.610	0.219	0.002	0.000	0.229	-0.138	-0.165	-0.021	0.229
17	C	23186	0.806	0.062	0.072	0.806	0.058	0.002	0.000	0.284	-0.180	-0.189	0.284	-0.079
18	C	23186	0.378	0.242	0.263	0.378	0.110	0.006	0.000	0.224	-0.013	-0.138	0.224	-0.132
19	A	23186	0.761	0.761	0.079	0.068	0.089	0.002	0.001	0.337	0.337	-0.213	-0.218	-0.103
20	B	23186	0.478	0.129	0.478	0.220	0.169	0.003	0.000	0.263	-0.061	0.263	-0.120	-0.157
21	D	23186	0.326	0.356	0.153	0.161	0.326	0.003	0.001	0.237	-0.069	-0.091	-0.116	0.237
22	D	23186	0.398	0.079	0.373	0.144	0.398	0.005	0.001	0.183	-0.159	0.006	-0.137	0.183
23	A	23186	0.386	0.386	0.216	0.186	0.207	0.005	0.001	0.226	0.226	-0.096	-0.020	-0.150
26	D	23186	0.784	0.086	0.074	0.054	0.784	0.001	0.000	0.382	-0.178	-0.223	-0.210	0.382
27	A	23186	0.629	0.629	0.120	0.222	0.027	0.002	0.000	0.340	0.340	-0.237	-0.157	-0.133
28	B	23186	0.567	0.136	0.567	0.177	0.116	0.005	0.000	0.331	-0.115	0.331	-0.208	-0.134
29	D	23186	0.254	0.104	0.187	0.451	0.254	0.003	0.000	0.230	-0.193	-0.212	0.087	0.230
30	B	23186	0.536	0.219	0.536	0.143	0.099	0.003	0.000	0.287	-0.108	0.287	-0.196	-0.098
31	C	23186	0.499	0.125	0.174	0.499	0.197	0.003	0.000	0.296	-0.161	-0.149	0.296	-0.092
32	A	23186	0.501	0.501	0.170	0.107	0.220	0.002	0.000	0.274	0.274	-0.100	-0.193	-0.094
33	D	23186	0.581	0.045	0.300	0.072	0.581	0.002	0.001	0.315	-0.208	-0.109	-0.234	0.315
34	B	23186	0.679	0.146	0.679	0.095	0.076	0.002	0.001	0.310	-0.176	0.310	-0.099	-0.194
35	B	23186	0.505	0.193	0.505	0.194	0.103	0.004	0.001	0.302	-0.179	0.302	-0.133	-0.085
36	B	23186	0.545	0.338	0.545	0.068	0.047	0.001	0.001	0.407	-0.254	0.407	-0.213	-0.127

Item Description			Proportions							Point Biserials				
Seq.	Key	N	P-Value	A	B	C	D	-	*	Tot. Corr.	A	B	C	D
37	C	23186	0.491	0.186	0.188	0.491	0.130	0.004	0.000	0.288	-0.093	-0.200	0.288	-0.081
38	B	23186	0.403	0.304	0.403	0.198	0.089	0.006	0.000	0.114	0.033	0.114	-0.091	-0.114
39	B	23186	0.462	0.310	0.462	0.130	0.094	0.003	0.001	0.216	-0.019	0.216	-0.161	-0.140
40	A	23186	0.511	0.511	0.146	0.159	0.179	0.004	0.000	0.432	0.432	-0.219	-0.241	-0.122
41	B	23186	0.452	0.170	0.452	0.136	0.236	0.005	0.001	0.100	-0.012	0.100	-0.140	0.013
42	C	23186	0.542	0.155	0.243	0.542	0.056	0.003	0.001	0.345	-0.199	-0.139	0.345	-0.158
43	C	23186	0.405	0.387	0.101	0.405	0.103	0.003	0.001	0.130	0.128	-0.238	0.130	-0.167
44	B	23186	0.454	0.206	0.454	0.262	0.072	0.005	0.001	0.192	-0.141	0.192	-0.012	-0.116
45	B	23186	0.430	0.133	0.430	0.366	0.068	0.003	0.000	0.149	-0.226	0.149	0.042	-0.057
46	D	23186	0.151	0.406	0.299	0.140	0.151	0.004	0.000	0.129	0.070	-0.062	-0.139	0.129
47	A	23186	0.720	0.720	0.073	0.142	0.060	0.004	0.001	0.463	0.463	-0.256	-0.255	-0.197
48	A	23186	0.318	0.318	0.267	0.228	0.180	0.007	0.000	0.309	0.309	-0.055	-0.153	-0.131
49	C	23186	0.606	0.197	0.136	0.606	0.056	0.004	0.001	0.371	-0.203	-0.157	0.371	-0.179
50	A	23186	0.238	0.238	0.234	0.157	0.365	0.006	0.001	0.193	0.193	-0.078	-0.122	-0.001
51	D	23186	0.191	0.197	0.147	0.459	0.191	0.006	0.001	0.180	-0.148	-0.177	0.113	0.180
52	C	23186	0.509	0.095	0.282	0.509	0.108	0.006	0.000	0.271	-0.189	-0.068	0.271	-0.142
53	C	23186	0.418	0.195	0.248	0.418	0.131	0.008	0.001	0.168	-0.092	-0.044	0.168	-0.065
54	A	23186	0.534	0.534	0.182	0.154	0.123	0.007	0.001	0.350	0.350	-0.159	-0.199	-0.107
55	A	23186	0.430	0.430	0.196	0.259	0.108	0.006	0.001	0.275	0.275	-0.079	-0.119	-0.149
56	C	23186	0.469	0.101	0.280	0.469	0.143	0.007	0.000	0.285	-0.184	-0.113	0.285	-0.086
57	D	23186	0.417	0.156	0.256	0.164	0.417	0.007	0.000	0.307	-0.204	-0.030	-0.156	0.307
58	D	23186	0.448	0.166	0.113	0.266	0.448	0.006	0.001	0.319	-0.110	-0.233	-0.085	0.319
59	A	23186	0.610	0.610	0.142	0.150	0.092	0.006	0.000	0.435	0.435	-0.242	-0.189	-0.188
60	B	23186	0.415	0.188	0.415	0.209	0.181	0.007	0.001	0.270	-0.122	0.270	-0.084	-0.121
61	A	23186	0.585	0.585	0.124	0.225	0.060	0.005	0.001	0.299	0.299	-0.136	-0.137	-0.169
62	B	23186	0.466	0.115	0.466	0.146	0.265	0.007	0.000	0.311	-0.133	0.311	-0.165	-0.110

