TECHNICAL REPORT

for the
Pennsylvania
System of School Assessment

2008 Science
Grades 4, 8, and 11

Provided by
Data Recognition Corporation

January 2009

Table of Contents

Glossary of Common Terms

The following table contains some terms used in the 2008 technical reports 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 the context of scaling, a latent-trait characteristic indicating the level of an individual on a particular construct or competence in a particular area. Following Rasch literature, ability is used as a generic term for the construct that is being measured by an exam. Competence, achievement, learning and status are alternative terms that are sometimes used, but all are subject to some degree of misinterpretation.
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. Also see Differential Item Functioning.
Construct Validity	A term used to indicate that the test scores are to be interpreted as indicating the test taker's standing on the psychological construct measured by the test. A construct is a theoretical variable inferred from multiple types of evidence, which might include the interrelations of the test scores with other variables, internal test structure, observations of response processes, as well as the content of the test.
Constructed-Response Item	See Open-Ended Item.

Term	Common Definition
Content Validity	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.
Criterion-Referenced Interpretation	When a score is interpreted as a measure of a student's performance as 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 in 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. For example, 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 which results in dividing the score range into various proficiency level ranges. Methods for establishing cut scores vary.
Differential Item Functioning	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 or, in some cases, different rates of choosing various item options. Also known as DIF. Also see Bias.
Distractor	A plausible, but 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.
Error of Measurement	The amount by which the score actually received (an observed score) differs from a hypothetical true score. Also see 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).
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.
Key	The correct response for an MC item.

Term	Common Definition
Linking	A generic term referring to one of a number of processes by which scores from one or more 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. Also see Equating.
Logit	The fundamental unit of measurement in the Rasch model used to express both item difficulties and person locations. When expressing person locations, logits are invariably transformed into Scale Scores through a simple linear transformation before reporting (also see Scaled Score). When expressing item difficulties, logits are transformed P-Value (also see P-Value). The logit difficulty scale is inversely related to P-Values. A higher logit value would represent a relatively harder item, while a smaller 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, and 97} is 81. The value of a mean can be influenced by extreme values in a score distribution.
Measure	A Rasch estimate (or calibration) for a parameter, i.e., a person ability-parameter estimate, or an item difficulty-parameter estimate.
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% 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 which 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. Also see Open-Ended Item.
N-count	Sometimes designated as N or n, it is the number 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 constructed-response (CR) item—refers to 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 items (MC) 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.

Term	Common Definition
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.
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. Also see Logit.
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 points possible and multiplying the result by 100. Percent Correct scores 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.
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 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.

Term	Common Definition
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.
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.
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, 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. Also see 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.

Term	Common Definition
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, & 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.
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	Abbreviated 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.
Subscale	On score reports, a subscale 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 subscale score (sometimes called "strand" score).
Technical Advisory Committee	Or TAC—a group of individuals, most often professionals in the field of testing, that 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 (also see Content Validity and Construct Validity).

PREFACE: An Overview of Recent and Future Assessments

The period from 2003 through 2006 brought significant structural changes in the test blueprint for the Pennsylvania System of School Assessment (PSSA). These changes necessitated extensive test development and field testing activity along with phased-in implementation in the operational assessment. Included in this process was the development and implementation of assessments in additional grade levels.

For reading and mathematics, content changes for grades 5, 8, and 11 were developed in 2003, field tested in spring 2004, and implemented in spring 2005. The *2005 PSSA Technical Report for Reading and Mathematics* provides a description of test development activities, review of open-ended tasks and multiple-choice items, field testing, selection of items, statistical analysis of assessment data, reliability, validity, standard setting, and other technical characteristics of the operational 2005 PSSA. Test development for the new grade levels of 4, 6, and 7 began in 2004, with field testing in 2005, and full implementation in 2006. Similarly, the *2006 PSSA Technical Report for Reading and Mathematics: Grades 4, 6, and 7* provides a complete description of test development activities, item review, field testing, statistical analysis, item selection, and technical characteristics of the operational 2006 PSSA for these grade levels. In 2007 the grade 3 reading and mathematics assessment became DRC's responsibility and is covered in the present technical report, along with grades 4 through 8, and 11.

Changes in the writing assessment were designed to sharpen the focus on what is assessed with respect to Academic Standards 1.4 and 1.5. To support this effort, a shift in grade levels assessed was made, moving from grades 6 and 9 to grades 5 and 8, thereby aligning assessment to the end of elementary and middle school years. The writing testing window was changed from fall to February for grades 5 and 8, making it consistent with grade 11. Mode-specific scoring guidelines replaced domain scoring, and the introduction of stimulus-based passages and associated multiple-choice items measuring revising and editing expanded the basis of the conventions score. An account of the development of writing prompts and stimulus-based, multiple-choice items, review processes, field testing and item analysis, standard setting, and other technical characteristics of the operational 2006 PSSA may be found in the *2006 PSSA Technical Report for Writing*.

The introduction of an operational science assessment in 2008 moved closer to reality with a major standalone field test at grades 4, 8, and 11 in April–May of 2007. A description of the development of science scenarios and related multiple-choice, short answer open-ended, and extended open-ended questions, item review processes, statistical analysis of field test data, and selection of items for the 2008 operational science test may be found in the *2008 PSSA Preliminary Technical Report for Science*.

To assist the reader in navigating through the year-to-year changes in all aspects of the PSSA, tables are presented along with explanatory text. Provided is an overview of the subject areas assessed, time of year the testing activity took place, and the type of testing that occurred (e.g., operational, field testing, grade 12 retest).

ASSESSMENT ACTIVITIES OCCURRING IN THE 2003–04 SCHOOL YEAR

Table P–1 outlines the operational assessments and field tests administered during the 2003–04 school year. (A spring operational assessment in mathematics and reading took place at grades 3, 5, 8, and 11.)

As a result of new Assessment Anchor Content Standards (Assessment Anchors) developed by the Pennsylvania Department of Education (PDE) during 2003, new test items were developed (see Chapter Two of the *2005 PSSA Technical Report for Reading and Mathematics*). Following the spring operational assessment, a separate, "standalone" field test of new items for grades 5, 8, and 11 was conducted. Note that grade 11 students also took an operational writing assessment in February, and grade 6 and grade 9 students participated in a fall writing assessment. Lastly, grade 12 students who as 11th graders in the preceding spring failed to attain at least the Proficient level in any subject area were offered an opportunity to retest.

Table P–1. Operational Assessment and Field Testing During the 2003–04 School Year

Grade	Assessment Activity	Date
3	Operational mathematics and reading with embedded field test (conducted by CTB/McGraw-Hill)	April 2004
5	Operational mathematics and reading	April 2004
	Standalone field test in mathematics and reading	April/May 2004
6	Operational writing	October 2004
8	Operational mathematics and reading	April 2004
	Standalone field test in mathematics and reading	April/May 2004
9	Operational writing	October 2004
11	Operational mathematics and reading	April 2004
	Standalone field test in mathematics and reading	April/May 2004
	Operational writing	February 2004
12	Retest opportunity for students who as grade 11 students in the spring of 2003 failed to reach at least the Proficient level in mathematics, reading, or writing	October/ November 2004

ASSESSMENT ACTIVITIES OCCURRING IN THE 2004–05 SCHOOL YEAR

Table P–2 displays the operational assessments and field tests that took place during the 2004–05 school year. The operational assessment at grades 5, 8, and 11 used items chosen from the spring 2004 field test. This was the first operational assessment that reflected the Pennsylvania Assessment Anchors and Eligible Content. Fulfilling the No Child Left Behind Act of 2001 (NCLB) requirement that states must implement a test at grades 3 through 8, a major field test in mathematics and reading was administered at grades 4, 6, and 7. Item development for these new grade levels took place during 2004.

The grades 6 and 9 writing assessment was reassessed in favor of moving the writing assessment to grades 5 and 8. This accounts for the separate (standalone) field test at these grade levels. There was also a test administration change from October to February. The writing assessment also underwent changes to align the test to the Academic Standards for writing. New writing prompts and stimulus-based, multiple-choice items were also field tested at grade 11 as part of the operational assessment, hence the reference to an "embedded" field test. No assessment activity of any kind occurred at grade 9. As in fall 2003, the retest opportunity at grade 12 continued.

Table P–2. Operational Assessment and Field Testing During the 2004–05 School Year

Grade	Assessment Activity	Date
3	Operational mathematics and reading with embedded field test (conducted by CTB/McGraw-Hill)	April 2005
4	Standalone field test for mathematics and reading	April 2005
5	Operational mathematics and reading with embedded field test	April 2005
	Standalone field test in writing	February 2005
6	Standalone field test for mathematics and reading	April 2005
7	Standalone field test for mathematics and reading	April 2005
8	Operational mathematics and reading with embedded field test	April 2005
	Standalone field test in writing	February 2005
11	Operational mathematics and reading with embedded field test	April 2005
	Operational writing with embedded field test	February 2005
12	Retest opportunity for students who as grade 11 students in the spring of 2004 failed to reach at least the Proficient level in mathematics, reading, or writing	October/ November 2004

ASSESSMENT ACTIVITIES OCCURRING IN THE 2005–06 SCHOOL YEAR

Table P–3 shows the assessment activities that occurred during the 2005–06 school year. Note that the reading and mathematics operational assessments ran consecutively from grades 3 through 8 and at grade 11. For grades 4, 6, and 7, it was the first year for operational assessments. Field testing for mathematics and reading was embedded as part of the operational assessment at each grade level. At grade 3, the reference to field testing with items developed by DRC reflects the transition process of shifting the assessment from CTB/McGraw-Hill to DRC in 2007. As in previous years, the retest opportunity at grade 12 continued.

The first operational assessments for writing at grades 5 and 8 took place this year while the grade 11 writing assessment continued in the same February test window. New this year for all three grade levels, the operational writing assessments featured mode-specific scoring guidelines; stimulus-based, multiple-choice items; and a grade-specific emphasis shift in writing modes assessed. See the *2006 PSSA Technical Report for Writing: Grades 5, 8, and 11* for further information about the new writing assessments. Since extensive field testing in February 2005 produced a pool of prompts for use over several years, no additional writing prompts were field tested in 2006. However, new multiple-choice items were field tested in the 2006 writing assessment.

Table P–3. Operational Assessment and Field Testing During the 2005–06 School Year

Grade	Assessment Activity	Date
3	Operational mathematics and reading with embedded field test of DRC-written items (conducted by CTB/McGraw-Hill)	April 2006
4	Operational mathematics and reading with embedded field test	March 2006
5	Operational mathematics and reading with embedded field test	March 2006
	Operational writing with embedded field test	February 2006
6	Operational mathematics and reading with embedded field test	March 2006
7	Operational mathematics and reading with embedded field test	March 2006
8	Operational mathematics and reading with embedded field test	March 2006
	Operational writing with embedded field test	February 2006
11	Operational mathematics and reading with embedded field test	March 2006
	Operational writing with embedded field test	February 2006
12	Retest opportunity for students who as grade 11 students in the spring of 2005 failed to reach at least the Proficient level in mathematics, reading, or writing	October/ November 2005

ASSESSMENT ACTIVITIES OCCURRING IN THE 2006–07 SCHOOL YEAR

Table P–4 shows the assessment plan for the 2006–07 school year. Note that the mathematics and reading assessments ran consecutively from grades 3 through 8 and at grade 11. For grades 4, 6, and 7, it was the second year for operational assessments and the first year in which these grade levels were included in the AYP calculations. Field testing for mathematics and reading continued to be embedded as part of the operational assessments at each grade level. This was the first year in which DRC was responsible for the grade 3 assessment, as the transition from CTB/McGraw-Hill was completed. As in previous years, the retest opportunity at grade 12 continued.

The operational assessment for writing at grades 5, 8, and 11 continued in the same February test window featuring the mode-specific scoring guidelines; stimulus-based, multiple-choice items; and a grade-specific emphasis in writing modes assessed, which were introduced in 2006. Since extensive field testing in February 2005 produced a pool of prompts for use over several years, no additional writing prompts needed to be field tested in 2007. However, new multiple-choice items were field tested in the 2007 writing assessment.

Following the spring operational assessments in writing and reading and mathematics, a separate, "standalone" field test in science occurred for grades 4, 8, and 11 with full implementation scheduled for 2008.

Table P–4. Operational Assessment and Field Testing During the 2006–07 School Year

Grade	Assessment Activity	Date
3	Operational mathematics and reading with embedded field test	March 2007
4	Operational mathematics and reading with embedded field test	March 2007
	Standalone field test in science	April/May 2007
5	Operational mathematics and reading with embedded field test	March 2007
	Operational writing with embedded field test	February 2007
6	Operational mathematics and reading with embedded field test	March 2007
7	Operational mathematics and reading with embedded field test	March 2007
8	Operational mathematics and reading with embedded field test	March 2007
	Operational writing with embedded field test	February 2007
	Standalone field test in science	April/May 2007
11	Operational mathematics and reading with embedded field test	March 2007
	Operational writing with embedded field test	February 2007
	Standalone field test in science	April/May 2007
12	Retest opportunity for students who as grade 11 students in the spring of 2006 failed to reach at least the Proficient level in mathematics, reading, or writing	October/ November 2006

ASSESSMENT ACTIVITIES OCCURRING IN THE 2007–08 SCHOOL YEAR

Table P–5 shows the assessment plan for the 2007–08 school year. Note that the mathematics and reading assessments ran consecutively from grades 3 through 8, and at grade 11. For grades 4, 6, and 7, it was the third year for operational assessments and the second year in which these grade levels were included in the AYP calculations. Field testing for mathematics and reading continued to be embedded as part of the operational assessments at each grade level. This was the second year in which DRC was responsible for the grade 3 assessment. As in previous years, the retest opportunity at grade 12 continued.

The operational assessment for writing at grades 5, 8, and 11 continued in the same February test window featuring the mode-specific scoring guidelines; stimulus-based, multiple-choice items; and a grade-specific emphasis in writing modes assessed, which were introduced in 2006. Since extensive field testing in February 2005 produced a pool of prompts for use over several years, no additional writing prompts needed to be field tested in 2007. However, new multiple-choice items were field tested in the 2008 writing assessment.

Joining the spring operational assessments in writing and reading and mathematics was science at grades 4, 8, and 11.

Table P–5. Operational Assessment and Field Testing During the 2007–08 School Year

Grade	Assessment Activity	Date
3	Operational mathematics and reading with embedded field test	March/April 2008
4	Operational mathematics and reading with embedded field test	March/April 2008
	Operational science with embedded field test	April/May 2008
5	Operational mathematics and reading with embedded field test	March/April 2008
	Operational writing with embedded field test	February 2008
6	Operational mathematics and reading with embedded field test	March/April 2008
7	Operational mathematics and reading with embedded field test	March/April 2008
8	Operational mathematics and reading with embedded field test	March/April 2008
	Operational writing with embedded field test	February 2008
	Operational science with embedded field test	April/May 2008
11	Operational mathematics and reading with embedded field test	March/April 2008
	Operational writing with embedded field test	February 2008
	Operational science with embedded field test	April/May 2008
12	Retest opportunity for students who as grade 11 students in the spring of 2007 failed to reach at least the Proficient level in mathematics, reading, or writing	October/ November 2007

ASSESSMENT ACTIVITIES PLANNED FOR THE 2008–09 SCHOOL YEAR

Table P–6 shows the assessment plan for the 2008–09 school year. The mathematics and reading assessments will continue to be operational for grades 3 through 8 and grade 11. Field testing for mathematics and reading will continue to be embedded as part of the operational assessments at each grade level. As in previous years, the retest opportunity at grade 12 will continue.

The operational assessment for writing at grades 5, 8, and 11 will continue with a February test window featuring mode-specific scoring guidelines; stimulus-based, multiple-choice items; and a grade-specific emphasis in writing modes assessed. An embedded field test of writing prompts will be incorporated in the 2009 assessment along with the general procedure of including a set of embedded field test of multiple-choice items.

The second operational assessment in science is scheduled to take place in April/May. Similar to the other operational assessments, field testing for science will be embedded as part of the operational assessments at each grade level.

Table P–6. Operational Assessment and Field Testing During the 2008–09 School Year (Planned)

Grade	Assessment Activity	Date
3	Operational mathematics and reading with embedded field test	March 2009
4	Operational mathematics and reading with embedded field test	March 2009
	Operational science with embedded field test	April/May 2009
5	Operational mathematics and reading with embedded field test	March 2009
	Operational writing with embedded field test	February 2009
6	Operational mathematics and reading with embedded field test	March 2009
7	Operational mathematics and reading with embedded field test	March 2009
8	Operational mathematics and reading with embedded field test	March 2009
	Operational writing with embedded field test	February 2009
	Operational science with embedded field test	April/May 2009
11	Operational mathematics and reading with embedded field test	March 2009
	Operational writing with embedded field test	February 2009
	Operational science with embedded field test	April/May 2009
12	Retest opportunity for students who as grade 11 students in the spring of 2008 failed to reach at least the Proficient level in mathematics, reading, or writing	October/ November 2008

Chapter One: Background of Pennsylvania System of School Assessment (PSSA)

This brief overview of assessment in Pennsylvania describes the original and subsequent legislative mandates, previous assessment programs, the history of the current program's development process, the program's intent and purpose, recent changes to the program, and the student population that participates in the assessments.

THE ORIGIN OF STATE ASSESSMENT IN PENNSYLVANIA

State assessment of student achievement came about as a result of legislation enacted in 1963. Generally known as the School District Reorganization Act (Act 299), the issue of whether large or small district size provided a better quality education led to the development of Section 299.1 of Act 299, which required the State Board of Education to

. . . develop or cause to be developed an evaluation procedure designed to measure objectively the adequacy and efficiency of the educational program offered by the public schools of the Commonwealth . . . The evaluation procedure shall be so constructed and developed as to provide each school district with relevant comparative data to enable directors and administrators to more readily appraise the educational performance and to effectuate without delay the strengthening of the district's educational program. Tests developed . . . shall be used for the purpose of providing a uniform evaluation of each school district . . .

In response to the legislative mandate, the State Board of Education contracted with Educational Testing Service of Princeton, New Jersey, to engage in a two-year process of surveying and interviewing stakeholders in business, industry, education, and the general public as to what constituted a quality education. This led to the State Board adoption of *The Goals of Quality Education* in 1965. In 1967, the Department of Education formed an organizational unit along with staff to begin developing appropriate measures and engaging in extensive field testing during the 1967–68 and 1968–69 school years.

Educational Quality Assessment (EQA) Program

The first state assessment of students in Pennsylvania took place in the 1969–70 school year. Initially, state assessment was a purely school-based evaluation in the form of the *Educational Quality Assessment (EQA)* program, which reported grade 5 and 11 school-level results in ten goal areas. Grade 8 was added in 1974. Measuring both cognitive and non-cognitive areas, the program operated from 1970 through 1988. As the program evolved, a matrix sampling design was used in measuring and reporting school results in subject areas such as reading, language arts, mathematics, science, health, social studies, and analytical thinking. Initially, it operated as a voluntary program, but in 1974 it became mandatory on a cyclical basis.

Testing for Essential Learning and Literacy Skills (TELLS)

The next major revision in state assessment was the advent of the state's first mandated competency testing program, *Testing for Essential Learning and Literacy Skills (TELLS)* in the 1984–85 school year. The impetus for a statewide essential skills test evolved from an October 1983 document entitled *Turning the Tide: An Agenda for Excellence in Pennsylvania Public Schools*. A two-pronged approach was advocated, calling for:

1. competency testing in grades 3, 5, and 8 as an "early warning system" to identify students with reading and mathematics difficulties and

2. state-funded remedial instruction to provide needed additional help.

In response to this and other recommendations, the State Board of Education added *Chapter 3: Student Testing* to its regulations on June 14, 1984. It required all public school students in grades 3, 5, and 8 to be given criterion-referenced tests in reading and mathematics. The second part of the program, remedial instruction, was mandated by Act 93-1984, and required districts to provide remedial instruction programs to students identified by the tests given under the State Board regulation. Subsequently, funds were distributed to districts and intermediate units for this part of the program. The *TELLS* and *EQA* testing programs coexisted until the *EQA* was concluded in 1988. The *TELLS* program continued through the spring of 1991.

THE PENNSYLVANIA SYSTEM OF SCHOOL ASSESSMENT (PSSA)

The Pennsylvania System of School Assessment (PSSA) program was instituted in 1992. The PSSA returned to a school evaluation model with reporting at the school level only. Test administration took place in February/March, and school district participation was every third year based on the strategic planning cycle. Reading and mathematics were assessed at grades 5, 8, and 11; districts could choose to participate in the writing assessment at grades 6 and 9. State Board revisions to Chapter 5 in November 1994 brought major changes to the PSSA, beginning with the spring 1995 assessment. These changes included

1. all districts were required to participate in the reading and mathematics assessment each year,
2. student-level reports were generated in addition to school reports, and
3. the grades 6 and 9 writing assessment became mandatory on a three-year cycle corresponding to the district's strategic planning cycle.

Pennsylvania Academic Standards and the PSSA

A major structural change took place in test content with the State Board of Education's adoption of the Pennsylvania Academic Standards for Reading, Writing, Speaking and Listening, and Mathematics in January 1999 (Pennsylvania State Board of Education, 1999). The Academic Standards, which are part of *Chapter 4 Regulations on Academic Standards and Assessment*, detailed what students should know (knowledge) and be able to do (skills) at various grade levels. Subsequently, the State Board approved a set of criteria defining Advanced, Proficient, Basic, and Below Basic levels of performance. Reading and mathematics performance level results were reported at both the student and school levels for the 2000 PSSA. At that point, the PSSA became a standards-based, criterion-referenced assessment measuring student attainment of the Academic Standards while simultaneously determining the extent to which school programs enabled students to achieve proficiency of the Academic Standards.

ASSESSMENT ANCHOR CONTENT STANDARDS, CONTENT STRUCTURE, AND NEW GRADE LEVELS

Assessment in 2005 was marked by major structural changes in the PSSA. Assessment Anchor Content Standards (Assessment Anchors) developed during the previous school year to clarify content structure and improve articulation between assessment and instruction were implemented in terms of test design and reporting. At the same time field testing of mathematics and reading occurred at grades 4, 6, and 7. As specified by PL 107-110, the *No Child Left Behind Act* of 2001

(NCLB), states, school districts, and schools must achieve a minimum level of improvement each year, known as adequately yearly progress, or AYP. Accordingly, the third year of calculations for AYP were conducted and reported for grades 5, 8, and 11.

The 2006 operational reading and mathematics assessment incorporated grades 4, 6, and 7 for the first time. The assessed grade levels for 2006 included grades 3 through 8 and 11. The fourth year of calculations for AYP were conducted and reported for grades 5, 8, and 11 and for the first time in grade 3.

In 2007 the operational reading and mathematics assessment continued in grades 3 through 8 and 11. AYP calculations for grades 4, 6, and 7 took place in 2007 when they were assessed for the second time.

In 2008 the operational reading and mathematics assessment continued in grades 3 through 8 and 11, utilizing the same content structure. AYP calculations continued for all grades in 2008.

Purposes of the PSSA

As outlined in Chapter 4 of the State Board Regulations, the purposes of the statewide assessment component of the PSSA are as follows:

1. Provide students, parents, educators, and citizens with an understanding of student and school performance.
3. Determine the degree to which school programs enable students to attain proficiency of Academic Standards.
4. Provide results to school districts (including charter schools) and Area Vocational Technical Schools (AVTSs) for consideration in the development of strategic plans.
5. Provide information to state policymakers, including the State Senate, the General Assembly, and the State Board, on how effective schools are in promoting and demonstrating student proficiency of Academic Standards.
6. Provide information to the general public on school performance.
7. Provide results to school districts (including charter schools and AVTSs) based upon the aggregate performance of all students, for students with an Individualized Education Program (IEP), and for those without an IEP.

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 reach proficiency in the Academic Standards.

THE PENNSYLVANIA WRITING ASSESSMENT

In 1990 the state initiated an on-demand writing assessment in which students wrote an essay in response to a particular topic or prompt. Offered to school districts on a voluntary basis, the writing assessment consisted of three modes of writing: narrative, informational, and persuasive. The test administration for grades 6 and 9 used a matrix sampling design; nine prompts (three per mode) were administered to students within a school, although each student responded to just one randomly distributed prompt. Scoring was based on a 6-point holistic scale. Student results were aggregated and reported at the school level only. In 1992 the writing assessment was incorporated as part of the PSSA. Beginning in 1995, districts were required to participate in the

writing assessment every third year in accordance with their strategic planning cycle. However, districts were also given the choice to participate more frequently. As a result, participation rose dramatically from the expected 167 districts (one-third) in any given year to 235 (47%) in 1995, 306 (61%) in 1996, 412 (82%) in 1997, 445 (89%) in 1998, and 449 (90%) in 1999.

With the advent of the Pennsylvania Academic Standards in 1999, major changes took place in the writing assessment, including alignment to the Academic Standards as well as changes in scoring method, prompts, testing date, and reporting. These changes, which are summarized below, were implemented in the 2000–01 school year and were followed by performance level reporting in the 2001–02 school year.

- The writing assessment became mandatory for all districts every year.
- Administration of the grades 6 and 9 writing assessment was changed from February to October.
- Scoring changed to a 4-point scale for each of five domains (focus, content, organization, style, and conventions).
- Prompts were different for grade 6 and grade 9 rather than being identical at the two grade levels.
- Within a grade level all students responded to two common prompts.
- The reporting model was greatly revised, and individual student reports were issued for the first time.
- A writing assessment for grade 11 was administered for the first time in February 2001.
- In 2002, performance levels were adopted for writing and implemented in the reporting of total writing results for the February grade 11 and fall 2002 grades 6 and 9 writing assessment.

The 2006 PSSA operational writing assessment featured additional revisions that included the following enhancements:

- A shift from grades 6 and 9 to grades 5 and 8, to provide better alignment to the end of elementary school and middle school.
- Grades 5 and 8 joined grade 11 in a February test window rather than the October window used previously for grades 6 and 9.
- Students responded to two writing prompts, which were evaluated in terms of (1) a mode-specific scoring guideline and (2) a conventions scoring guideline instead of the former domain scoring.
- Stimulus-based revising/editing multiple-choice items were incorporated to provide a more reliable and valid measure of the Conventions Academic Standard.

The 2008 PSSA operational writing assessment continued with the same structure and time of year as in 2006 and 2007.

THE PENNSYLVANIA SCIENCE ASSESSMENT

In accordance with the NCLB requirement to implement an operational science assessment in 2008, a major test development effort in science took place during 2006, followed by a large-scale, standalone field test in April/May of 2007. A full implementation of an operational science assessment at grades 4, 8, and 11 occurred in April–May 2008.

Several historical milestones were significant to the development of a science test in Pennsylvania. These include:

- Adoption of Act 16 or Pennsylvania Senate Bill 652 in 2000, which redefined the PSSA "as a test developed and implemented by the Department of Education to determine only academic achievement relating directly to objective Academic Standards in the areas of reading, mathematics, and science." (see the *Science Assessment Handbook*, PDE, November 2006).
- Pennsylvania State Board of Education adoption of *Science and Technology Standards* on July 12, 2001 and the *Environment and Ecology Standards* on January 5, 2002.

Aligned to the *Pennsylvania Science Assessment Anchor Content Standards* and Eligible Content, the science test is designed to measure and report results in four major categories:

A. The Nature of Science,

B. Biological Sciences,

C. Physical Sciences, and

D. Earth and Space Sciences.

At grade 4, test questions consist of standalone multiple-choice and 2-point short answer open-ended items. At grades 8 and 11 multiple-choice questions consist of both standalone and scenario-based types. All open-ended items at grade 8 are standalone 2-point questions, while grade 11 has stand alone 2-point questions and scenario-based 4-point questions. A science scenario consists of a description of a class project, an experiment, or other research. Scenarios typically contain text, graphs, charts and/or tables. Students use their content knowledge and science process skills to answer a set of multiple-choice items and, at grade 11 only, a 4-point extended open-ended item related to the scenario. More information may be found in the following two Pennsylvania Department of Education publications available on the PDE website: *2007-2008 PSSA Assessment Handbook* and *2008 PSSA Science Item and Scoring Sampler* (one per assessed grade level).

An extensive description of the science test development activities, field testing, and statistical analyses may be found in the *2008 PSSA Preliminary Technical Report for Science*.

Chapter Two: New Test Development for Science

ASSESSMENT ANCHOR CONTENT STANDARDS AND ELIGIBLE CONTENT

The PSSA science assessment is based on the Assessment Anchor Content Standards (Assessment Anchors) as defined by the Eligible Content. The Assessment Anchors are rooted in the Academic Standards adopted by the State Board of Education in January of 2002, and the standards—under two documents: Science and Technology Standards and Environment and Ecology Standards— cover seventeen major categories describing what students need to know. The purpose of the Assessment Anchors is to articulate essential and assessable elements, and to provide clarity for instruction and for the focus of the state assessment in grades 4, 8, and 11.

Table 2–1. The Pennsylvania Academic Standards for Science

3.1 Unifying Themes	**4.1** Watersheds and Wetlands
3.2 Inquiry and Design	**4.2** Renewable and Nonrenewable Resources
3.3 Biological Sciences	**4.3** Environmental Health
3.4 Physical Science, Chemistry, and Physics	**4.4** Agriculture and Society
3.5 Earth Sciences	**4.5** Integrated Pest Management
3.6 Technology Education	**4.6** Ecosystems and their Interactions
3.7 Technological Devises	**4.7** Threatened, Endangered and Extinct Species
3.8 Science, Technology and Human Endeavors	**4.8** Humans and the Environment
	4.9 Environmental Laws and Regulations

The Assessment Anchors clarify what is expected across each grade span and focus the content of the standards into what is assessable on a large-scale test. By narrowing the Reporting Categories down to fewer groupings there are more items per Reporting Category making interpretations about what students actually know more reliable. Rather than attempting to report results for all 17 standards, the categories are organized into only four. These categories are similar to those used by the National Assessment of Educational Progress (NAEP) and The Third International Mathematics and Science Study (TIMSS). [However, the PSSA organizes the categories differently.]