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

Appendix B: 2012 Grade 12 Fall Mathematics Retest
Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit MS	InFit ZSTD	OutFit MS	OutFit ZSTD
1	0.5099	0.0148	1.27	9.9	1.39	9.9
2	-0.6361	0.0140	0.95	-9.9	0.95	-9.8
3	-1.3389	0.0152	0.94	-8.4	0.93	-7.3
4	-0.0015	0.0141	1.12	9.9	1.15	9.9
5	-1.8111	0.0168	0.92	-8.4	0.84	-9.9
6	-0.0218	0.0140	0.98	-5.4	0.98	-4.0
7	0.5774	0.0150	1.15	9.9	1.22	9.9
8	0.6689	0.0152	1.20	9.9	1.30	9.9
9	0.4961	0.0148	1.14	9.9	1.21	9.9
10	-0.7668	0.0142	0.95	-9.9	0.95	-8.8
11	-0.5254	0.0140	0.94	-9.9	0.93	-9.9
12	-1.5037	0.0157	0.87	-9.9	0.87	-9.9
13	-0.0536	0.0140	1.03	7.6	1.05	8.4
14	0.4551	0.0147	1.11	9.9	1.18	9.9
15	1.0155	0.0163	1.45	9.9	1.69	9.9
16	1.1298	0.0168	1.03	3.1	1.10	7.3
17	-1.6472	0.0162	0.84	-9.9	0.84	-9.9
18	0.2943	0.0144	1.05	9.9	1.09	9.9
19	-1.7410	0.0165	0.99	-1.4	0.95	-3.6
20	-0.1681	0.0140	1.02	5.0	1.03	5.2
21	0.1338	0.0142	0.97	-5.8	0.97	-4.4
22	0.5627	0.0150	1.23	9.9	1.30	9.9
23	0.1235	0.0142	1.03	6.0	1.04	6.4
26	-1.7409	0.0165	0.88	-9.9	0.82	-9.9
27	-1.2563	0.0150	1.07	9.9	1.07	7.2
28	-0.7260	0.0141	0.98	-4.2	0.97	-4.5
29	0.7582	0.0155	0.98	-3.5	1.00	-0.2
30	-0.3524	0.0139	1.00	-0.2	1.00	-0.1
31	-0.3168	0.0139	0.99	-1.4	0.99	-1.9
32	-0.1288	0.0140	1.03	6.2	1.03	5.3
33	-0.6225	0.0140	0.96	-8.0	0.97	-5.8
34	-1.3214	0.0152	1.02	2.4	1.03	2.7
35	-0.4409	0.0139	0.99	-1.9	1.00	-0.6
36	-0.6709	0.0141	0.93	-9.9	0.92	-9.9
37	-0.4281	0.0139	1.01	1.5	1.01	1.0
38	0.2079	0.0143	1.14	9.9	1.19	9.9
39	0.0981	0.0142	1.10	9.9	1.12	9.9
40	-0.8768	0.0143	1.01	1.1	0.99	-0.8
41	-0.1126	0.0140	1.14	9.9	1.16	9.9
42	-0.2526	0.0139	0.97	-7.6	0.97	-6.6
43	0.5943	0.0150	1.30	9.9	1.41	9.9
44	0.1632	0.0142	1.13	9.9	1.16	9.9
45	0.1775	0.0143	1.14	9.9	1.18	9.9
46	1.3056	0.0176	0.90	-9.6	1.03	2.2
47	-1.9632	0.0175	1.22	9.9	1.08	5.5
48	0.1491	0.0142	0.92	-9.9	0.91	-9.9
49	-0.9302	0.0144	0.96	-8.4	0.95	-7.8
50	0.5687	0.0150	0.92	-9.9	0.95	-5.8
51	1.2503	0.0173	1.02	1.7	1.12	8.3
52	-0.2924	0.0139	1.01	3.1	1.02	3.4
53	0.0494	0.0141	1.09	9.9	1.11	9.9

Seq.	Anchored Measure	Measure SE	InFit MS	InFit ZSTD	OutFit MS	OutFit ZSTD
54	-0.6816	0.0141	0.98	-4.4	0.98	-4.2
55	-0.4214	0.0139	1.03	6.9	1.03	5.7
56	-0.4175	0.0139	1.01	2.9	1.01	2.3
57	-0.2016	0.0139	0.97	-6.1	0.97	-5.0
58	-0.2890	0.0139	0.97	-6.1	0.97	-5.3
59	-1.1322	0.0147	0.98	-4.1	0.95	-6.4
60	-0.1275	0.0140	1.00	-0.5	1.00	0.4
61	-0.7990	0.0142	1.00	-0.5	1.00	0.4
62	-0.3242	0.0139	0.98	-4.0	0.98	-3.3

Appendix C: 2012 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	23186	0.540	0.623	0.263	0.075	0.029	0.010	0.039	0.000	0.000	0.408	-0.376	0.189	0.208	0.181	0.140
25	4	23186	0.941	0.385	0.415	0.075	0.125	0.000	0.058	0.001	0.000	0.492	-0.497	0.208	0.153	0.297	0.058
63	4	23186	0.526	0.555	0.372	0.064	0.008	0.000	0.067	0.001	0.000	0.500	-0.482	0.339	0.259	0.125	0.041

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

Appendix D: 2012 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	B	18305	0.452	0.315	0.452	0.041	0.190	0.002	0.000	0.277	-0.139	0.277	-0.202	-0.082
2	A	18305	0.299	0.299	0.598	0.057	0.045	0.001	0.000	0.169	0.169	0.044	-0.268	-0.172
3	B	18305	0.874	0.038	0.874	0.064	0.023	0.001	0.000	0.415	-0.217	0.415	-0.252	-0.221
4	C	18305	0.488	0.113	0.240	0.488	0.155	0.003	0.001	0.196	-0.085	0.025	0.196	-0.221
5	A	18305	0.443	0.443	0.391	0.098	0.065	0.002	0.001	0.168	0.168	0.017	-0.173	-0.155
6	A	18305	0.548	0.548	0.137	0.167	0.146	0.002	0.001	0.327	0.327	-0.139	-0.164	-0.146
7	D	18305	0.679	0.073	0.152	0.093	0.679	0.002	0.000	0.383	-0.252	-0.110	-0.248	0.383
9	A	18305	0.551	0.551	0.212	0.041	0.195	0.001	0.000	0.282	0.282	-0.160	-0.253	-0.060
10	C	18305	0.520	0.275	0.129	0.520	0.074	0.001	0.000	0.263	-0.040	-0.233	0.263	-0.128
11	B	18305	0.814	0.050	0.814	0.068	0.067	0.001	0.000	0.423	-0.245	0.423	-0.227	-0.207
12	A	18305	0.570	0.570	0.110	0.195	0.123	0.002	0.000	0.321	0.321	-0.218	-0.147	-0.092
13	B	18305	0.446	0.049	0.446	0.190	0.313	0.002	0.000	0.259	-0.245	0.259	-0.181	-0.006
14	C	18305	0.588	0.220	0.093	0.588	0.096	0.001	0.001	0.217	-0.004	-0.238	0.217	-0.112
15	B	18305	0.436	0.227	0.436	0.246	0.088	0.002	0.000	0.267	-0.197	0.267	-0.006	-0.159
16	A	18305	0.481	0.481	0.114	0.154	0.249	0.003	0.000	0.242	0.242	-0.178	-0.239	0.056
17	A	18305	0.815	0.815	0.064	0.055	0.062	0.003	0.000	0.426	0.426	-0.247	-0.222	-0.209
18	D	18305	0.709	0.110	0.109	0.069	0.709	0.003	0.000	0.456	-0.169	-0.236	-0.304	0.456
19	C	18305	0.240	0.048	0.416	0.240	0.290	0.004	0.001	0.035	-0.258	0.042	0.035	0.054
20	A	18305	0.616	0.616	0.242	0.087	0.051	0.004	0.001	0.358	0.358	-0.100	-0.260	-0.245
21	C	18305	0.468	0.093	0.355	0.468	0.080	0.003	0.001	0.287	-0.229	-0.048	0.287	-0.183
22	B	18305	0.340	0.196	0.340	0.311	0.147	0.006	0.001	0.137	-0.002	0.137	-0.033	-0.124
23	B	18305	0.330	0.212	0.330	0.245	0.206	0.006	0.000	0.097	-0.079	0.097	0.012	-0.034
24	D	18305	0.625	0.084	0.203	0.084	0.625	0.004	0.000	0.411	-0.211	-0.142	-0.288	0.411
26	D	18305	0.706	0.150	0.083	0.060	0.706	0.001	0.000	0.453	-0.190	-0.262	-0.272	0.453
27	D	18305	0.324	0.335	0.256	0.083	0.324	0.002	0.000	0.109	0.029	0.015	-0.253	0.109
28	B	18305	0.683	0.135	0.683	0.117	0.062	0.002	0.001	0.441	-0.176	0.441	-0.285	-0.209
29	A	18305	0.611	0.611	0.114	0.103	0.169	0.002	0.001	0.395	0.395	-0.246	-0.223	-0.117
30	A	18305	0.696	0.696	0.090	0.115	0.095	0.003	0.001	0.472	0.472	-0.242	-0.263	-0.209
31	C	18305	0.749	0.060	0.111	0.749	0.077	0.002	0.001	0.442	-0.251	-0.244	0.442	-0.198
32	B	18305	0.752	0.098	0.752	0.100	0.049	0.002	0.000	0.486	-0.252	0.486	-0.259	-0.257
33	B	18305	0.295	0.179	0.295	0.282	0.240	0.003	0.000	0.092	-0.073	0.092	0.036	-0.065
35	C	18305	0.726	0.064	0.145	0.726	0.059	0.006	0.000	0.374	-0.279	-0.126	0.374	-0.203

Item Description			Proportions							Point Biserials				
Seq.	Key	N	P-Value	A	B	C	D	-	*	Tot. Corr.	A	B	C	D
36	D	18305	0.412	0.162	0.109	0.310	0.412	0.007	0.000	0.322	-0.076	-0.238	-0.106	0.322
37	A	18305	0.614	0.614	0.093	0.233	0.052	0.007	0.000	0.304	0.304	-0.254	-0.039	-0.225
38	D	18305	0.678	0.097	0.108	0.110	0.678	0.007	0.001	0.457	-0.230	-0.232	-0.210	0.457
39	A	18305	0.366	0.366	0.334	0.244	0.047	0.008	0.001	0.240	0.240	0.033	-0.158	-0.257
40	D	18305	0.430	0.084	0.304	0.173	0.430	0.008	0.001	0.197	-0.222	-0.013	-0.056	0.197
41	B	18305	0.358	0.252	0.358	0.089	0.293	0.008	0.001	0.180	-0.056	0.180	-0.290	0.065
42	B	18305	0.549	0.167	0.549	0.168	0.109	0.008	0.000	0.318	-0.084	0.318	-0.215	-0.125
43	C	18305	0.393	0.291	0.190	0.393	0.117	0.009	0.000	0.177	-0.005	-0.044	0.177	-0.187

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

Appendix E: 2012 Grade 12 Fall Reading Retest
Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit MS	InFit ZSTD	OutFit MS	OutFit ZSTD
1	0.1593	0.0161	1.04	7.7	1.08	9.9
2	1.1669	0.0176	1.13	9.9	1.34	9.9
3	-2.3343	0.0246	0.96	-2.1	0.77	-9.1
4	0.3823	0.0162	1.16	9.9	1.23	9.9
5	0.4612	0.0162	1.16	9.9	1.25	9.9
6	-0.0645	0.0161	1.01	1.7	1.02	2.9
7	-0.9935	0.0176	1.02	2.0	1.01	1.0
9	-0.2491	0.0162	1.06	9.9	1.08	9.7
10	0.0134	0.0161	1.07	9.9	1.09	9.9
11	-1.8560	0.0213	1.00	-0.3	0.91	-4.4
12	0.1401	0.0161	1.03	6.2	1.06	7.6
13	0.3554	0.0161	1.06	9.9	1.12	9.9
14	0.0612	0.0161	1.12	9.9	1.15	9.9
15	0.2396	0.0161	1.05	8.5	1.08	9.9
16	0.4163	0.0162	1.12	9.9	1.18	9.9
17	-1.7310	0.0206	0.91	-7.2	0.82	-9.7
18	-1.0640	0.0179	0.91	-9.9	0.89	-9.1
19	1.7243	0.0198	1.34	9.9	2.11	9.9
20	-0.3866	0.0164	0.97	-4.6	0.97	-3.9
21	0.2143	0.0161	1.04	6.9	1.07	9.0
22	1.0914	0.0174	1.22	9.9	1.49	9.9
23	1.3462	0.0182	1.37	9.9	1.79	9.9
24	-0.5818	0.0167	0.95	-7.9	0.94	-6.8
26	-1.2946	0.0186	1.03	2.9	1.01	0.4
27	1.1464	0.0175	1.24	9.9	1.48	9.9
28	-1.2671	0.0185	1.10	9.9	1.06	4.0
29	-0.5016	0.0165	0.96	-6.1	0.95	-5.2
30	-1.3653	0.0189	1.10	9.4	1.03	1.9
31	-1.4159	0.0191	0.98	-2.2	0.94	-3.8
32	-1.6253	0.0201	1.06	5.4	0.96	-2.1
33	1.3417	0.0182	1.27	9.9	1.67	9.9
35	-1.1487	0.0181	0.98	-2.1	0.97	-1.9
36	0.0936	0.0161	1.01	2.5	1.01	1.2
37	-0.2286	0.0162	1.01	2.1	1.02	2.5
38	-1.1128	0.0180	1.01	1.4	0.97	-2.7
39	0.3438	0.0161	1.04	7.0	1.08	9.5
40	0.9730	0.0171	1.31	9.9	1.49	9.9
41	0.7411	0.0166	1.10	9.9	1.24	9.9
42	-0.4882	0.0165	1.09	9.9	1.11	9.9
43	0.7238	0.0166	1.15	9.9	1.29	9.9

Appendix F: 2012 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	18305	1.619	0.092	0.351	0.402	0.154	0.031	0.000	0.001	0.000	0.546	-0.367	-0.294	0.272	0.313
25	3	18305	1.273	0.200	0.386	0.357	0.058	0.077	0.000	0.019	0.000	0.588	-0.501	-0.083	0.386	0.239
34	3	18305	1.409	0.151	0.369	0.400	0.080	0.047	0.000	0.002	0.000	0.571	-0.467	-0.156	0.354	0.254
44	3	18305	1.672	0.111	0.252	0.492	0.145	0.051	0.000	0.002	0.000	0.546	-0.413	-0.234	0.248	0.303