Below is a table showing how the four Reporting Categories for the Assessment Anchors for the Pennsylvania System of School Assessment (PSSA) in science are related to the standards.

Table 2–2. The Relationship between the Assessment Anchor Content Standards for Science and the Academic Standards for Science

Reporting Categories		*Related Standards*																
		3.1	**3.2**	**3.3**	**3.4**	**3.5**	**3.6**	**3.7**	**3.8**	**4.1**	**4.2**	**4.3**	**4.4**	**4.5**	**4.6**	**4.7**	**4.8**	**4.9**
A.	**The Nature of Science**	✓	✓				✓	✓	✓				✓		✓	✓	✓	
B.	**Biological Sciences**	✓		✓							✓	✓			✓	✓		
C.	**Physical Sciences**		✓		✓		✓											
D.	**Earth and Space Sciences**		✓		✓	✓		✓		✓	✓						✓	

The Assessment Anchor's coding is read like an outline. The code includes the content, grade level, Reporting Category, Assessment Anchor, descriptor (Sub-Assessment Anchor), and Eligible Content. Thus S.4.A.1.3.1 would be: grade 4 science, Reporting Category A, Assessment Anchor 1, descriptor (Sub-Assessment Anchor) 3, and Eligible Content 1.

Distributed across the four Reporting Categories are a dozen Assessment Anchors. Each of the 12 Assessment Anchors exists at each grade, with the descriptors (Sub-Assessment Anchors) and Eligible Content varying to reflect grade-level appropriateness. There are 21 descriptors (Sub-Assessment Anchors) at grade 4, 23 descriptors (Sub-Assessment Anchors) at grade 8, and 23 descriptors (Sub-Assessment Anchors) at grade 11, and the Assessment Anchor documents form the basis of the test design for the grades undergoing new test development.

Total science scores reported at the student level will be based on the core (common) sections. Also reported will be the student's science performance levels. School and district-level scores will be reported at the Eligible Content level under the Assessment Anchors and based on the core (common) and matrix sections, excluding the embedded field test items.

Achieve conducted a preliminary review of the anchors in 2003, and produced a follow-up report on the anchors in 2005. More information about the Assessment Anchors and the Eligible Content can be found by referencing the Pennsylvania Science Assessment Anchors located on PDE's website at; www.pde.state.pa.us/a_and_t

OVERVIEW OF THE INITIAL SPRING 2007 PSSA SCIENCE FIELD TEST

The 2007 PSSA Science Field Test employed two types of test items: multiple-choice, and open-ended. These item types assess different levels of knowledge and provide different kinds of information about science achievement. Psychometrically, multiple-choice items are very useful and efficient tools for collecting information about a student's academic achievement. Open-ended performance tasks are less efficient in the sense that they generally generate fewer scorable points in the same amount of testing time. They do, however, provide tasks that are more realistic and better sample higher-level skills. The design of the operational 2008 PSSA for science achieves a reasonable balance between the two item types. Furthermore, well-constructed scoring guides have made it possible to include open-ended tasks in large-scale assessments such as the PSSA. Trained scorers can apply the scoring guides to efficiently score large numbers of student papers in a highly reliable way.

SCIENCE SCENARIOS FOR GRADES 8 AND 11

In addition to standalone multiple-choice and open-ended items, the science assessment includes scenarios at grades 8 and 11. In consideration of the multidisciplinary and interdisciplinary nature of science, science scenarios create stronger connections between the Nature of Science and Science Content and the MC and OE items within the scenario to allow the assessment to efficiently address and utilize the connections between content domains. A science scenario contains text, graphics, charts, and/or tables, and uses these elements to describe the results of a class project, an experiment, or other similar research. Students use the information found in a science scenario as a platform from which to answer both multiple-choice and open-ended questions (grade 11 only). Scenarios and questions reach beyond simple fact recollection; they are designed to challenge students to think and to apply knowledge and skills learned in their classrooms. Scenarios are designed to reflect multi-dimensional classroom activities that incorporate higher cognitive levels of understanding. Science scenarios challenge students to

interpret stimulus content and to apply existing knowledge to new data while using science knowledge and process skills to arrive at their answers.

MULTIPLE-CHOICE ITEMS

The majority of the science items that were included on the 2007 PSSA science field tests were multiple-choice (selected-response items), either as standalone multiple-choice items or as scenario-based multiple-choice items. This item type is especially efficient for measuring a broad range of content. In the PSSA science assessment, each multiple-choice item has four response options, only one of which is correct. The student is awarded one point for choosing the correct response. Distractors typically represent incorrect concepts, incorrect logic, or incorrect application of a scientific principle.

Multiple-choice items are used to assess a variety of skill levels, from short-term recall of facts to the application of science content. PSSA items involving application emphasize the requirement to utilize science content to find an answer, rather than simply recalling information from memory.

OPEN-ENDED ITEMS

Standalone open-ended items (constructed-response tasks) require students to read a description of a scientific problem and to develop an appropriate solution. At grade 11, scenario-based open-ended items are similar, but with scenario-based open-ended items students also have to consider the stimulus material presented in the associated scenario. Open-ended items require about five minutes per task, while the 2- or 3-part scenario-based open-ended items at grade 11 require a total of about 10 minutes. At grade 11 in scenario-based open-ended items, successive components of the open-ended item are designed to measure The Nature of Science and then either Biological Science, Physical Science, or Earth and Space Science.

Open-ended tasks are especially useful for measuring students' skills in science. They offer the opportunity to present real-life situations that require students to solve problems using science abilities learned in the classroom. Students must read the task carefully, identify the necessary information, devise a method of solution and enter the solution directly in the answer document, and when required, offer an explanation. This provides insight into the students' science knowledge, abilities, and reasoning processes.

The open-ended science items are scored on a 0–2-point scale with an item-specific scoring guideline, and each task is carefully constructed with a scoring guide reflecting the task requirements. At grade 11, scenario-based open-ended items combine two 0–2-point scale items into one compound 0–4-point scale item with two of the points associated with The Nature of Science and two of the points associated with Biological Science, Physical Science, or Earth and Space Science. The general guidelines describe a hierarchy of responses, which represent the three score levels. Each item-specific scoring guideline outlines the requirements at each score point, and each item-specific scoring guideline is based on the Science Scoring Guidelines for Open-Ended Items. (See Appendix A or the grade-specific *Science Item and Scoring Sampler,* available on the PDE website.)

Chapter Three: Item Development Process

In 2003, the existing Science, Technology, Environment, and Ecology (STEE) test was put on hold and PDE began efforts to develop a new science assessment. In the winter of 2006 a series of cognitive labs or item pilots were conducted across Pennsylvania with the primary focus of ascertaining language and contextual issues within the draft open-ended test items (grade 4) and scenario-based multiple-choice items (grade 8 and 11) and scenario-based open-ended items (grade 11), the relative difficulty of the test items, the time required to complete the individual tasks, and the opportunity to know factors related to the implementation of the new science Assessment Anchors and Eligible Content by the participating schools.

Following the series of successful cognitive labs or item pilots, DRC developed another set of test items for the proposed voluntary, standalone field test. During the development phase, PDE made the determination to change the designation of the field test from a voluntary assessment to a census-based assessment. Leading up to the administration of the standalone field test both content and bias, fairness, and sensitivity review were conducted in Pennsylvania with Pennsylvania educators. Then in the spring of 2007, the initial standalone field test was administered to the census populations at grades 4, 8, and 11 followed by a rangefinding for the open-ended items. After the scoring was completed, an item review with data was conducted for the field test items administered in 2007.

Table 3–1 provides a timeline of these major activities, which are described in some detail in this chapter as well as in Chapters Four and Five.

Table 3–1. General Timeline Associated with Assessment of Science at Grades 4, 8, and 11 Leading up to the Operational 2008 PSSA

Dates for the 2007 Standalone Field Test and 2008 Operational Test	
Detailed item development plan submitted to PDE from DRC	Aug. 2005
Draft Assessment Anchors, Eligible Content, and Test Blueprints for grades 4, 8, and 11 approved by PDE	Sep. 2005
Item Writer Training and Item Development	Sep. 2005 —June 2006
Cognitive Lab or Item Pilots for scenarios and OE items	Feb/March 2006
PDE and DRC considered Cognitive Lab or Item Pilots results and feedback from Pennsylvania students and educators on the development of future items.	March — June 2006
Item and Bias, Fairness, and Sensitivity Review	July 2006
Revisions to standalone field test items were reviewed by PDE; forms were prepared for field testing	Aug. — Dec. 2006
Item Samplers released on PDE website	Sep./Oct. 2006
Final typeset field test forms were reviewed by PDE and forms were sent to printer	Dec. 2006 — Feb. 2007
Standalone Field Testing, grades 4, 8, and 11	May 2007
Rangefinding	May/June 2007
Item Data Review	Aug. 2007
Operational and Embedded Field Test Item Selection for 2008 Administration	Sep. — Oct. 2007
Achieve reviews selected Operational items	Oct. — Nov. 2007
Operational Forms Review	Nov. — Dec. 2007
Operational Testing and Embedded Field Testing	May 2008

TEST CONTENT BLUEPRINT FOR THE 2008 OPERATIONAL SCIENCE TEST

The PSSA is based on the Pennsylvania Academic Standards as defined by the Eligible Content. The PSSA science assessment for 2008 reflects the Assessment Anchor Content Standards, which were designed as a means of improving the articulation of curricular, instructional, and assessment practices. The Anchors serve to clarify the Academic Standards assessed on the PSSA and to communicate "assessment limits," or the range of knowledge and skills from which the PSSA would be designed. Relevant to item development are the refinement and clarification embodied in the Assessment Anchors (PDE, 2004).

The Assessment Anchors are rooted in the Academic Standards adopted by the State Board of Education in January of 2002, and the standards—under two documents: Science and Technology Standards and Environment and Ecology Standards— cover seventeen major categories describing what students need to know. Rather than attempting to report results for all 17 standards, the categories are organized into only four. These categories are similar to those used by the National Assessment of Educational Progress (NEAP) and The Third International Mathematics and Science Study (TIMSS). [However, the PSSA organizes the categories differently.]

Achieve conducted a preliminary review of the anchors in 2003, and produced a follow-up report on the anchors in 2005. More information about the Assessment Anchors and the Eligible Content can be found by referencing the Pennsylvania Science Assessment Anchors located on PDE's website at; http://www.pde.state.pa.us/a_and_t/cwp/

More information on the Assessment Anchors can be found in Chapter Two.

2008 OPERATIONAL LAYOUT FOR SCIENCE

The first operational administration of the PSSA science test took place April 28 through May 9, 2008. Critical to the preparation for this operational assessment, the design of the operational assessment had to be configured to meet NCLB requirements as well as other test development and psychometric requirements. The preliminary science PSSA plan was developed in 2004 through the collaborative efforts of Data Recognition Corporation (DRC) and the Pennsylvania Department of Education (PDE) based on the recommendations of the Pennsylvania Technical Advisory Committee (TAC). At grades 4 and 8, the science assessment is in one test booklet and one separate answer booklet. The test booklet contains multiple-choice items and at grade 8 contains stimulus scenario text. The answer booklet contains scannable pages for multiple-choice (MC) responses (answer grids), open-ended (OE) items with response spaces, and demographic data collection areas. At grade 11, the science assessment is in one integrated test/answer booklet with items and scenario text appearing with scannable multiple-choice answer grids, OE response space, and demographic data collection areas.

All MC items are worth 1 point. Standalone OE items receive a maximum of 2 points (scale of 0–2) and scenario-based OE items (at grade 11 only) receive a maximum of 4 points (scale of 0–4). Each test form contains common items (identical on all forms) along with matrix/embedded field test items. The common items consist of a set of "core" items taken by all students. The matrix items and the embedded field test items are unique, in most instances, to a form. That is, there are several instances in which a matrix or embedded field test OE item appears on more than one form.

At grades 4 and 8, the 2008 PSSA science assessment is comprised of 18 forms per grade. At grade 11, the 2008 PSSA science assessment is comprised of 12 forms. All of the forms at a grade contain the common items identical for all students and sets of generally unique ("matrix") items that fulfill several purposes. These purposes include:

1. Expanding the total pool of items for school-level reporting,
2. Field testing new items,
3. Using items from the previous year's assessment for the purpose of linking.

The following three tables display the 2008 operational test design for science.

Table 3–2. 2008 Science Test Plan per Operational Form for Grade 4

CORE		MATRIX			CORE & MATRIX*	EMBEDDED FIELD TEST	
No. Core Nature of Science Points per Form	No. Core Science Content (Bio, Phy, and E&S) Points per Form	No. Matrix Nature of Science Points per Form	No. Matrix Science Content (Bio, Phy, and E&S) Points per Form	Total No. of Matrix Points per Form	Total No. of Science Points per Form (Core & Matrix) (66 +4) =	Total No. of Embedded FT Points per Form	Total No. of Forms
33	33	0–2	2–4	4	70	6	18
Core items per form (56 MC/5 (2 pt.) OE) = 61 Items Total Core/Matrix item positions per form = 61 Core + 3 Matrix Embedded Field Test items are included on each form (4 MC and 1(2pt) OE). Total item positions per form = 69 item positions Item types will be distributed according to DOK requirements of Assessment Anchors and Eligible Content.							

Table 3–3. 2008 Science Test Plan per Operational Form for Grade 8

CORE		MATRIX*			CORE & MATRIX*	EMBEDDED FIELD TEST	
No. Core Nature of Science Points per Form	No. Core Science Content (Bio, Phy, and E&S) Points per Form	No. Matrix Nature of Science Points per Form	No. Matrix Science Content (Bio, Phy, and E&S) Points per Form	Total No. of Matrix Points per Form	Total No. of Science Points per Form (Core & Matrix) (66 + 4) =	Total No. of Embedded FT Points per Form	Total No. of Forms
33	33	0–2	2–4	4	70	10	18
Core items per form (56 MC / 5(2 pt.) OE) = 61 items Total Core/Matrix item positions per form (61 core + 3 matrix)							
2 Core Nature of Science and 2 Core Science Content MC items will be bundled into a science scenario set.							
Embedded Field Test items are included on each form (4 MC and 1(2pt) OE standalone items and 4 MC scenario-based items). Total (61 core + 3 matrix + 9 FT) = 73 item positions per form. Item types will be distributed according to DOK requirements of Assessment Anchors and Eligible Content.							