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

Appendix G: 2012 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	B	18453	0.343	0.225	0.343	0.129	0.299	0.004	0.000	0.102	-0.044	0.102	-0.089	0.008
2	A	18453	0.357	0.357	0.396	0.136	0.109	0.002	0.000	0.169	0.169	-0.020	-0.206	0.001
3	D	18453	0.728	0.092	0.090	0.088	0.728	0.001	0.000	0.395	-0.179	-0.240	-0.189	0.395
4	A	18453	0.350	0.350	0.275	0.235	0.137	0.003	0.000	0.250	0.250	0.000	-0.171	-0.133
5	A	18453	0.481	0.481	0.176	0.089	0.251	0.003	0.000	0.272	0.272	-0.130	-0.218	-0.054
6	A	18453	0.811	0.811	0.073	0.076	0.038	0.002	0.000	0.381	0.381	-0.285	-0.185	-0.125
7	B	18453	0.474	0.076	0.474	0.285	0.163	0.002	0.000	0.222	-0.177	0.222	-0.073	-0.079
8	D	18453	0.364	0.078	0.131	0.424	0.364	0.002	0.000	0.216	-0.172	-0.148	-0.011	0.216
9	B	18453	0.302	0.380	0.302	0.189	0.125	0.004	0.000	0.103	0.059	0.103	-0.108	-0.097
10	C	18453	0.791	0.071	0.087	0.791	0.049	0.001	0.000	0.480	-0.253	-0.282	0.480	-0.226
11	A	18453	0.550	0.550	0.166	0.102	0.179	0.003	0.000	0.247	0.247	-0.166	-0.196	0.001
12	A	18453	0.582	0.582	0.234	0.113	0.067	0.003	0.000	0.364	0.364	-0.095	-0.241	-0.247
13	A	18453	0.512	0.512	0.216	0.215	0.054	0.004	0.000	0.221	0.221	-0.161	-0.064	-0.079
14	C	18453	0.731	0.096	0.092	0.731	0.077	0.004	0.000	0.481	-0.255	-0.245	0.481	-0.243
16	D	18453	0.517	0.241	0.123	0.117	0.517	0.001	0.000	0.434	-0.148	-0.254	-0.214	0.434
17	B	18453	0.426	0.208	0.426	0.167	0.198	0.002	0.000	0.180	-0.056	0.180	-0.184	0.009
18	B	18453	0.495	0.217	0.495	0.212	0.073	0.003	0.000	0.240	-0.024	0.240	-0.119	-0.232
19	C	18453	0.408	0.194	0.234	0.408	0.158	0.005	0.000	0.297	-0.130	-0.105	0.297	-0.136
21	C	18453	0.371	0.067	0.153	0.371	0.406	0.003	0.000	0.088	-0.229	-0.202	0.088	0.183
22	D	18453	0.182	0.507	0.131	0.176	0.182	0.004	0.000	0.115	0.224	-0.254	-0.180	0.115
23	B	18453	0.339	0.341	0.339	0.227	0.089	0.004	0.000	0.254	0.008	0.254	-0.159	-0.195
24	A	18453	0.254	0.254	0.434	0.187	0.116	0.008	0.000	0.068	0.068	0.170	-0.129	-0.197
26	C	18453	0.594	0.198	0.139	0.594	0.059	0.010	0.000	0.301	-0.092	-0.208	0.301	-0.119
27	C	18453	0.607	0.164	0.123	0.607	0.103	0.002	0.000	0.419	-0.242	-0.221	0.419	-0.135
28	B	18453	0.466	0.258	0.466	0.148	0.117	0.011	0.000	0.350	-0.047	0.350	-0.230	-0.211
29	A	18453	0.625	0.625	0.235	0.073	0.064	0.001	0.000	0.280	0.280	-0.054	-0.224	-0.217
30	C	18453	0.317	0.355	0.240	0.317	0.086	0.002	0.000	0.137	0.108	-0.098	0.137	-0.254
31	C	18453	0.765	0.064	0.114	0.765	0.055	0.002	0.000	0.400	-0.259	-0.193	0.400	-0.192
32	A	18453	0.519	0.519	0.135	0.175	0.169	0.003	0.000	0.239	0.239	-0.190	-0.101	-0.040
33	B	18453	0.492	0.151	0.492	0.252	0.098	0.007	0.000	0.338	-0.126	0.338	-0.108	-0.244
34	B	18453	0.590	0.110	0.590	0.223	0.075	0.001	0.001	0.339	-0.264	0.339	-0.091	-0.173
35	B	18453	0.670	0.057	0.670	0.162	0.107	0.004	0.000	0.266	-0.284	0.266	-0.083	-0.087
36	C	18453	0.646	0.055	0.128	0.646	0.168	0.002	0.000	0.327	-0.226	-0.135	0.327	-0.155

Item Description			Proportions							Point Biserials				
Seq.	Key	N	P-Value	A	B	C	D	-	*	Tot. Corr.	A	B	C	D
37	C	18453	0.681	0.115	0.103	0.681	0.100	0.002	0.000	0.147	-0.039	-0.162	0.147	-0.021
41	C	18453	0.347	0.176	0.378	0.347	0.096	0.003	0.000	0.130	-0.052	-0.027	0.130	-0.095
42	A	18453	0.504	0.504	0.106	0.130	0.258	0.002	0.000	0.125	0.125	-0.199	-0.117	0.091
43	B	18453	0.400	0.144	0.400	0.243	0.208	0.005	0.000	0.204	-0.109	0.204	-0.124	-0.018
44	C	18453	0.373	0.145	0.296	0.373	0.178	0.008	0.000	0.206	-0.207	-0.031	0.206	-0.032
46	D	18453	0.642	0.127	0.121	0.072	0.642	0.038	0.000	0.352	-0.026	-0.233	-0.211	0.352
47	D	18453	0.366	0.221	0.200	0.210	0.366	0.002	0.000	0.271	0.096	-0.180	-0.238	0.271
48	C	18453	0.511	0.160	0.195	0.511	0.130	0.003	0.000	0.331	-0.246	-0.132	0.331	-0.062
49	B	18453	0.459	0.174	0.459	0.153	0.211	0.003	0.000	0.267	-0.089	0.267	-0.204	-0.058
50	B	18453	0.343	0.307	0.343	0.163	0.180	0.007	0.000	0.213	0.010	0.213	-0.167	-0.113
51	D	18453	0.368	0.299	0.167	0.162	0.368	0.002	0.000	0.318	0.023	-0.250	-0.186	0.318
52	B	18453	0.432	0.177	0.432	0.325	0.062	0.004	0.000	0.175	-0.164	0.175	0.005	-0.105
53	B	18453	0.452	0.195	0.452	0.193	0.155	0.004	0.000	0.339	-0.016	0.339	-0.222	-0.201
54	B	18453	0.551	0.150	0.551	0.163	0.131	0.005	0.000	0.369	-0.269	0.369	-0.103	-0.139
55	D	18453	0.455	0.107	0.163	0.271	0.455	0.003	0.000	0.387	-0.209	-0.207	-0.111	0.387
56	D	18453	0.515	0.263	0.099	0.119	0.515	0.004	0.000	0.432	-0.165	-0.258	-0.196	0.432
57	C	18453	0.464	0.114	0.171	0.464	0.248	0.004	0.000	0.300	-0.258	-0.210	0.300	0.033

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

Appendix H: 2012 Grade 12 Fall Science Retest Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit MS	InFit ZSTD	OutFit MS	OutFit ZSTD
1	0.7558	0.0166	1.16	9.9	1.31	9.9
2	0.8051	0.0168	1.16	9.9	1.24	9.9
3	-1.2315	0.0176	0.91	-9.9	0.89	-9.5
4	0.5535	0.0162	1.01	2.0	1.05	6.3
5	0.0855	0.0157	1.03	5.8	1.05	7.5
6	-1.5659	0.0189	0.83	-9.9	0.74	-9.9
7	0.2437	0.0158	1.09	9.9	1.11	9.9
8	0.6262	0.0164	1.07	9.9	1.12	9.9
9	1.1487	0.0178	1.25	9.9	1.42	9.9
10	-1.8633	0.0204	1.01	0.6	0.82	-9.9
11	-0.0095	0.0157	1.05	9.9	1.07	9.9
12	-0.5346	0.0160	0.96	-6.7	0.96	-6.0
13	0.0551	0.0157	1.07	9.9	1.09	9.9
14	-1.2790	0.0178	0.87	-9.9	0.78	-9.9
16	-0.3888	0.0158	0.93	-9.9	0.92	-9.9
17	0.5650	0.0162	1.16	9.9	1.22	9.9
18	0.0949	0.0157	1.06	9.9	1.07	9.9
19	0.2119	0.0158	0.99	-2.1	1.00	-0.3
21	0.9254	0.0171	1.31	9.9	1.47	9.9
22	1.4445	0.0191	0.99	-0.8	1.14	8.6
23	0.3038	0.0159	0.98	-3.4	0.99	-0.8
24	1.0459	0.0175	1.09	9.9	1.29	9.9
26	-0.3606	0.0158	0.98	-3.0	0.98	-2.4
27	-0.8800	0.0166	1.00	-0.1	0.98	-1.9
28	-0.2237	0.0157	0.99	-2.4	0.98	-2.5
29	-0.7255	0.0163	1.02	3.7	1.04	4.4
30	0.8090	0.0168	1.11	9.9	1.21	9.9
31	-1.3746	0.0181	0.87	-9.9	0.84	-9.9
32	-0.0661	0.0157	1.05	9.9	1.06	9.9
33	0.0023	0.0157	0.98	-4.9	0.98	-3.8
34	-0.5083	0.0160	0.97	-5.1	0.97	-4.8
35	-0.9395	0.0167	1.03	4.0	1.05	5.7
36	-0.8034	0.0164	0.98	-3.0	0.98	-1.9
37	-0.6582	0.0162	1.05	7.8	1.07	9.2
41	0.6906	0.0165	1.13	9.9	1.22	9.9
42	0.1765	0.0158	1.17	9.9	1.21	9.9
43	0.4428	0.0160	1.08	9.9	1.11	9.9
44	0.5413	0.0162	1.07	9.9	1.11	9.9
46	-0.7069	0.0163	0.95	-8.6	0.94	-7.7
47	0.3471	0.0159	0.99	-2.1	0.99	-1.5
48	-0.1359	0.0157	0.98	-4.4	0.98	-3.5
49	-0.1776	0.0157	1.04	8.3	1.05	7.7
50	0.5895	0.0163	1.04	6.0	1.08	9.4
51	0.2225	0.0158	0.95	-9.1	0.95	-7.2
52	0.2722	0.0158	1.10	9.9	1.13	9.9
53	-0.2063	0.0157	1.00	-0.3	1.00	-0.8
54	-0.7040	0.0162	1.05	7.5	1.05	6.2
55	-0.3866	0.0158	1.01	2.0	1.00	0.4
56	-0.5246	0.0160	0.97	-5.5	0.96	-5.4
57	-0.0743	0.0157	1.00	1.0	1.01	1.5

Appendix I: 2012 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	18453	0.961	0.273	0.493	0.234			0.101	0.000	0.004	0.000	0.399	-0.384	0.117	0.266		
20	4	18453	1.194	0.326	0.324	0.217	0.099	0.035	0.086	0.000	0.003	0.000	0.512	-0.467	0.031	0.234	0.249	0.182
25	4	18453	0.830	0.458	0.318	0.167	0.051	0.006	0.112	0.000	0.004	0.000	0.473	-0.433	0.126	0.270	0.219	0.104
38	2	18453	0.543	0.498	0.461	0.041			0.115	0.000	0.004	0.000	0.469	-0.465	0.394	0.184		
39	2	18453	0.651	0.495	0.360	0.146			0.090	0.000	0.002	0.000	0.490	-0.487	0.279	0.310		
40	2	18453	0.460	0.647	0.247	0.106			0.106	0.000	0.004	0.000	0.402	-0.402	0.258	0.262		
45	4	18453	0.565	0.601	0.270	0.098	0.027	0.004	0.118	0.000	0.004	0.000	0.462	-0.456	0.253	0.256	0.176	0.091
58	2	18453	0.445	0.644	0.267	0.089			0.128	0.000	0.004	0.000	0.418	-0.426	0.306	0.241		
59	2	18453	0.733	0.444	0.378	0.178			0.145	0.000	0.003	0.000	0.438	-0.429	0.209	0.292		

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

Appendix J: 2012 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	D	6173	0.181	0.395	0.179	0.240	0.181	0.005	0.000	0.171	-0.010	-0.096	-0.058	0.171
2	B	6173	0.617	0.166	0.617	0.171	0.046	0.001	0.000	0.320	-0.224	0.320	-0.124	-0.125
3	D	6173	0.437	0.277	0.092	0.192	0.437	0.002	0.000	0.316	-0.081	-0.219	-0.145	0.316
4	C	6173	0.784	0.071	0.044	0.784	0.099	0.002	0.000	0.364	-0.200	-0.201	0.364	-0.190
5	A	6173	0.320	0.320	0.395	0.155	0.128	0.002	0.000	0.169	0.169	0.012	-0.198	-0.036
6	B	6173	0.634	0.087	0.634	0.143	0.135	0.001	0.000	0.393	-0.160	0.393	-0.231	-0.184
7	D	6173	0.428	0.114	0.181	0.273	0.428	0.004	0.000	0.356	-0.234	-0.117	-0.123	0.356
8	A	6173	0.698	0.698	0.075	0.140	0.084	0.003	0.000	0.411	0.411	-0.239	-0.188	-0.214
9	A	6173	0.719	0.719	0.108	0.083	0.088	0.002	0.000	0.416	0.416	-0.250	-0.249	-0.138
10	D	6173	0.502	0.146	0.111	0.238	0.502	0.003	0.000	0.334	-0.085	-0.172	-0.193	0.334
11	B	6173	0.536	0.244	0.536	0.108	0.106	0.005	0.000	0.357	-0.149	0.357	-0.227	-0.135
12	D	6173	0.521	0.126	0.167	0.181	0.521	0.005	0.000	0.370	-0.208	-0.182	-0.121	0.370

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

Appendix K: 2012 Grade 12 Fall Writing Retest
Multiple-Choice Rasch Item Statistics

Seq.	Anchored Measure	Measure SE	InFit MS	InFit ZSTD	OutFit MS	OutFit ZSTD
1	3.7495	0.0696	4.82	9.9	9.90	9.9
2	-0.0203	0.0341	2.20	9.9	9.65	9.9
3	1.3398	0.0406	3.03	9.9	9.90	9.9
4	-2.0199	0.0358	1.56	9.9	2.30	4.1
5	2.7799	0.0554	5.11	9.9	9.90	9.9
6	-0.3482	0.0336	1.93	9.9	4.59	9.9
7	0.8831	0.0376	2.31	9.9	9.38	9.9
8	-1.0374	0.0337	1.68	9.9	3.13	9.5
9	-1.0934	0.0338	1.68	9.9	2.90	8.5
10	0.8863	0.0376	2.71	9.9	9.90	9.9
11	0.2508	0.0348	2.14	9.9	7.05	9.9
12	1.0398	0.0385	2.94	9.9	9.90	9.9

Appendix L: 2012 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	6173	1.900	0.297	0.513	0.182	0.008	0.668	-0.553	0.101	0.479	0.201
13	4	6173	1.904	0.306	0.494	0.188	0.011	0.699	-0.573	0.088	0.502	0.233
14	4	6173	1.735	0.394	0.481	0.121	0.004	0.660	-0.579	0.254	0.447	0.153
14	4	6173	1.758	0.389	0.471	0.134	0.006	0.691	-0.595	0.223	0.483	0.183

Appendix M: 2012 Grade 12 Fall Writing Retest
Percentage Agreement

Prompt	Composition			Revising & Editing		
	% Exact Agreement	% Adjacent Agreement	% Exact + Adjacent	% Exact Agreement	% Adjacent Agreement	% Exact + Adjacent
1	91	9	100	89	11	100
2	87	13	100	86	14	100