Table 3–4. 2008 Science Test Plan per Operational Form for Grade 11

CORE		MATRIX*			CORE & MATRIX*	EMBEDDED FIELD TEST	
No. Core Nature of Science Points per Form	No. Core Science Content (Bio, Phy, and E&S) Points per Form	No. Matrix Nature of Science Points per Form	No. Matrix Science Content (Bio, Phy, and E&S) Points per Form	Total No. of Matrix Points per Form	Total No. of Science Points per Form (Core & Matrix) (72 + 4) =	Total No. of Embedded FT Points per Form	Total No. of Forms
36	36	0–2	2–4	4	76	15	12
Core items per form (48 MC / 6(2 pt.) OE / 3 (4 pt.) OE) = 57 items Total Core/Matrix item positions per form (57 core + 4 matrix)							
2 Core Nature of Science, 2 Core Science Content MC items, and one 4pt OE item will be bundled into a science scenario set. Three Core and one Embedded Field Test Scenario per form.							
Embedded Field Test items are included on each form (5 MC and 1(2pt) OE standalone items and 4 MC and 1(4pt) OE scenario-based items). Total (57 core + 4 matrix + 11 FT) = 72 item positions per form. Matrix will consist of only MC items. Item types will be distributed according to DOK requirements of Assessment Anchors and Eligible Content.							

Since an individual student's score is based solely on the common (or core) items, the total number of operational points is 66 for grades 4 and 8 and 72 for grade 11. The total score is obtained by combining the points from the core MC and OE portions of the test as follows:

Table 3–5. 2008 Science Core Plan per Grade

Grade	Standalone MC Items	Scenario-based MC Items	Standalone OE Items	Scenario-based OE Items	Total Points
4	56	0	5 items X 2-points =10 points	0 items X 4-points =0 points	66
8	52	4	5 items X 2-points =10 points	0 items X 4-points =0 points	66
11	36	12	6 items X 2-points =12 points	3 items X 4-points =12 points	72

For more information concerning the process used to convert the operational layout into forms (form construction), see Chapter Six. For more information about operational layout across forms and across years (form equivalency) see Chapter Ten.

TEST SESSIONS AND TIMING

The test window for the 2008 operational assessment extended from April 28 through May 9, 2008, including make-ups. The science assessments consist of 2 sections at grades 4 and 8, and three sections at grade 11. Test administration recommendations call for each section to be scheduled as one assessment session, although schools are permitted to combine multiple sections in a single session. Administration guidelines stipulate that the sections be administered

in the sequence in which they are printed in the booklets. The following tables outline the assessment schedule and estimated times for each section ("MC" refers to multiple-choice and "OE" refers to open-ended items). The estimated testing times do not include time for administrative tasks that occur during the pre- and post- administration activities.

Table 3–6. Science – Administration and Testing Times

Grade Level	Test Sections											
	Suggested Times In Minutes											
	1				2				3			
	# of Items	Administration (Total)	Administrative (Pre & Post)	Student Testing	# of Items	Administration (Total)	Administrative (Pre & Post)	Student Testing	# of Items	Administration (Total)	Administrative (Pre & Post)	Student Testing
4	31 MC 4 OE	60-65	10	50-55	31 MC 3 OE	55-60	10	45-50				
8	33 MC 4 OE	65-70	10	55-60	33 MC 3 OE	60-65	10	50-55				
11	21 MC 3 OE	65-70	10	55-60	20 MC 4 OE	60-65	10	50-55	20 MC 4 OE	60-65	10	50-55

During the assessment, students may request an extended assessment period if they indicate that they have not completed the task. Such requests are granted if the assessment administrator finds the request to be educationally valid.

REPORTING CATEGORIES AND POINTS DISTRIBUTIONS

The science assessment results will be reported in four categories, coded as A through D:

- **A.** Nature of Science
- **B.** Biological Science
- **C.** Physical Science
- **D.** Earth and Space Science

The distribution of science items into these four categories is shown in the following table.

Table 3–7. Science Reporting Categories and Proposed 2008 Point Distribution

REPORTING CATEGORY	GRADE		
	4	8	11
Nature of Science	~50% 33 pts	~50% 33 pts	~50% 36 pts
Biological Science	~17% 11 pts	~17% 11 pts	~17% 12 pts
Physical Science	~17% 11 pts	~17% 11 pts	~17% 12 pts
Earth and Space Science	~17% 11 pts	~17% 11 pts	~17% 12 pts
Total	66 pts	66 pts	72 pts

The Reporting Categories are further subdivided for specificity and Eligible Content limits. Each subdivision is coded by adding an additional numeral, e.g., A.1. These subdivisions are called "Assessment Anchors", "Descriptors (Sub-Assessment Anchors)", and "Eligible Content."

ASSESSMENT ANCHOR CONTENT STANDARDS SUBSUMED WITHIN REPORTING CATEGORIES

Distributed across the four Reporting Categories are a dozen Sub-Reporting Categories. Each of the 12 Assessment Anchors exists at each grade, with the Assessment Anchors and Eligible Content varying to reflect grade-level appropriateness. The numbers of Assessment Anchors targeted by grade level are 21 at grade 4, 23 at grade 8, and 23 at grade 11.

Total science scores reported at the student level are based on the core (common) sections. School and district-level scores are reported at the Eligible Content level under the Assessment Anchors and are based on the core (common) and matrix sections, excluding the embedded field test items.

TEST DEVELOPMENT CONSIDERATIONS

Alignment to the PSSA Assessment Anchors and Eligible Content, grade-level appropriateness (interest level, etc.), Depth of Knowledge, cognitive level, item/task level of complexity, estimated difficulty level, relevancy of context, rationale for distractors, style, accuracy, and correct terminology were major considerations in the item development process. *The Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999) and the *Principles of Universal Design* (Thompson, Johnstone, & Thurlow, 2002) guided the development process. In addition, DRC's *Bias, Fairness, and Sensitivity Guidelines* were used for developing items free of bias, fairness, and sensitivity issues. All items were reviewed for fairness by bias and sensitivity committees and for content by Pennsylvania educators and field-specialists. Items were also reviewed for adherence to the Principles of Universal Design and adherence to the guidelines outlined in the Pennsylvania publication *Principles, Guidelines and Procedures for Developing Fair Assessment Systems: Pennsylvania Assessment Through Themes* (PATT) by test

development staff trained by representatives from the National Center for Educational Outcomes (NCEO). A discussion concerning bias, fairness, and sensitivity can be found in the section below.

Bias, Fairness, and Sensitivity

At every stage of the item and test development process, item development vendor (DRC) employed procedures that are designed to ensure that items and tests meet Standard 7.4 of the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999).

> *Standard 7.4: Test developers should strive to identify and eliminate language, symbols, words, phrases, and content that are generally regarded as offensive by members of racial, ethnic, gender, or other groups, except when judged to be necessary for adequate representation of the domain.*

In meeting Standard 7.4, DRC employed a series of internal quality steps. The steps included providing specific training for item and test developers, item writers, and reviewers on how to write, review, revise, and edit items for issues of bias, fairness, and sensitivity (as well as for technical quality). The training included an awareness of and sensitivity to issues of cultural diversity. In addition to providing *internal* training in reviewing items in order to eliminate potential bias, training was also provided to *external* review panels of minority experts, teachers, and other stakeholders.

DRC's guidelines for bias, fairness, and sensitivity includes instruction concerning how to eliminate language, symbols, words, phrases, and content that might be considered offensive by members of racial, ethnic, gender, or other groups. Areas of bias that are specifically targeted include, but are not limited to the following: stereotyping, gender, regional/geographic, ethnic/cultural, socioeconomic/class, religious, experiential, and biases against a particular age group (ageism) and against persons with disabilities. Topics that should be avoided are also noted, and a balance in gender and ethnic emphasis within the pool of available items was also maintained.

Universal Design

In addition, the Principles of Universal Design were incorporated throughout all phases of the science item development process to allow participation of the widest possible range of students in the PSSA. The following checklist was used by item developers as a guideline:

1. Items are written to measure what they are intended to measure.
2. Items are written to respect the diversity of the assessment population.
3. Items are written to have a clear format for text.
4. Stimuli and items have clear pictures and graphics.
5. Items are written to have concise and readable text.
6. Items are written to allow changes to format, such as Braille, without changing meaning or difficulty.

A more extensive description of the application of Universal Design principles is described in Chapter Four.

Depth of Knowledge

An important element in statewide assessment is the alignment between the overall assessment system and the state's standards. A methodology developed by Norman Webb (1999) offers a comprehensive model that can be applied to a wide variety of contexts. With regard to the alignment between standards statements and the assessment instruments, Webb's criteria include five categories, one dealing with content. Within the content category is a useful set of levels for evaluating depth of knowledge. According to Webb (1999, p.7–8) "depth-of-knowledge consistency between standards and assessments indicates alignment if what is elicited from students on the assessment is as demanding cognitively as what students are expected to know and do as stated in the standards." The four levels of cognitive complexity (depth of knowledge) are:

- Level 1: Recall
- Level 2: Skill/Concept
- Level 3: Strategic Thinking
- Level 4: Extended Thinking

Depth-of-knowledge levels were incorporated in the item writing and review process, and items were coded with respect to the level they represented.

Test Item Writers and Item Writer Training

DRC selected and trained item writers. Qualified writers were college graduates with teaching experience and a demonstrated base of knowledge in the content area. Many of the item writers were science content assessment specialists and curriculum specialists. The writers were trained individually and had previous experience in writing multiple-choice and open-ended response items. Writers who created science scenarios had prior experience writing scenarios. Prior to developing items for the PSSA, the cadre of item writers was trained with regard to:

- Pennsylvania Assessment Anchors and Eligible Content
- Webb's Four Levels of Cognitive Complexity: Recall, Basic Application of Skill/Concept, Strategic Thinking, and Extended Thinking
- General scoring guidelines for science
- Specific and General Guidelines for Item Writing
- Bias, Fairness, and Sensitivity
- Principles of Universal Design
- Item Quality Technical Style Guidelines
- Reference Information
- Sample Items

Test Item Readability

Careful attention was given to the readability of the items and scenarios (grade 8 and grade 11) to make certain that the assessment focus of the item and the scenario did not shift based on the difficulty of reading the item or scenario. The issue of readability was addressed for all items during the final editing of items and at the item content review. Vocabulary was also addressed at the Bias, Fairness, and Sensitivity Review, although the focus was on how certain words or phrases may represent a possible source of bias or issues of fairness or sensitivity.

Process of Item Construction

To ensure that the items produced were sufficient in number and adequately distributed across subcategories and levels of difficulty, item writers were informed of the required quantities of items. As items were written, an item authoring card was completed. It contained information about the item, such as grade level, content Reporting Category, Assessment Anchor, and Eligible Content. Based on the item writer's classroom teaching experience, knowledge of the content-area curriculum, and cognitive demands required by the item, estimates were recorded for level of cognitive complexity and difficulty level. Items were written to provide for a range of difficulty.

As part of the item construction process, each item was reviewed by content specialists and editors at DRC. Content specialists and editors evaluated each item to make sure that it measured the intended Eligible Content and/or Assessment Anchor. They also assessed each item to make certain that it was appropriate to the intended grade and that it provided and cued only one correct answer. In addition, the difficulty level, Depth of Knowledge, graphics, language demand, and distractors were also evaluated. Other elements considered in this process include, but were not limited to the following: Universal Design, bias, source of challenge, grammar/punctuation, and PSSA style.

A flow chart summarizing the item and test development processes used appears in Appendix B.

Cognitive Labs (Item Pilots)

In collaboration with the PDE, DRC conducted a science cognitive lab/item pilots in selected schools throughout the Commonwealth from February 27 through March 17, 2006. A sample of 507 students from urban, suburban, and rural school districts from across the Commonwealth participated in the PSSA Science Item Tryout Project. The impetus for this study was Pennsylvania's response to the mandatory science assessment component of the *No Child Left Behind* legislation to create a rigorous science test for grades 4, 8, and 11 by 2008. The primary purpose of the cognitive lab or item tryout was to pilot the use of the new science scenarios at grade 8 and grade 11, and to pilot the multiple-choice items at grade 4.

The project involved development of science scenarios, refinement of science test items, creation of survey questions, and design of interview protocols to be administered using a cognitive laboratory technique. The cognitive laboratory technique was developed in the early 1980s through an interdisciplinary effort by survey methodologists and psychologists (Willis 1999). Different models of the cognitive process have evolved over the years, but all have four major processes in common: 1) comprehension of the question; 2) retrieval of relevant information; 3) decision process; and 4) response process (Tourangearu, 1984).

In the development and execution of the cognitive laboratory project, DRC customized the techniques employed specifically to meet the goal and expectations of the PDE. The goal of the project was to gather relevant information about the thinking processes of students enrolled in science in grades 4, 8, and 11, to create a better science assessment for Pennsylvania students.

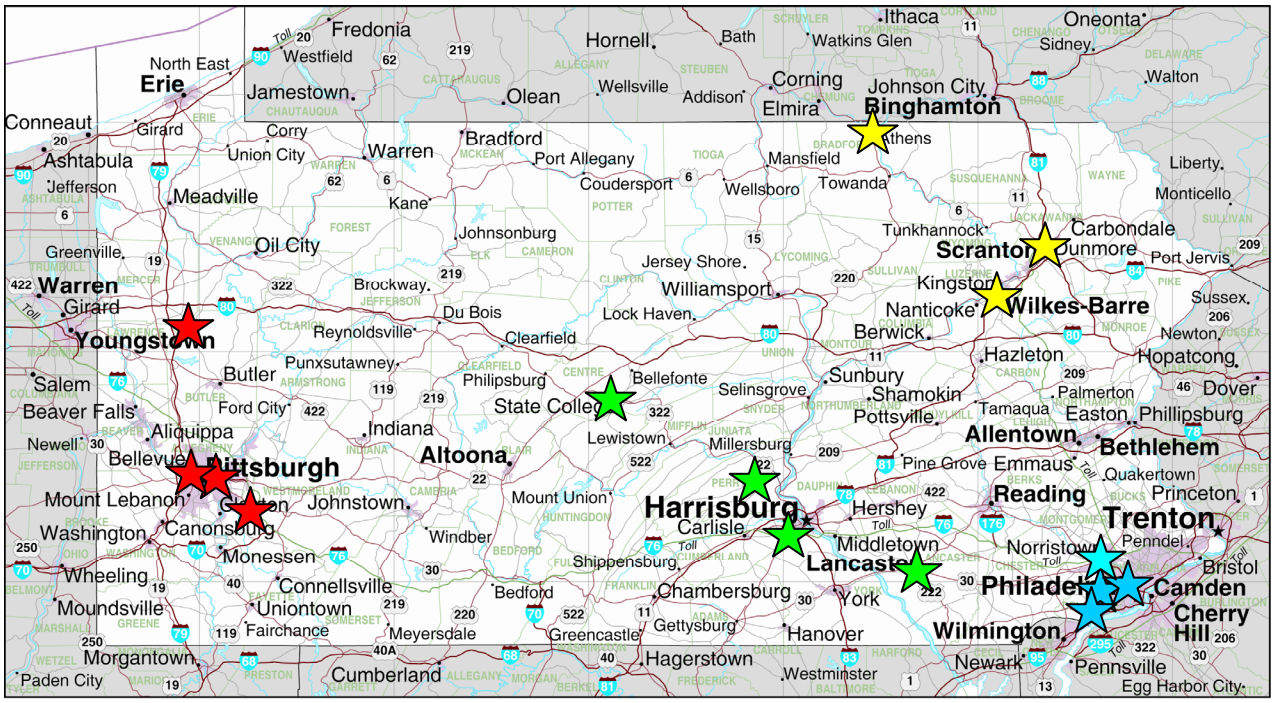
Logistics and Demographics

PDE provided DRC with a list of the Science, Technology, Environment, and Ecology Assessment Advisory Committee (STEEAAC) members who agreed to participate and to facilitate the PSSA Science Item Tryout Project in their respective districts. Disbursed throughout Pennsylvania, participating districts provided a representative sample of students enrolled in science in grades 4, 8, and 11 in urban, suburban, and rural schools. Participating schools are listed in Table 3–8 and graphically displayed in Figure 3–1.

Table 3–8. Participating Districts by Region

REGION OF COMMONWEATH	SCHOOL DISTRICT
Central	Manheim Township Newport State College Area West Shore Wilkes Barre Area
Western	Athens Area Grove City Area Penn Hills Pittsburgh Public Schools
Eastern	Haverford Township Lower Merion Mid-Valley Philadelphia City SD Upper Merion

Figure 3–1. Location of Testing Sites



Process and Procedures—The Pilot Test

Two parallel forms of the science assessment were designed for each grade level, with a designated administration time of thirty minutes. No attempt was made to replicate the design of a PSSA science operational test for the cognitive lab or pilot test, because of testing-time limitations and the objectives of this study. The items were representative of items from each of the proposed PSSA's four reporting strands (i.e., Nature of Science, Biological Science, Physical Science, and Earth and Space Science). All test items were approved by the PDE before inclusion in the PSSA Science Item Tryout Project

In fourth grade, each form of the test consisted of ten multiple-choice items, 70% of which included graphs, graphics, charts, or tables with relevant information associated with the item. All four reporting strands were assessed in each fourth grade test form. In grades 8 and 11, age/grade-appropriate science scenarios were developed. The scenarios included graphics, charts, tables, graphs, and diagrams to support the scenario text. A set of test items associated with each science scenario was developed. In grade 8, each test form included items from all four reporting strands. In grade 11, scenarios in test Form A assessed the biological, earth/space, and nature of science reporting strands, while test Form B assessed the physical, earth/space, and nature of science reporting strands.

Scenarios and questions reached beyond simple fact recollection; they were designed to challenge students to think and to apply knowledge and skills learned in their classrooms. The science scenarios were specifically designed and based on Pennsylvania Assessment Anchors and Eligible Content. Scenarios were designed to reflect multi-dimensional classroom activities that incorporate higher cognitive levels of understanding. Each scenario was stimulus-based having passages with graphics, charts, graphs, or a combination of all three media. Science scenarios challenged students to interpret passage content, while using science knowledge and process skills to arrive at their answers.

Implementation and Test Administration

Two classrooms within one geographic region participated in the project each day. At least two test development specialists were present at all but one school district during the pilot study project sessions; in addition, representatives from PDE attended most sessions. The Pennsylvania PSSA Science Item Tryout Project field work occurred during a three-week window, beginning on February 27 and concluding on March 16.

Results

Based on the information gathered at the cognitive labs or item pilots, the construction of the science scenarios was further refined. As the development process on the science scenarios continued, PDE provided additional approvals on the content and format of the science scenarios before the scenarios were reviewed by committees of Pennsylvania Educators in summer 2006.

Item Content Review in July 2006

Prior to field testing, all newly developed test items were submitted to content committees for review. The content committees consisted of Pennsylvania teachers and subject-area supervisors from school districts throughout the Commonwealth of Pennsylvania, some with postsecondary university affiliations. The primary responsibility of the content committee was to evaluate items with regard to quality and content classification, including grade-level appropriateness, estimated

difficulty, Depth of Knowledge, and source of challenge. They also suggested revisions and made recommendations for reclassification of items. In some cases when an item was deleted, the committee suggested a replacement item and/or reviewed a suggested replacement item provided by the facilitators. The committee also reviewed the items for adherence to the Principles of Universal Design, including language demand and issues of bias, fairness, and sensitivity.

The content review was held July 11–13, 2006 in Grantville, Pennsylvania. Committee members were selected and approved by PDE, and PDE-approved invitations were sent to participants by DRC. PDE also selected internal PDE staff members for attendance. The meeting commenced with a welcome by PDE and DRC. This was followed by an overview of the test development process by DRC. DRC also provided training on the procedures and forms to be used for item content review.

DRC science assessment specialists facilitated the reviews and were assisted by representatives of PDE. Committee members, grouped by grade level, worked through and reviewed the items for quality and content, as well as for the following categories:

1. Anchor Alignment (classified as Full, Partial, or No)
2. Content Limits (classified as Yes or No)
3. Grade-Level Appropriateness (classified as at grade level, below, or above grade level)
4. Difficulty Level (classified as Easy, Medium, or Hard)
5. Depth of Knowledge (classified as Recall, Application, Strategic Thinking)
6. Appropriate Source of Challenge (classified as Yes or No)
7. Correct Answer (classified as Yes or No)
8. Quality of Distractors (classified as Yes or No)
9. Graphics (classified as Yes or No)
10. Appropriate Language Demand (classified as Yes or No)
11. Freedom from Bias (classified as Yes or No)

The members then came to consensus and assigned a status to each item as a group: Approved, Accepted with Revision, Move to Another Assessment Anchor or Grade, or Rejected. All comments were recorded, and a master rating sheet was completed. Committee facilitators recorded the committee consensus on the Item Review Tally Form. The form may be found in Appendix C.

Security was addressed by adhering to a strict set of procedures. Items in binders were distributed for committee review by number and signed for by each member on a daily basis. All attendees, with the exception of PDE staff, were required to sign a Confidentiality Agreement. All materials not in use at any time were stored in a locked room. Secure materials that did not need to be retained after the meetings were deposited in secure barrels and their contents were shredded.

Bias, Fairness, and Sensitivity Reviews

Prior to field testing, all newly developed test items were also submitted to a Bias, Fairness, and Sensitivity Committee for review. This took place on July 10–14, 2006. The committee's primary responsibility was to evaluate items as to acceptability with regard to bias, fairness, and sensitivity issues. They also made recommendations for changes or deletion of items in order to remove the potential for issues of bias, fairness, and/or sensitivity. An expert, multi-ethnic committee composed of men and women was trained by DRC to review items for bias, fairness, and sensitivity issues. Table 3–9 shows the gender and race/ethnicity composition of the bias committee that reviewed the science items and scenarios for bias, fairness, and sensitivity.

Table 3–9. Demographic Composition of the 2006 Bias, Fairness, and Sensitivity Committee for Science

Committee Member	Gender	Race/Ethnicity
1	Female	American Indian or Alaskan Native
2	Female	White Non-Hispanic /Caucasian American
3	Male	Black/African American non-Hispanic
4	Female	Black/African American non-Hispanic
5	Male	White Non-Hispanic /Caucasian American
6	Male	Asian or Pacific Islander
7	Female	White Non-Hispanic /Caucasian American
8	Female	Black/African American non-Hispanic
9	Female	Latino/Hispanic American
Totals	6 Females 3 Males	3 African American 1 Asian American 3 Caucasian American 1 Hispanic American 1 Native American

Training materials included a manual developed by DRC (DRC, 2003–2007). Members of the committee also had expertise with special needs students and English Language Learners. PDE staff members were also trained and participated in the review. All items were read by a cross-section of committee members. Each member noted bias, fairness, and/or sensitivity comments on tracking sheets and on the item, if needed, for clarification. The committee then discussed each of the issues as a group and came to consensus as to which issues should represent the view of the committee. All consensus comments were then compiled, and the actions taken on these items were recorded and submitted to PDE. This review followed the same security procedures as outlined above, except that the materials were locked up and stored at the DRC offices in Harrisburg.

The results from the Bias, Fairness, and Sensitivity committee review of science are summarized in Table 3–10.

Table 3–10. Number of Items—2006 Bias, Fairness, and Sensitivity Committee Review for Science

Grade	Scenarios			Items		
	Accepted As Is	Accepted With Revision	Rejected	Accepted As Is	Accepted With Revision	Rejected
4				335	24	1
8	4	0	0	359	10	0
11	7	1	0	340	10	0
Total	11	1	0	1034	44	1

ITEM AUTHORING AND TRACKING

Initially, items were prepared on PSSA Item Cards and used for preliminary sorting and review. Although very similar, the PSSA Item Card for Multiple-Choice Items differs from the PSSA Item Card for Open-Ended Items in that the former has a location at the bottom of the card for comments regarding the distractors. Blank examples of these two cards are shown in Appendix D. In both instances a column next to the right margin provides for codes to identify the subject area, grade, content categories, passage information (in the case of reading), item type, Depth of Knowledge (cognitive complexity), estimated difficulty, answer key (MC items), and calculator use (mathematics).

All items undergoing field testing were entered into the DRC Item Viewer and Authoring Network ™ (IVAN), which is a comprehensive, secure, online item banking system. It accommodates item writing, item viewing and reviewing, and item tracking and versioning. IVAN manages the transition of an item from its developmental stage to its approval for use within a test form. The system supports an extensive item history that includes item usage within a form, item-level notes, content categories and subcategories, item statistics from both classical and Rasch item analyses, and classifications derived from analyses of differential item functioning (DIF). A sample IVAN Item Card is presented in Appendix D.

Chapter Four: Universal Design Procedures Applied in the 2008 PSSA Test Development Process

Universally designed assessments allow participation of the widest possible range of students and contribute to valid inferences about participating students. Principles of Universal Design are based on the premise that each child in school is a part of the population to be tested, and that testing results should not be affected by disability, gender, race, or English language ability (Thompson, Johnstone & Thurlow, 2002). At every stage of the item and test development process, including the 2007 field test, procedures were employed to ensure that items and subsequent tests were designed and developed using the elements of universally designed assessments developed by the National Center for Educational Outcomes (NCEO).

Federal legislation addresses the need for universally designed assessments. The *No Child Left Behind* Act (Elementary and Secondary Education Act) requires that each state must "provide for the participation in [statewide] assessments of all students" [Section 1111(b)(3)(C)(ix)(l)]. Both Title 1 and IDEA regulations call for universally designed assessments that are accessible and valid for all students, including students with disabilities and students with limited English proficiency. The benefits of universally designed assessments not only apply to these groups of students, but to all individuals with wide-ranging characteristics.

DRC's test development team was trained by the National Center for Educational Outcomes (NCEO) in the elements of Universal Design as it relates to developing large-scale statewide assessments. Pennsylvania committees involved in content review included some members who were familiar with the unique needs of students with disabilities and students with limited English proficiency. Likewise some members of the Bias, Fairness, and Sensitivity Committee were conversant with these issues. What follows are the Universal Design guidelines followed during all stages of the item development process for the 2008 PSSA.

ELEMENTS OF UNIVERSALLY DESIGNED ASSESSMENTS

After a review of research relevant to the assessment development process and the Principles of Universal Design (Center for Universal Design, 1997), NCEO has produced seven elements of Universal Design as they apply to assessments (Thompson, Johnstone & Thurlow, 2002). These elements served to guide PSSA science item development.

- **Inclusive Assessment Population**
  The PSSA target population includes all students at the assessed grades attending Commonwealth schools. For state, district, and school accountability purposes, the target population includes every student except those who will participate in accountability through an alternate assessment.
- **Precisely Defined Constructs**
  An important function of well-designed assessments is that they actually measure what they are intended to measure. The Pennsylvania Assessment Anchor Content Standards (Assessment Anchors) provided clear descriptions of the constructs to be measured by the PSSA at the assessed grade levels. Universally designed assessments must remove all non-construct-oriented cognitive, sensory, emotional, and physical barriers.
- **Accessible, Non-biased Items**
  DRC conducted both internal and external reviews of items and test specifications to ensure that they did not create barriers because of lack of sensitivity to disability, culture,

or other subgroups. Items and test specifications were developed by a team of individuals who understand the varied characteristics of items that might create difficulties for any group of students. Accessibility is incorporated as a primary dimension of test specifications, so that accessibility was woven into the fabric of the test rather than being added after the fact.

- **Amenable to Accommodations**
  Even though items on universally designed assessments are accessible for most students, there are some students who continue to need accommodations. This essential element of universally designed assessment requires that the test is compatible with accommodations and a variety of widely-used adaptive equipment and assistive technology.
  (See the section on Assessment Accommodations later in Chapter Four.)

- **Simple, Clear, and Intuitive Instructions and Procedures**
  Assessment instructions should be easy to understand, regardless of a student's experience, knowledge, language skills, or current concentration level. Knowledge questions that are posed within complex language can invalidate the test if students cannot understand how they are expected to respond to a question. To meet this guideline, directions and questions were prepared in simple, clear, and understandable language that underwent multiple reviews.

- **Maximum Readability and Comprehensibility**
  A variety of guidelines exist to ensure that text is maximally readable and comprehensible. Readability and comprehensibility are affected by many characteristics, including student background, sentence difficulty, text organization, and others. All of these features were considered as item text was developed.

  Plain language is a concept now being highlighted in research on assessments. Plain language has been defined as language that is straightforward and concise. The following strategies for editing text to produce plain language were used during the editing process of the new PSSA items.

  - Reduction of excessive length
  - Use of common words
  - Avoidance of ambiguous words
  - Avoidance of irregularly spelled words
  - Avoidance of proper names
  - Avoidance of inconsistent naming and graphic conventions
  - Avoidance of unclear signals about how to direct attention

- **Maximum Legibility**
  Legibility is the physical appearance of text, the way that the shapes of letters and numbers enable people to read text easily. Bias results when tests contain physical features that interfere with a student's focus on or understanding of the constructs that test items are intended to assess. A style guide developed and updated annually (DRC, 2004–2008) was utilized, with PDE approval, which included dimensions of style consistent with Universal Design.

GUIDELINES FOR UNIVERSALLY DESIGNED ITEMS

All test items written and reviewed adhered closely to the following guidelines for Universal Design. Item writers and reviewers used a checklist during the item development process to ensure that each aspect was attended to.