Appendix N: 2012 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	-5.8990	1.8333	700	378	1	0.0	1	0.0	1
1	-4.6755	1.0139	700	209	0	0.0	1	0.0	1
2	-3.9544	0.7266	700	150	0	0.0	1	0.0	1
3	-3.5209	0.6011	700	124	0	0.0	1	0.0	1
4	-3.2051	0.5274	700	109	0	0.0	1	0.0	1
5	-2.9538	0.4777	700	99	0	0.0	1	0.0	1
6	-2.7432	0.4416	700	91	1	0.0	2	0.0	1
7	-2.5607	0.4139	700	85	4	0.0	6	0.0	1
8	-2.3987	0.3919	708	81	16	0.1	22	0.1	1
9	-2.2522	0.3739	738	77	38	0.2	60	0.3	1
10	-2.1181	0.3590	766	74	56	0.2	116	0.5	1
11	-1.9939	0.3463	792	71	132	0.6	248	1.1	1
12	-1.8777	0.3355	816	69	175	0.8	423	1.8	1
13	-1.7684	0.3261	838	67	202	0.9	625	2.7	2
14	-1.6648	0.3179	859	66	285	1.2	910	3.9	3
15	-1.5661	0.3107	880	64	388	1.7	1298	5.6	5
16	-1.4715	0.3043	899	63	451	1.9	1749	7.5	7
17	-1.3807	0.2986	918	62	474	2.0	2223	9.6	9
18	-1.2931	0.2936	936	61	560	2.4	2783	12.0	11
19	-1.2082	0.2891	954	60	583	2.5	3366	14.5	13
20	-1.1258	0.2850	971	59	601	2.6	3967	17.1	16
21	-1.0457	0.2814	987	58	627	2.7	4594	19.8	18
22	-0.9674	0.2782	1003	57	657	2.8	5251	22.6	21
23	-0.8908	0.2753	1019	57	647	2.8	5898	25.4	24
24	-0.8157	0.2727	1035	56	700	3.0	6598	28.5	27
25	-0.7419	0.2704	1050	56	739	3.2	7337	31.6	30
26	-0.6694	0.2684	1065	55	723	3.1	8060	34.8	33
27	-0.5978	0.2666	1080	55	712	3.1	8772	37.8	36
28	-0.5271	0.2651	1094	55	771	3.3	9543	41.2	39
29	-0.4573	0.2637	1109	54	771	3.3	10314	44.5	43
30	-0.3880	0.2626	1123	54	796	3.4	11110	47.9	46
31	-0.3193	0.2616	1137	54	774	3.3	11884	51.3	50
32	-0.2511	0.2608	1151	54	851	3.7	12735	54.9	53
33	-0.1833	0.2602	1165	54	797	3.4	13532	58.4	57
34	-0.1157	0.2597	1179	54	802	3.5	14334	61.8	60
35	-0.0484	0.2594	1193	54	822	3.5	15156	65.4	64
36	0.0189	0.2593	1207	54	773	3.3	15929	68.7	67
37	0.0861	0.2593	1221	54	741	3.2	16670	71.9	70
38	0.1534	0.2595	1235	54	762	3.3	17432	75.2	74
39	0.2208	0.2598	1249	54	608	2.6	18040	77.8	76
40	0.2884	0.2603	1263	54	661	2.9	18701	80.7	79
41	0.3563	0.2609	1277	54	651	2.8	19352	83.5	82
42	0.4245	0.2617	1291	54	565	2.4	19917	85.9	85
43	0.4933	0.2627	1305	54	516	2.2	20433	88.1	87
44	0.5626	0.2639	1319	54	439	1.9	20872	90.0	89
45	0.6326	0.2653	1334	55	403	1.7	21275	91.8	91
46	0.7034	0.2669	1348	55	371	1.6	21646	93.4	93
47	0.7752	0.2688	1363	55	284	1.2	21930	94.6	94
48	0.8480	0.2710	1378	56	271	1.2	22201	95.8	95
49	0.9221	0.2734	1393	56	193	0.8	22394	96.6	96
50	0.9976	0.2762	1409	57	174	0.8	22568	97.3	97

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
51	1.0747	0.2794	1425	58	148	0.6	22716	98.0	98
52	1.1538	0.2830	1441	58	117	0.5	22833	98.5	98
53	1.2350	0.2871	1458	59	86	0.4	22919	98.8	99
54	1.3188	0.2918	1475	60	86	0.4	23005	99.2	99
55	1.4055	0.2972	1493	61	36	0.2	23041	99.4	99
56	1.4957	0.3034	1512	63	32	0.1	23073	99.5	99
57	1.5898	0.3105	1531	64	29	0.1	23102	99.6	99
58	1.6888	0.3188	1552	66	18	0.1	23120	99.7	99
59	1.7934	0.3284	1573	68	12	0.1	23132	99.8	99
60	1.9049	0.3397	1596	70	10	0.0	23142	99.8	99
61	2.0248	0.3531	1621	73	9	0.0	23151	99.8	99
62	2.1550	0.3691	1648	76	9	0.0	23160	99.9	99
63	2.2983	0.3885	1678	80	9	0.0	23169	99.9	99
64	2.4583	0.4122	1711	85	8	0.0	23177	100.0	99
65	2.6403	0.4419	1748	91	4	0.0	23181	100.0	99
66	2.8521	0.4798	1792	99	2	0.0	23183	100.0	99
67	3.1058	0.5294	1844	109	2	0.0	23185	100.0	99
68	3.4212	0.5966	1909	123	0	0.0	23185	100.0	99
69	3.8329	0.6913	1994	143	1	0.0	23186	100.0	99
70	4.4077	0.8340	2113	172	0	0.0	23186	100.0	100
71	5.3215	1.1148	2302	230	0	0.0	23186	100.0	100
72	6.7020	1.8952	2587	391	0	0.0	23186	100.0	100