1. **Items measure what they are intended to measure.** Item writing training included assuring that writers and reviewers had a clear understanding of Pennsylvania's Assessment Anchors and Eligible Content. During all phases of test development, items were presented with content standard information to ensure that each item reflected the intended Assessment Anchor as defined by the Eligible Content. Careful consideration of the content standards was important in determining which skills involved in responding to an item were extraneous and which were relevant to what was being tested. (Necessary to the items is the requirement that students utilize reading skills.)
2. **Items respect the diversity of the assessment population.** As stated earlier, to develop items that avoid content that might unfairly advantage or disadvantage any student subgroup, item writers, test developers, and reviewers were trained to write and review items for issues of bias, fairness, and sensitivity. Training also included an awareness of, and sensitivity to, issues of cultural and regional diversity.
3. **Items have a clear format for text.** Decisions about how items are presented to students must allow for maximum readability for all students. Appropriate fonts and point sizes were employed with minimal use of italics, which is far less legible and is read considerably more slowly than standard typeface. Captions, footnotes, keys, and legends were at least a 12-point size (11pt for grade 11). Legibility was enhanced by sufficient spacing between letters, words, and lines. Blank space around paragraphs and between columns and staggered right margins were used.
4. **Stimuli and items have clear pictures and graphics.** When pictures and graphics were used, they were designed to provide essential information in a clear and uncluttered manner. Illustrations were placed directly next to the information to which they referred, and labels were used where possible. Sufficient contrast between background and text, with minimal use of shading, increased readability for students with visual difficulties. Color was not used to convey important information.
5. **Items have concise and readable text.** Linguistic demands of stimuli and items can interfere with a student's ability to demonstrate knowledge of the construct being assessed. During item writing and review, the following guidelines were used.
    - Simple, clear, commonly-used words were used whenever possible.
    - Extraneous text was omitted.
    - Vocabulary and sentence complexity were appropriate for the grade level assessed.
    - Technical terms and abbreviations were used only if related to the content being measured.
    - Definitions and examples were clear and understandable.
    - Idioms were avoided unless idiomatic speech was being assessed.
    - The questions to be answered were clearly identifiable.

6. **Items allow changes to format without changing meaning or difficulty.** In case large-print and/or Braille versions of the assessment will be produced, attention was given to using items that allow for Braille. If appropriate, specific accommodations will be permitted such as signing to a student, the use of oral presentation under specified conditions, and the use of various assistive technologies.
7. **The test has an overall appearance that is clean and organized.** Images, pictures, and text that may not be necessary (e.g., sidebars, overlays, callout boxes, visual crowding, shading) and that could be potentially distracting to students were avoided. Also avoided were purely decorative features that did not serve a purpose. Information was organized in a manner consistent with an academic English framework with a left-right, top-bottom flow.

Item Formatting

The PSSA field test booklets were formatted to maximize accessibility for all students by using text that is in a point size and font style that is easily readable. Shading, spacing, graphics, charts, and number of items were limited per page so that there is sufficient white space on each page. Whenever possible, graphics, pictures, diagrams, charts, and tables were positioned on the page with the associated test items. High contrast for text and background was used where possible to convey pertinent information. Field test forms were published on dull-finish paper to avoid the glare encountered on glossy paper. Close attention was paid to the binding of the PSSA test booklets to ensure that they would lay flat for two-page viewing and ease of reading and handling.

Consistency across PSSA assessments was maintained on the field test forms as follows:

- High contrast and clarity was used to convey detailed information.
- Typically, shading was avoided; when necessary for content purposes, 10% screens were used as the standard.
- Overlaid print on diagrams, charts, and graphs were avoided.
- Charts, graphs, diagrams, and tables were clearly labeled with titles and with short descriptions where applicable.
- Only relevant information was included in diagrams, pictures, and graphics.
- Symbols used in keys and legends were meaningful and provided reasonable representations of the topic they depicted.
- Pictures that required physical measurement were true to size.

ASSESSMENT ACCOMMODATIONS

While universally designed assessments provide for participation of the widest range of students, many students require accommodations in order to participate in the regular assessment. Clearly, the intent of providing accommodations for students is to ensure that students are not unfairly disadvantaged during testing and that the accommodations used during instruction, if appropriate, are made available as students take the test. The literature related to assessment accommodations is still evolving and often focuses on state policies regulating accommodations rather than on providing empirical data that supports the reliability and validity of the use of

accommodations. On a yearly basis, the Pennsylvania Department of Education examines accommodations policies and current research to ensure that valid, acceptable accommodations are available for students. An accommodations manual for the PSSA entitled *2008 Accommodations Guidelines for Students with IEPs and Students with 504 Plans* (PDE, January 2008) was developed for use with the 2008 PSSA.

Chapter Five: Field Test Procedures

The purpose of administering field test items is to obtain statistics for them so they can be reviewed before becoming operational. Based on this statistical review, many of the field test items in the 2007 PSSA Science Field Test were selected for use as common or matrix items in the 2008 PSSA.

EMBEDDED FIELD TEST ITEMS FOR 2008

The 2007 PSSA science forms consisted of field test items only, with no common core. Starting with the 2008 administration, like other PSSA subjects, the science assessment is built with core and embedded field test items.

The operational 2008 PSSA science forms contain common items (identical on all forms) along with matrix/embedded field test items. The common items consist of a set of "core" items taken by all students. The matrix and field test items were embedded and were unique, in most instances, to a form; however, there were instances in which a matrix or embedded field test item appeared on more than one form in order to balance the layout of forms. The development of the field test items embedded in the 2008 operational forms followed the item development procedures outlined Chapter Three and Chapter Four.

STANDALONE FIELD TEST ITEMS FOR 2007

In 2007, the PSSA science test for grades 4, 8, and 11 consisted of 12 standalone field test forms per grade. Construction of the field test forms took place from October 2006 through January 2007, following the Item Content Review consisting of Pennsylvania educators, and the Bias, Fairness, and Sensitivity review consisting of national as well as Pennsylvania review experts. All items and graphics without an "Accepted" status were revised according to committee recommendation and PDE approval. DRC designed the field test format and received PDE approval. Twelve forms, labeled 01–12, were then constructed for each grade with items distributed across the Reporting Categories and, as reasonably possible, across the Assessment Anchors.

Grade 4 Science Test Plan: Standalone Field Test

Each student taking the grade 4 standalone field test was administered 25 multiple-choice items and four open-ended items. In both sections, the multiple-choice items were administered before the open-ended items. The estimated testing time for the grade 4 standalone field test was approximately 45–50 minutes. Table 5–1 provides a summary of the spring 2007 grade 4 standalone science field tests.

Table 5–1. Grades 4 2007 Science Field Test Plan by Section

Section	No. of Standalone MC Items per Form	No. of Scenario-based MC Items per Form	No. of Standalone OE Items per Form	No. of Scenario-based OE Items per Form	Total No. of MC Items per Form	Total No. of OE Items per Form	No. of Field Test Forms
1	12	0	2	0	25	4	12
2	13	0	2	0			

Grade 8 Science Test Plan: Standalone Field Test

Each student taking the grade 8 standalone field test was administered one scenario, 24 multiple-choice items and four open-ended items. In both sections, the multiple-choice items were administered before the open-ended items. The estimated testing time for the grade 8 standalone field test was approximately 45–50 minutes. Table 5–2 provides a summary of the spring 2007 grade 8 standalone science field tests.

Table 5–2. Grades 8 2007 Science Field Test Plan by Section

Section	No. of Standalone MC Items per Form	No. of Scenario-based MC Items per Form	No. of Standalone OE Items per Form	No. of Scenario-based OE Items per Form	Total No. of MC Items per Form	Total No. of OE Items per Form	No. of Field Test Forms
1	8	4	2	0	24	4	12
2	12	0	2	0			

Grade 11 Science Test Plan: Standalone Field Test

Each student taking the grade 11 standalone field test was administered one scenario, 24 multiple-choice items and two open-ended items. In both sections, the multiple-choice items were administered before the open-ended items. The estimated testing time for the grade 11 standalone field test was approximately 45–50 minutes. Table 5–3 provides a summary of the spring 2007 grade 8 standalone science field tests.

Table 5–3. Grades 11 2007 Science Field Test Plan by Section

Section	No. of Standalone MC Items per Form	No. of Scenario-based MC Items per Form	No. of Standalone OE Items per Form	No. of Scenario-based OE Items per Form	Total No. of MC Items per Form	Total No. of OE Items per Form	No. of Field Test Forms
1	8	4	1	0	24	2	12
2	12	0	1	0			

One *Directions for Administration Manual* for each grade was written and printed for the administrators. The forms were then printed, and packets of forms were spiraled and shipped according to the sampling plan described later in this chapter. At grades 4 and 8, each student received a test booklet and one separate answer booklet. The answer booklet was used to respond to the multiple-choice and open-ended items and to collect demographic information. The multiple-choice items and scenario stimulus-text were placed within the test booklets. At grade 11, the contents of the answer booklet and the test booklet were combined into one integrated science booklet. This configuration was utilized to mirror the plans for the operational assessment.

Student Questionnaire

A 5-item student questionnaire was included on the back of each student answer booklet, and students were given approximately 5 minutes to answer the questions following the completion of the last section of the field testing event. The purpose of the questionnaire was to gain further feedback on the field test from the student's perspective. The completed student questionnaires were returned to DRC and the results were summarized.

STATISTICAL ANALYSIS OF ITEM DATA

All field tested items were analyzed statistically following conventional item analysis methods. This analysis was useful to evaluate the overall performance of the pool of field test items. For MC items, traditional or classical item statistics included the point-biserial correlation (Pt Bis) for the correct and incorrect responses, percent correct (P-Value), and the percent responding to incorrect responses (distractors). For OE items the statistical indices included the item-test correlation, the point-biserial correlation for each score level, percent in each score category or level, and the percent of non-scorable responses.

In general, more capable students are expected to respond correctly to easy items and less capable students are expected to respond incorrectly to difficult items. If either of these situations does *not* occur, the item will be reviewed by DRC test development staff and committees of Pennsylvania educators to determine the nature of the problem and the characteristics of the students affected. The primary way of detecting such conditions is through the point-biserial correlation coefficient for dichotomous (MC) items and the item-total correlation for polytomous (OE) items. In each case the statistic will be positive if the total test mean score is higher for the students who respond correctly to MC items (or attain a higher OE item score) and negative when the reverse is true.

Item statistics are used as a means of detecting items that deserve closer scrutiny, rather than being a mechanism for automatic retention or rejection. For a MC item to be flagged as needing further internal review, the criteria included any of the following:

- Point-biserial correlation for the correct response of less than 0.25
- Point-biserial correlation for any incorrect response greater than 0.0
- Percent correct less than 30% or greater than 90%
- Percent responding to any incorrect responses greater than the percent correct

For an OE item to be flagged as needing further internal review, the criteria included any of the following:

- Item-test correlation less than 0.40
- Percent in any score category less than 10% or greater than 40%
- Non-scorable responses greater than 10 percent

Item analysis results for multiple-choice and open-ended field test items are presented in Appendices F through G.

DIFFERENTIAL ITEM FUNCTIONING

Differential item functioning (DIF) occurs when two examinees with the same ability level but different group membership do not have the same probability of answering the item correctly. This pattern of results may suggest the presence of item bias. As a statistical concept, however, DIF can be differentiated from *item bias*, which is a content issue that can arise when an item presents negative group stereotypes, uses language that is more familiar to one subpopulation than to another, or is presented in a format that disadvantages certain learning styles. While the source of item bias is usually plain to trained judges, DIF may have no clear cause. However, studying how DIF arises and how it presents itself has an effect on how best to detect and correct it.

Limitations of Statistical Detection

No statistical procedure should be used as a substitute for rigorous, hands-on reviews by content and bias specialists. The statistical results can help to organize the review so the effort is concentrated on the most problematic cases; however, no items should be automatically rejected simply because a statistical method flagged them or accepted because they were not flagged.

Statistical detection of DIF is an inexact science. There have been a variety of methods proposed for detecting DIF, but no one statistic can be considered either necessary or sufficient. Different methods are more or less successful depending on the situation. No analysis can guarantee that a test is free of bias, but almost any thoughtful analysis will uncover the most flagrant problems.

A fundamental shortcoming of all of the statistical methods used in DIF evaluation is that all are intrinsic to the test being evaluated. If a test is unbiased overall but contains one or two DIF items, any method will locate the problems. If, however, all items on the test show consistent DIF to the disadvantage of a given subpopulation, a statistical analysis of the items will not be able to separate DIF effects from true differences in achievement.

Mantel-Haenszel Procedure for Differential Item Functioning

The *Mantel-Haenszel* procedure for detecting differential item functioning is a commonly used technique in educational testing. It does not depend on the application or the fit of any specific measurement model. However, it does have significant philosophical overlap with the Rasch model since it uses total score to organize the analysis.

The procedure as implemented by DRC contrasts a focal group with a reference group. While it makes no practical difference in the analysis which group is defined as the focal group, the group most apt to be disadvantaged by a biased measurement is typically defined as the focal group. In these analyses, the focal group was female for gender-based DIF and black for ethnicity-based DIF; reference groups were male and white, respectively. The Mantel-Haenszel (MH) statistic (Mantel & Haenszel, 1959) for each item is computed from a contingency table. It has two groups (focal and reference) and two outcomes (right or wrong). The ability groups are defined by the core test's score distribution for the total examinee populations.

The basic MH statistic is a single degree of freedom chi-square that compares the observed number in each cell to the expected number. The expected counts are computed to ensure that the analysis is not confounded with differences in the achievement level of the two groups.

For OE items, a comparable statistic is computed based on the standardized mean difference (SMD) (Dorans, Schmitt & Bleistein, 1992), computed as the differences in mean scores for the focal and reference groups if both groups had the same score distribution.

To assist the review committees in interpreting the analyses, the items are assigned a severity code based on the magnitude of the MH statistic. Items classified as A+ or A- have little or no statistical indication of differential item functioning. Items classified as B+ or B- have some indication of DIF and may not require revision. Items classified as C+ or C- have strong evidence of DIF and should be reviewed and revised if they are to be used again. The plus sign indicates that the item favors the focal group and a minus sign indicates that the item favors the reference group.

Counts of the number of items from each grade and content area that were assigned to each severity code are shown below in Table 5–4. DIF analyses were conducted only on field test items.

Table 5–4. 2008 DIF Summary

Multiple-Choice Male/Female DIF Counts						
2008						
Grade	**A**	**B-**	**B+**	**C-**	**C+**	**Total**
4	70	1	1	0	0	72
8	140	3	0	0	1	144
11	96	8	3	0	1	108

Multiple-Choice White/Black DIF Counts						
2008						
Grade	**A**	**B-**	**B+**	**C-**	**C+**	**Total**
4	68	3	0	1	0	72
8	138	4	0	2	0	144
11	105	3	0	0	0	108

Open-Ended Male/Female DIF Counts						
2008						
Grade	**A**	**B-**	**B+**	**C-**	**C+**	**Total**
4	17	0	1	0	0	18
8	15	0	3	0	0	18
11	20	1	3	0	0	24

Open-Ended White/Black DIF Counts						
2008						
Grade	**A**	**B-**	**B+**	**C-**	**C+**	**Total**
4	11	5	0	2	0	18
8	7	5	0	6	0	18
11	20	1	0	3	0	24

REVIEW OF ITEMS WITH DATA

In the preceding section on Statistical Analysis of Item Data, it was stated that test development content-area specialists used certain statistics from item and DIF analyses of the 2006 field test to identify items for further internal review. Specific flagging criteria for this purpose were specified in the previous section. Generally, items not identified were those that had good statistical characteristics and, consequently, were regarded as statistically acceptable. Following the internal review, a committee of Pennsylvania educators was convened to review the entire pool of field test items using the same criteria.