Appendix O: 2012 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	-5.7809	1.8364	700	451	1	0.0	1	0.0	1
1	-4.5494	1.0195	700	250	0	0.0	1	0.0	1
2	-3.8168	0.7345	700	180	0	0.0	1	0.0	1
3	-3.3717	0.6108	700	150	3	0.0	4	0.0	1
4	-3.0442	0.5385	700	132	8	0.0	12	0.1	1
5	-2.7810	0.4901	700	120	12	0.1	24	0.1	1
6	-2.5583	0.4553	700	112	37	0.2	61	0.3	1
7	-2.3633	0.4288	700	105	67	0.4	128	0.7	1
8	-2.1886	0.4080	700	100	89	0.5	217	1.2	1
9	-2.0290	0.3913	700	96	145	0.8	362	2.0	2
10	-1.8814	0.3775	700	93	188	1.0	550	3.0	2
11	-1.7433	0.3661	700	90	233	1.3	783	4.3	4
12	-1.6128	0.3566	719	88	317	1.7	1100	6.0	5
13	-1.4886	0.3485	750	86	298	1.6	1398	7.6	7
14	-1.3696	0.3416	779	84	333	1.8	1731	9.5	9
15	-1.2549	0.3358	807	82	312	1.7	2043	11.2	10
16	-1.1438	0.3308	834	81	362	2.0	2405	13.1	12
17	-1.0358	0.3266	861	80	341	1.9	2746	15.0	14
18	-0.9303	0.3231	887	79	399	2.2	3145	17.2	16
19	-0.8269	0.3202	912	79	417	2.3	3562	19.5	18
20	-0.7252	0.3178	937	78	436	2.4	3998	21.8	21
21	-0.6248	0.3159	962	78	469	2.6	4467	24.4	23
22	-0.5254	0.3145	986	77	514	2.8	4981	27.2	26
23	-0.4269	0.3135	1010	77	528	2.9	5509	30.1	29
24	-0.3289	0.3129	1034	77	584	3.2	6093	33.3	32
25	-0.2311	0.3127	1058	77	630	3.4	6723	36.7	35
26	-0.1333	0.3129	1082	77	630	3.4	7353	40.2	38
27	-0.0352	0.3134	1107	77	755	4.1	8108	44.3	42
28	0.0633	0.3144	1131	77	735	4.0	8843	48.3	46
29	0.1625	0.3158	1155	78	813	4.4	9656	52.8	51
30	0.2628	0.3175	1180	78	864	4.7	10520	57.5	55
31	0.3643	0.3198	1205	78	868	4.7	11388	62.2	60
32	0.4673	0.3224	1230	79	852	4.7	12240	66.9	65
33	0.5723	0.3256	1256	80	917	5.0	13157	71.9	69
34	0.6795	0.3293	1282	81	876	4.8	14033	76.7	74
35	0.7893	0.3336	1309	82	810	4.4	14843	81.1	79
36	0.9022	0.3386	1337	83	766	4.2	15609	85.3	83
37	1.0188	0.3443	1365	85	659	3.6	16268	88.9	87
38	1.1395	0.3509	1395	86	537	2.9	16805	91.8	90
39	1.2653	0.3585	1426	88	432	2.4	17237	94.2	93
40	1.3969	0.3673	1458	90	336	1.8	17573	96.0	95
41	1.5355	0.3776	1492	93	258	1.4	17831	97.4	97
42	1.6825	0.3896	1528	96	150	0.8	17981	98.2	98
43	1.8398	0.4040	1567	99	132	0.7	18113	99.0	99
44	2.0099	0.4213	1609	103	80	0.4	18193	99.4	99
45	2.1961	0.4425	1654	109	48	0.3	18241	99.7	99
46	2.4036	0.4694	1705	115	24	0.1	18265	99.8	99
47	2.6400	0.5045	1763	124	13	0.1	18278	99.9	99
48	2.9181	0.5527	1831	136	11	0.1	18289	99.9	99
49	3.2619	0.6246	1916	153	11	0.1	18300	100.0	99
50	3.7249	0.7472	2029	183	4	0.0	18304	100.0	99

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
51	4.4773	1.0297	2214	253	1	0.0	18305	100.0	99
52	5.7238	1.8425	2520	452	0	0.0	18305	100.0	100

Appendix P: 2012 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	-5.5677	1.8335	1050	187	1	0.0	1	0.0	1
1	-4.3439	1.0140	1050	103	0	0.0	1	0.0	1
2	-3.6226	0.7267	1050	74	0	0.0	1	0.0	1
3	-3.1892	0.6010	1050	61	0	0.0	1	0.0	1
4	-2.8736	0.5271	1050	54	1	0.0	2	0.0	1
5	-2.6227	0.4772	1050	49	1	0.0	3	0.0	1
6	-2.4127	0.4409	1050	45	15	0.1	18	0.1	1
7	-2.2308	0.4130	1050	42	15	0.1	33	0.2	1
8	-2.0696	0.3908	1050	40	46	0.2	79	0.4	1
9	-1.9241	0.3726	1050	38	74	0.4	153	0.8	1
10	-1.7910	0.3574	1050	36	99	0.5	252	1.4	1
11	-1.6680	0.3445	1050	35	188	1.0	440	2.4	2
12	-1.5531	0.3335	1050	34	247	1.3	687	3.7	3
13	-1.4452	0.3239	1050	33	250	1.4	937	5.1	4
14	-1.3430	0.3154	1058	32	361	2.0	1298	7.0	6
15	-1.2460	0.3080	1068	31	376	2.0	1674	9.1	8
16	-1.1532	0.3014	1077	31	437	2.4	2111	11.4	10
17	-1.0641	0.2955	1086	30	410	2.2	2521	13.7	13
18	-0.9784	0.2902	1095	30	430	2.3	2951	16.0	15
19	-0.8956	0.2854	1104	29	449	2.4	3400	18.4	17
20	-0.8153	0.2812	1112	29	471	2.6	3871	21.0	20
21	-0.7374	0.2773	1120	28	443	2.4	4314	23.4	22
22	-0.6615	0.2738	1127	28	491	2.7	4805	26.0	25
23	-0.5874	0.2706	1135	28	475	2.6	5280	28.6	27
24	-0.5149	0.2678	1142	27	527	2.9	5807	31.5	30
25	-0.4439	0.2652	1150	27	491	2.7	6298	34.1	33
26	-0.3742	0.2629	1157	27	485	2.6	6783	36.8	35
27	-0.3057	0.2608	1164	27	479	2.6	7262	39.4	38
28	-0.2382	0.2589	1170	26	509	2.8	7771	42.1	41
29	-0.1716	0.2572	1177	26	549	3.0	8320	45.1	44
30	-0.1058	0.2557	1184	26	554	3.0	8874	48.1	47
31	-0.0408	0.2544	1191	26	531	2.9	9405	51.0	50
32	0.0237	0.2533	1197	26	550	3.0	9955	53.9	52
33	0.0877	0.2524	1204	26	559	3.0	10514	57.0	55
34	0.1511	0.2516	1210	26	535	2.9	11049	59.9	58
35	0.2142	0.2509	1216	26	558	3.0	11607	62.9	61
36	0.2771	0.2505	1223	26	597	3.2	12204	66.1	65
37	0.3397	0.2501	1229	25	563	3.1	12767	69.2	68
38	0.4023	0.2500	1236	25	565	3.1	13332	72.2	71
39	0.4647	0.2500	1242	25	594	3.2	13926	75.5	74
40	0.5272	0.2501	1248	25	542	2.9	14468	78.4	77
41	0.5899	0.2504	1255	25	499	2.7	14967	81.1	80
42	0.6526	0.2509	1261	26	487	2.6	15454	83.7	82
43	0.7157	0.2515	1268	26	467	2.5	15921	86.3	85
44	0.7792	0.2523	1275	26	435	2.4	16356	88.6	87
45	0.8430	0.2533	1281	26	368	2.0	16724	90.6	90
46	0.9075	0.2545	1287	26	305	1.7	17029	92.3	91
47	0.9726	0.2559	1294	26	283	1.5	17312	93.8	93
48	1.0384	0.2575	1300	26	241	1.3	17553	95.1	94
49	1.1052	0.2593	1307	26	190	1.0	17743	96.2	96
50	1.1730	0.2614	1314	27	188	1.0	17931	97.2	97