The review of the items with data by Pennsylvania educators was conducted by committees composed of PDE staff and grade-level classroom teachers and administrators. The review took place on August 20–21, 2007. In this session, committee members were first trained by DRC staff, with regard to a historical perspective on the science assessment and the statistical indices used in item evaluation. This was followed by a discussion with examples concerning reasons why a committee might recommend that an item might be retained or rejected regardless of the statistics. The committee review process involved a brief exploration of possible reasons for the statistical profile of an item (e.g., possible bias, grade appropriateness, instructional issues) and a decision regarding acceptance. DRC content-area test development specialists facilitated the review of the items by grade level. Each grade-level committee reviewed the pool of field test items and made recommendations on each item and scenario (grade 8 and grade 11). Further discussion on how this information was used is covered in Chapter Six.

Chapter Six: Operational Forms Construction for 2008

FINAL SELECTION OF ITEMS AND 2008 PSSA SCIENCE FORMS CONSTRUCTION

When the final selection of items for the operational 2008 science test was ready to begin, the candidate items that emerged from the 2007 standalone field test had undergone multiple reviews, including:

- Reviews by DRC science test development specialists and curriculum specialists
- Formal bias, fairness, and sensitivity review by the Bias, Fairness, and Sensitivity Committee consisting of an expert, multi-ethnic group of men and women with members also having expertise with special needs students and English Language Learners
- Formal review by the science content committees consisting of Pennsylvania educators, including teachers as well as district personnel
- PDE review
- Item data review by members of the PDE science teacher committees

The end product of the above process was an "item status" designation for each field tested item. All items having an item status code of "Acceptable/Active" were candidates to be selected for the 2008 PSSA. To have an item status code of "Acceptable/Active" meant that the item met the following criteria:

- Appropriately aligned with its designated Assessment Anchor Content Standard (Assessment Anchor) and sub-classifications
- Acceptable in terms of bias/fairness/sensitivity issues, including differential item functioning (for gender and race)
- Free of major psychometric flaws, including a special review of flagged items

Next, all relevant information regarding the acceptable items, including associated graphics, was entered into the IVAN (item banking) system. From the IVAN system and other database sources, Excel files were created at each grade. These files contained all relevant content-specific codes and statistical characteristics. The IVAN system also created for each acceptable item a card displaying the item, any associated graphic, and all relevant content-specific codes and item statistics for use by the science test development specialists and psychometric services staff.

DRC test development specialists reviewed the test design blueprint, including the number of items per strand for each content-area test.

Psychometricians provided science test development specialists with an overview of the psychometric guidelines for forms construction, including guidelines for selecting matrix items.

Senior DRC science test development specialists reviewed all items in the operational pool to make an initial selection for common (core) and matrix sections according to test blueprint requirements and psychometric guidelines. No changes were made to any item since even slight alterations could affect how an item performs on subsequent testing.

For the common items, this meant that the combination of MC and OE items would yield the appropriate range of points while tapping an appropriate variety of the Assessment Anchors and related Eligible Content within each Reporting Category. Items selected in the first round were examined with regard to how well they went together as a set. Of particular concern were the following:

- One item providing cues as to the correct answer to another item
- Context redundancy
- Presence of "clang" (distractors not unique from one another)
- Diversity of names and artwork for gender and ethnicity

The first round of items was then evaluated for statistical features such as an acceptable point bi-serial correlation and whether correct answers were distributed equally—that is, whether approximately 25 % of correct answers appeared in each of the four possible positions (A, B, C, or D). Selected items that were psychometrically problematic resulted in a search by the senior reviewer for suitable replacements. At this point, the second round of items was analyzed. If necessary, this iterative process between content-based selections and statistical properties continued in an effort to reach the best possible balance.

The process for selecting operational matrix items was a little different. The chief consideration was that items in the matrix section of the various forms, together with the common items, would yield a greater overall pool of items from which reliable sub-category results could be generated for school-level reporting. Once again the cardinal principle was the selection of an appropriate number of items to properly cover the sub-categories. The science test development specialist's task was to distribute these items in matrix sections across the 18 forms (12 forms at grade 11) so that the OE item and set of MC items assigned to a particular form would go well with one another and reflect the same content and statistical considerations as previously outlined. Additionally, the forms needed to display similar difficulty levels. Scenario-based items were not considered for use as matrix items, and, at grade 11, OE items were not considered for use as matrix items.

Once the recommendations were finalized for the common/core and matrix items, they were submitted to PDE for review. Department staff provided feedback, which could be in the form of approval or recommendations for replacing certain items. Any item replacement was accomplished by the collective effort of the test development specialists, psychometricians, and PDE staff until final PDE approval.

SPECIAL FORMS USED IN THE 2008 PSSA

Braille and Large Print

Students with visual impairments were able to respond to test materials that were available in either **Braille** or **large print**. At each grade level assessed, one form was selected for the creation of Braille and large-print editions. School district personnel ordered Braille or large-print assessment materials directly from the Pennsylvania Training and Technical Assistance Network (PaTTAN) in Harrisburg. They could also contact PaTTAN for technical assistance regarding students with visual impairments.

School personnel were directed to transcribe all student answers (MC and OE) into scannable answer documents exactly as the student responded. No alterations or corrections of student work were permitted, and the answer document had to have the identical form designation.

Instructions for the appropriate use of these special forms are detailed in the *2008 Accommodations Guidelines for Students with IEPs and Students with 504 Plans* (PDE, January 2008) available on the PDE website at www.pde.state.us.

Chapter Seven: Test Administration Procedures

TEST SESSIONS, TIMING, AND LAYOUT

The test window for the 2008 science assessment was from April 28 through May 9, 2008, including make-ups. The assessment consisted of two sections. Additional information concerning testing time and test layouts can be found in Chapter Five.

SHIPPING, PACKAGING, AND DELIVERY OF MATERIALS

There were two shipments sent out by Data Recognition Corporation (DRC). Shipment one was delivered by March 31, 2008, and contained the *Handbook for Assessment Coordinators and Administrators* and the *Directions for Administration Manuals* for each grade tested at a school. Shipment two was delivered by April 14, 2008, and contained the Spanish-translation *Directions for Administration Manuals*, administrative materials (e.g., return shipping labels, District/School Labels, Do Not Score Labels, and Student Precode Labels) and secure materials (e.g., test booklets and answer booklets). DRC ensured that all assessment materials were assembled correctly prior to shipping. DRC Operations staff used the automated Operations Materials Management System (Ops MMS) to assign secure materials to a district at the time of ship out. This system used barcode technology to provide an automated quality check between items requested for a site and items shipped to a site. A shipment box manifest was produced for and placed in each box shipped. DRC Operations staff double-checked all box contents with the box manifest prior to the box being sealed for shipment to ensure accurate delivery of materials. DRC Operations staff performed lot acceptance sampling on both shipments. Districts and schools were selected at random and examined for correct and complete packaging and labeling. This sampling represented a minimum of 10 percent of all shipping sites.

DRC's materials management system, along with the systems of shippers, allowed DRC to track the items from the point of shipment from DRC's warehouse facility to receipt at the district, school, or testing site. All DRC shipping facilities, materials processing facilities, and storage facilities are secure. Access is restricted by security code. Non-DRC personnel are escorted by a DRC employee at all times. Only DRC inventory control personnel have access to stored secure materials. DRC employees are trained in and made aware of the high level of security that is required.

The assessments for grades 4, 8, and 11 were shipped together. DRC packed 510,323 assessment booklets, approximately 59,918 manuals, and 53,889 non-secure materials for 3,002 testing sites. DRC used UPS to deliver materials to the testing sites.

MATERIALS RETURN

The materials return window was May 14–June 27, 2008. DRC used UPS for all returns.

TEST SECURITY MEASURES

Test security is essential to obtaining reliable and valid scores for accountability purposes. The 2008 PSSA included a Test Security Affidavit that was to be signed and returned by every principal or director where testing materials were shipped. DRC received 3,140 signed Test Security Affidavits of the 3,200 affidavits sent to the testing sites participating in the 2008 Science PSSA. The purpose of the affidavit was to serve as a tool to document that the

individuals responsible for administering the assessments both understood and acknowledged the importance of test security and accountability. The Test Security Affidavit attested that all security measures were followed concerning the handling of secure materials. Some of the security measures included:

- The contents of the test were not discussed, disseminated, described, or otherwise revealed to anyone.
- The contents of the test were not kept, copied, or reproduced.
- All booklets were kept in a locked, secure storage area at both the district and school levels.

SAMPLE MANUALS

Copies of the *Handbook for Assessment Coordinators and Administrators* and the *Directions for Administration Manuals* can be found on the Pennsylvania Department of Education website at www.pde.state.pa.us.

ASSESSMENT ACCOMMODATIONS

An accommodations manual entitled *2008 Accommodations Guidelines for Students with IEPs and Students with 504 Plans* (PDE, January 2008) was developed for use with the 2008 PSSA. Additional information regarding assessment accommodations can be found in Chapter Four of this report.

Chapter Eight: Processing and Scoring

RECEIPT OF MATERIALS

Receipt of PSSA science materials began on May 16, 2008, and concluded on June 24, 2008. DRC's Operations Material Management System (Ops MMS) was utilized to receive secure materials securely, accurately, and efficiently. This system features advanced automation and cutting-edge barcode scanners. Captured data were organized into reports, which provided timely information with respect to suspected missing material.

The first step in the Ops MMS was the Box Receipt System. When a shipment arrived at DRC, the boxes were removed from the carrier's truck and passed under a barcode reader, which read the barcode contained on the return label and identified the district and school. If the label could not be read automatically, a floor operator entered the information into the system manually. The data collected in this process were stored in the Ops MMS database. After the barcode data were captured, the boxes were placed on a pallet and assigned a corresponding pallet number. A "three way match" among the district box count, the carrier box count, and the DRC return box count was conducted to verify a box return accuracy rate of 100%.

Once the box receipt process was completed, the materials separation phase began. Warehouse personnel opened the district boxes and sorted the contents by grade and status (used/unused) into new boxes. Once filled, a sorted box's documents were loaded into an automated counter, which recorded a booklet count for each box. An on-demand DRC box label was produced that contained a description of each box's contents and quantity in both barcode and human-readable format. This count remained correlated to the box as an essential quality control step throughout secure booklet processing and provided a target number for all steps of the check-in process.

Once labeled, the sorted and counted boxes proceeded to booklet check-in. This system used streamfeeder automation to carry documents past oscillating scanners that captured data from up to two representative barcodes and stored it in the Ops MMS database.

- The secure booklet check-in operator used a hand scanner to scan the counted box label. This procedure input material type and quantity parameters for what the Ops MMS should expect within a box. It then loaded the box's contents into the streamfeeder.
- The documents were fed past oscillating scanners that captured either a security code or both a security code and a pre-code, depending upon material type. A human operator monitored an Ops MMS screen, which displayed scan errors, an ordered accounting of what was successfully scanned, and the document count for each box.
- When all materials were scanned and the correct document count was reached, the box was sealed and placed on a pallet. If the correct document count was not reached, or if the operator encountered difficulties with material scanning, the box and its contents were delivered to an exception handling station for resolution.

This check-in process occurred immediately upon receipt of materials; therefore, DRC provided immediate feedback to districts and schools regarding any missing materials based on actual receipts versus expected receipts.

Upon completion of secure booklet check-in, DRC produced a Missing Materials Report that listed all schools in each participating district and the number of booklets, by grade, for each school that were not returned to DRC.

After scannable materials were processed through Book Receipt, the materials became available to the DRC Document Processing Center Log-in staff for document log-in. Based on a pre-determined sampling and calibration plan, the staff prioritized answer documents using the following process:

- A DRC scannable barcode batch header was scanned, and a batch number was assigned to each box of answer documents.
- The DRC box label barcode was scanned into the system to link the box and answer documents to the newly created batch and to create a Batch Control Sheet.
- The DRC box label barcode number, along with the number of answer documents in the box, was printed on the Batch Control Sheet for document tracking purposes. All documents that were linked to the box barcode were assigned to the batch number and tracked through all processing steps. As documents were processed, DRC staff dated and initialed the Batch Control Sheet to indicate that proper processing and controls were observed.
- Before the answer documents were scanned, all batches went through a quality inspection to ensure batch integrity and correct document placement.

After a quality check in the DRC Document Processing Log-in area, the spines were cut off the scannable documents, and the pages were sent to DRC's Imaging and Scoring System.

SCANNING OF MATERIALS

DRC used its image scanning system to capture constructed-response items as images. These were then loaded into the image scoring system for both the handscoring of constructed-response items and for the capture of multiple-choice and demographic data.

DRC's image scanners were calibrated using a standard deck of scannable pages with 16 known levels of gray. On a predefined page location, the average pixel darkness was compared to the standard calibration to determine the level of gray. Marks with an average darkness level of 4 or above on a scale of 16 (0 through F) were determined to be valid responses, per industry standard. If multiple marks were read for a single item and the difference of the grayscale reads was greater than four levels, the lighter mark was discarded. If the multiple marks had fewer than four levels of grayscale difference, the response was flagged systematically and forwarded to an editor for resolution.

Customized scanning programs for all scannable documents were prepared to read the answer documents and to electronically format the scanned information. Before materials arrived, all image scanning programs went through a quality review process that included scanning of mock data from production booklets to ensure proper data collection.

- DRC's image scanners read selected-response, demographic, and identification information. The image scanners also used barcode readers to read pre-printed barcodes from a label on the booklet.
- The scannable documents were automatically fed into the image scanners where pre-defined processing criteria determined which fields were to be captured electronically. Constructed-response images were separated out for image-based scoring.
- During scanning, a unique serial number was printed on each sheet of paper. This serial number was used for document integrity and to maintain sequencing within a batch of answer documents.
- A monitor randomly displayed images, and the human operator adjusted or cleaned the scanner when the scanned image did not meet DRC's strict quality standards for image clarity.
- All images passed through a process and a software clean-up program that despeckled, deskewed, and desmeared the images. A random sample of images was reviewed for image quality approval. If any document failed to meet image quality standards, the document was returned for rescanning.
- Page scan verification was performed to ensure that all pre-defined portions of the answer documents were represented in their entirety in the image files. If a page was missing, the entire answer document was flagged for resolution.

After each batch was scanned, answer documents were processed through a computer-based edit program to detect potential errors as a result of smudges, multiple marks, and omits in predetermined fields. Marks that did not meet the pre-defined editing standards were routed to editors for resolution.