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
51	1.2420	0.2638	1321	27	129	0.7	18060	97.9	98
52	1.3123	0.2665	1328	27	102	0.6	18162	98.4	98
53	1.3841	0.2696	1336	27	88	0.5	18250	98.9	99
54	1.4577	0.2730	1343	28	44	0.2	18294	99.1	99
55	1.5334	0.2769	1351	28	54	0.3	18348	99.4	99
56	1.6112	0.2813	1359	29	35	0.2	18383	99.6	99
57	1.6917	0.2862	1367	29	23	0.1	18406	99.7	99
58	1.7752	0.2918	1375	30	17	0.1	18423	99.8	99
59	1.8622	0.2982	1384	30	10	0.1	18433	99.9	99
60	1.9532	0.3054	1394	31	10	0.1	18443	99.9	99
61	2.0490	0.3136	1403	32	2	0.0	18445	100.0	99
62	2.1503	0.3231	1414	33	5	0.0	18450	100.0	99
63	2.2582	0.3342	1425	34	0	0.0	18450	100.0	99
64	2.3742	0.3471	1436	35	1	0.0	18451	100.0	99
65	2.4999	0.3625	1449	37	1	0.0	18452	100.0	99
66	2.6379	0.3810	1463	39	0	0.0	18452	100.0	99
67	2.7916	0.4037	1479	41	1	0.0	18453	100.0	99
68	2.9659	0.4322	1497	44	0	0.0	18453	100.0	100
69	3.1684	0.4694	1517	48	0	0.0	18453	100.0	100
70	3.4121	0.5203	1542	53	0	0.0	18453	100.0	100
71	3.7208	0.5955	1574	61	0	0.0	18453	100.0	100
72	4.1481	0.7227	1617	74	0	0.0	18453	100.0	100
73	4.8642	1.0118	1690	103	0	0.0	18453	100.0	100
74	6.0852	1.8326	1814	187	0	0.0	18453	100.0	100

Appendix Q: 2012 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	-6.7127	1.8363	700	184	23	0.4	23	0.4	1
23	-5.4815	1.0195	700	102	85	1.4	108	1.7	1
24	-4.7482	0.7354	769	74	153	2.5	261	4.2	3
25	-4.3012	0.6129	814	61	253	4.1	514	8.3	6
26	-3.9706	0.5420	847	54	253	4.1	767	12.4	10
27	-3.7028	0.4954	874	50	203	3.3	970	15.7	14
28	-3.4742	0.4624	897	46	143	2.3	1113	18.0	17
29	-3.2721	0.4379	917	44	133	2.2	1246	20.2	19
30	-3.0887	0.4194	935	42	78	1.3	1324	21.4	21
31	-2.9190	0.4050	952	41	62	1.0	1386	22.5	22
32	-2.7597	0.3938	968	39	35	0.6	1421	23.0	23
33	-2.6082	0.3850	983	39	18	0.3	1439	23.3	23
34	-2.4627	0.3782	998	38	57	0.9	1496	24.2	24
35	-2.3217	0.3730	1012	37	88	1.4	1584	25.7	25
36	-2.1841	0.3691	1026	37	122	2.0	1706	27.6	27
37	-2.0490	0.3663	1039	37	166	2.7	1872	30.3	29
38	-1.9156	0.3643	1053	36	194	3.1	2066	33.5	32
39	-1.7833	0.3630	1066	36	150	2.4	2216	35.9	35
40	-1.6518	0.3623	1079	36	159	2.6	2375	38.5	37
41	-1.5207	0.3620	1092	36	131	2.1	2506	40.6	40
42	-1.3897	0.3619	1105	36	119	1.9	2625	42.5	42
43	-1.2587	0.3619	1118	36	74	1.2	2699	43.7	43
44	-1.1278	0.3618	1132	36	35	0.6	2734	44.3	44
45	-0.9969	0.3616	1145	36	37	0.6	2771	44.9	45
46	-0.8663	0.3612	1158	36	72	1.2	2843	46.1	45
47	-0.7360	0.3604	1171	36	107	1.7	2950	47.8	47
48	-0.6065	0.3594	1184	36	193	3.1	3143	50.9	49
49	-0.4778	0.3580	1197	36	254	4.1	3397	55.0	53
50	-0.3502	0.3564	1209	36	260	4.2	3657	59.2	57
51	-0.2238	0.3546	1222	35	296	4.8	3953	64.0	62
52	-0.0987	0.3528	1234	35	304	4.9	4257	69.0	66
53	0.0251	0.3509	1247	35	269	4.4	4526	73.3	71
54	0.1476	0.3492	1259	35	170	2.8	4696	76.1	75
55	0.2690	0.3477	1271	35	100	1.6	4796	77.7	77
56	0.3895	0.3466	1283	35	42	0.7	4838	78.4	78
57	0.5093	0.3458	1295	35	24	0.4	4862	78.8	79
58	0.6288	0.3456	1307	35	29	0.5	4891	79.2	79
59	0.7483	0.3459	1319	35	40	0.6	4931	79.9	80
60	0.8682	0.3468	1331	35	65	1.1	4996	80.9	80
61	0.9890	0.3485	1343	35	87	1.4	5083	82.3	82
62	1.1113	0.3510	1355	35	114	1.8	5197	84.2	83
63	1.2356	0.3544	1368	35	123	2.0	5320	86.2	85
64	1.3627	0.3589	1381	36	124	2.0	5444	88.2	87
65	1.4934	0.3645	1394	36	107	1.7	5551	89.9	89
66	1.6288	0.3715	1407	37	47	0.8	5598	90.7	90
67	1.7699	0.3801	1421	38	38	0.6	5636	91.3	91
68	1.9183	0.3907	1436	39	13	0.2	5649	91.5	91
69	2.0759	0.4035	1452	40	5	0.1	5654	91.6	92
70	2.2449	0.4191	1469	42	5	0.1	5659	91.7	92
71	2.4284	0.4382	1487	44	23	0.4	5682	92.0	92

Raw Score	Measure	Measure SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Percentile
72	2.6305	0.4617	1507	46	43	0.7	5725	92.7	92
73	2.8568	0.4905	1530	49	46	0.7	5771	93.5	93
74	3.1146	0.5260	1556	53	73	1.2	5844	94.7	94
75	3.4137	0.5686	1586	57	83	1.3	5927	96.0	95
76	3.7643	0.6156	1621	62	74	1.2	6001	97.2	97
77	4.1704	0.6558	1661	66	71	1.2	6072	98.4	98
78	4.6124	0.6674	1706	67	38	0.6	6110	99.0	99
79	5.0440	0.6410	1749	64	6	0.1	6116	99.1	99
80	5.4272	0.5957	1787	60	1	0.0	6117	99.1	99
81	5.7556	0.5511	1820	55	0	0.0	6117	99.1	99
82	6.0386	0.5142	1848	51	0	0.0	6117	99.1	99
83	6.2878	0.4855	1873	49	1	0.0	6118	99.1	99
84	6.5126	0.4637	1896	46	2	0.0	6120	99.1	99
85	6.7199	0.4476	1916	45	4	0.1	6124	99.2	99
86	6.9148	0.4360	1936	44	3	0.0	6127	99.3	99
87	7.1013	0.4284	1954	43	8	0.1	6135	99.4	99
88	7.2829	0.4243	1973	42	10	0.2	6145	99.5	99
89	7.4623	0.4235	1991	42	14	0.2	6159	99.8	99
90	7.6425	0.4259	2009	43	2	0.0	6161	99.8	99
91	7.8262	0.4319	2027	43	0	0.0	6161	99.8	99
92	8.0168	0.4419	2046	44	0	0.0	6161	99.8	99
93	8.2183	0.4567	2066	46	0	0.0	6161	99.8	99
94	8.4361	0.4779	2088	48	0	0.0	6161	99.8	99
95	8.6784	0.5081	2112	51	0	0.0	6161	99.8	99
96	8.9580	0.5523	2140	55	0	0.0	6161	99.8	99
97	9.2993	0.6209	2174	62	0	0.0	6161	99.8	99
98	9.7557	0.7414	2220	74	3	0.0	6164	99.9	99
99	10.4973	1.0234	2294	102	4	0.1	6168	99.9	99
100	11.7338	1.8383	2418	184	5	0.1	6173	100.0	99