- Experienced DRC Document Processing Center Editing staff reviewed all potential errors detected during scanning and made necessary corrections to the data file. The imaging system displayed each suspected error. The editing staff then inspected the image and made any needed corrections using the unique serial number printed on the document during scanning.
- Upon completion of editing, quality control reports were run to ensure that all detected potential errors were reviewed again and a final disposition was determined.

Before batches of answer documents were extracted for scoring, a final edit was performed to ensure that all requirements for final processing were met. If a batch contained errors, it was flagged for further review before being extracted for scoring and reporting.

- During this processing step, the actual number of documents scanned was compared to the number of answer documents assigned to the box during book receipt. Count discrepancies between book receipt and answer documents scanned were resolved at this time.
- Once all requirements for final processing were met, the batch was released for scoring and student level processing.

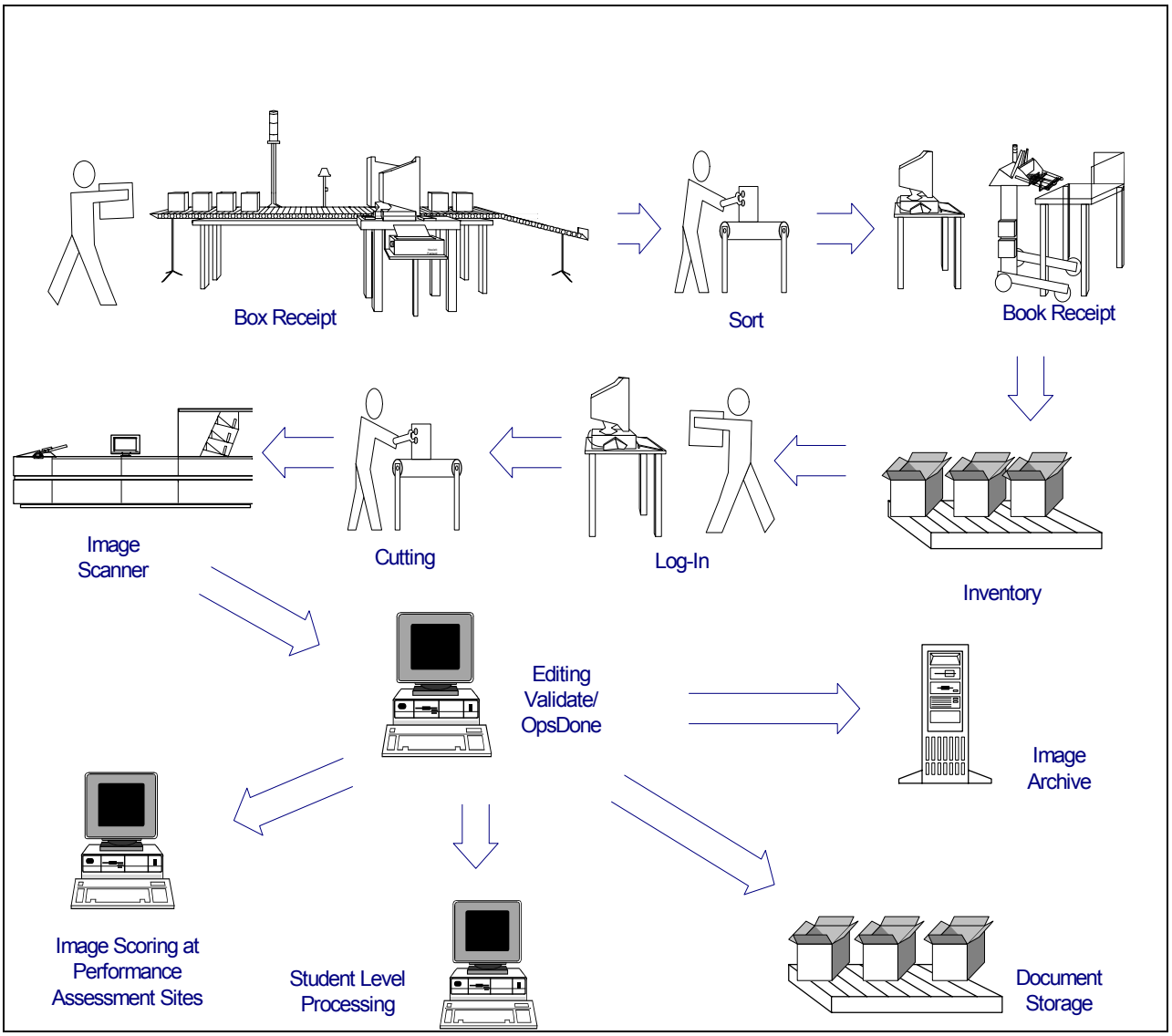
Table 8–1 shows the number of answer booklets received through booklet check-in and the number of booklets that contained student responses that were scanned and scored.

Table 8–1. Counts of 2008 PSSA Science Materials Received – Grades 4, 8, and 11

	Science Booklets Received	Used Science Booklets Scanned
Grade 4	164,743	128,838
Grade 8	174,425	141,748
Grade 11	174,752	137,859

Figure 8–1 illustrates the production workflow for DRC's Ops MMS and Image Scanning and Scoring System from receipt of materials through all processing of materials and the presentation of scanned images for scoring.

Figure 8–1. Workflow System



MATERIALS STORAGE

Upon completion of processing, student answer booklets are boxed for security purposes and final storage:

- Project-specific box labels were created containing unique customer and project information, materials type, batch number, pallet/box number, and the number of boxes for a given batch.
- Boxes were stacked on project-specific pallets that were labeled with a list of its contents and delivered to the Materials Distribution Center for final secure storage.
- Materials will be destroyed one year after contract year ends with PDE written approval.

SCORING MULTIPLE-CHOICE ITEMS

The scoring process included the scoring of multiple-choice items against the answer key and the aggregation of raw scores from the constructed responses. A student's raw score is the actual number of points achieved by the student for tested elements of an assessment. From the raw scores, the scale scores were calculated.

The student file was scored against the finalized and approved multiple-choice answer key. Items were scored as right, wrong, omitted, or double-gridded (more than one answer was bubbled for an item). Sections of the test were evaluated as a whole and an attempt status was determined for each student. The score program defined all data elements at the student level for reporting.

RANGEFINDING

After student documents were received and processed, DRC's Performance Assessment Services (PAS) staff began to assemble groups of responses that exemplified the different score points represented in the 0–2 item-specific scoring guidelines for the short responses and the Part A & B, 0–2, item-specific scoring guidelines for the scenario responses. Papers were pulled for the grades 4, 8, and 11.

Once examples for all the score points were identified, sets were assembled for each item. These sets were copied for use at rangefinding, held May 28–30, 2008, at the Sheraton Harrisburg-Hershey in Harrisburg. The rangefinding committees consisted of Pennsylvania educators, PDE staff members, DRC Test Development staff, and DRC Performance Assessment Services staff.

The joint session began with a review of the history of the 2008 assessment and then broke into grade-level groups. Copies of the student example sets were presented to the committees, one item at a time. The committees reviewed and scored the student samples together to ensure that everyone was interpreting the scoring guidelines consistently. Committee members then went on to score responses independently and those scores were discussed until a consensus was reached. Only responses for which a good agreement rate was attained were used in training the readers. Discussions of the responses used the language of the scoring guidelines, assuring PDE and all involved that the score point examples clearly illustrated the specific requirements of each score level. DRC PAS staff made notes of how and why the committees made these score point decisions, and this information was used by the individual scoring directors in reader training.

DRC and PDE discussed scoring guideline edits that the committees suggested. Changes approved by PDE were then made by DRC Test Development and the scoring guidelines were used by PAS staff in the preparation of materials and training of readers.

READER RECRUITMENT/QUALIFICATIONS

DRC retains a number of experienced readers from year to year, and those readers made up approximately 60% of the reader pool (N=225) for this project. To complete the reader staff for this project, DRC placed advertisements in local papers, minority publications, teacher newsletters, and at regional colleges and universities. Open houses were held and applications for reader positions were screened by the DRC recruiting staff. Candidates were personally interviewed and a mandatory, on-demand writing sample was collected, along with references and proof of a four-year college degree. In this screening process, preference was given to candidates with previous experience scoring large-scale assessments and with degrees emphasizing expertise in science. Since readers had to have a strong content-specific background, the reader pool consisted of educators and other professionals who were valued for their experience, but who were also required to set aside their own biases about student performance and accept the scoring standards. All readers on this assessment held at least a four-year degree.

LEADERSHIP RECRUITMENT/QUALIFICATIONS

Scoring directors and team leaders were chosen by the project director from a pool consisting of experienced individuals who were successful readers and leaders on other DRC contracts and had strong backgrounds in scoring science. Those selected demonstrated organization, leadership, and management skills. The scoring directors and a majority of the team leaders had at least five years of leadership experience on the PSSA. All scoring directors, team leaders, and readers were required to sign confidentiality forms before any training or handling of secure materials began.

Each room of readers was assigned a scoring director. This individual was monitored by the project manager and project content coordinator and led the hand scoring for the duration of the project. The scoring director assisted in rangefinding, worked with supervisors to create training materials, conducted the team leader training, and was responsible for training the readers. The scoring director also made sure that reports were available and interpreted reports for the readers. The scoring director supervised the team leaders.

Team leaders assisted the scoring director with reader training and monitoring by working with their teams in small group discussions and answering individual questions that readers may not have felt comfortable asking in a large group. Once readers had qualified, the team leaders were responsible for maintaining the accuracy and workload of team members. The ongoing monitoring identified those readers who were having difficulty scoring accurately and resulted in the reader receiving one-on-one retraining or in pairing that reader with a stronger reader. This process corrected any inaccuracies in scoring and, if not, that reader was released from the project.

TRAINING

After rangefinding was completed, DRC's PAS staff compiled the approved scoring guidelines and the scored student examples from the committees into packets used for training the readers.

Responses that illustrated the scoring concepts were used in a scoring guide. The item-specific scoring guidelines served as the reader's constant reference. Readers were instructed how to apply the guidelines and were required to demonstrate a clear comprehension of each anchor set by performing well on the training materials that were presented for each grade and item. These sets consisted entirely of examples of student responses reviewed by the rangefinding committee.

Team leaders assisted the scoring directors with the training and monitoring of readers. During their training, all materials were reviewed and discussed, and anticipated reader questions and concerns were addressed. Team leaders were required to annotate all of their training responses with the official annotations received from the content committee members at the rangefinding meetings. To facilitate scoring consistency, it was imperative that each team leader imparted the same rationale for each response that other team leaders used. A ratio of one team leader for each 8–10 readers ensured adequate monitoring of the readers.

Reader training consisted of the scoring director providing a review of the scoring guidelines and anchor papers assembled for each item. Readers were instructed how to apply the item-specific scoring guides and were required to demonstrate a clear comprehension of each item's anchor set by performing well on the training sets that were presented for each item.

HANDSCORING PROCESS

Student responses were scored independently and by multiple readers with a 10% double read, plus team leader read behind, to ensure reliability. The 10% double reads were randomly chosen by the imaging system at the item level. The PDE determined the required number of reads.

Readers scored the imaged student responses on PC monitors at the Minnetonka and Woodbury, Minnesota, Scoring Centers. Readers were seated at tables with two imaging stations at each table. Image distribution was controlled, thus ensuring that they were sent to designated readers qualified to score those items. Readers read each response and keyed the scores.

To handle possible alerts (i.e., student responses indicating potential issues related to the student's safety and well-being that may require attention at the state or local level), the imaging system allowed readers to forward responses needing attention to the scoring director. These alerts were reviewed by the project director, who then notified that student's school and the PDE of this occurrence. However, PDE did not receive the student's responses or other identifying information on that student. Readers, also, had no information on the student's personal identity.

Once handscoring was completed, PAS compiled anecdotal item reviews of the field test prompts for all grade levels. This information was handed on to DRC Test Development.

QUALITY CONTROL

Reader accuracy was monitored throughout the scoring session by producing both daily and on-demand reports, ensuring that an acceptable level of scoring accuracy was maintained. Inter-reader reliability was tracked and monitored with multiple quality control reports that were reviewed by quality assurance analysts. These reports were generated at the handscoring center and were reviewed by the scoring directors, team leaders, project content specialists, and project directors. The following reports were used during the scoring of the constructed responses:

The Reader Monitor Report monitored how often readers were in exact agreement and ensured that an acceptable agreement rate was maintained. This report provided daily and cumulative exact and adjacent inter-reader agreement and the percentage of responses requiring resolution.

The Score Point Distribution Report monitored the percentage of responses given each of the score points. For example, this daily and cumulative report showed how many 0s, 1s, and 2s, a reader had given to all the responses he or she had scored at the time the report was produced. It also indicated the number of responses read by each reader so that production rates could be monitored.

The Item Status Report monitored the progress of handscoring. This report tracked each response and indicated the status of each item (i.e., "needs second reading," "complete"). This report ensured that all discrepancies were resolved by the end of the project.

The Response Read by Reader Report identified all responses scored by an individual reader. This report was useful if any responses needed rescoring because of reader drift.

The Read-Behind Log was used by the team leader/scoring director to monitor reader reliability. Student responses were randomly selected and team leaders read scored items from each team member. If the team leader disagreed with the reader's score, remediation occurred. This proved to be a very effective type of feedback because it was done with "live" items scored by a particular reader.

Table 8–2 shows the exact and adjacent agreement rates of the readers for grades 4, 8, and 11.

Table 8–2 Inter-rater Agreement for 2008 Grades 4, 8, & 11 Science

Grade	Item	% Exact Agreement	% Adjacent Agreement	% Exact + Adjacent
4	1	82	18	100
	2	92	8	100
	3	81	19	100
	4	90	10	100
	5	88	12	100
8	1	80	20	100
	2	97	3	100
	3	80	20	100
	4	84	16	100
	5	87	13	100
11	1	81	18	99
	*2a	77	23	100
	*2b	72	28	100
	*3a	72	28	100
	*3b	87	13	100
	4	76	24	100
	5	93	7	100
	6	79	20	99
	*7a	86	14	100
	*7b	81	19	100
	8	89	11	100
	9	78	21	99

* Denotes scenario items

Tables 8–3 through 8–5 show the percentages awarded for each possible score point. All items are scored with a 0–2 score point range. Scenario items have each part scored 0–2 and parts a and b are combined for the total score.

Table 8–3. Grade 4 Score Point Distribution

Item	%0	%1	%2	%B/NS*
1	13	37	49	1
2	19	24	55	1
3	20	44	35	1
4	8	19	72	1
5	23	57	19	1

Table 8–4. Grade 8 Score Point Distribution

Item	%0	%1	%2	%B/NS*
1	6	55	38	2
2	23	57	14	7
3	38	38	21	4
4	14	35	48	3
5	29	40	28	4

Table 8–5. Grade 11 Score Point Distribution

Item	%0	%1	%2	% B/NS
1	33	39	18	9
2a	49	38	12	10
2b	29	48	23	
3a	41	47	12	11
3b	76	22	2	
4	16	43	34	7
5	33	12	36	19
6	25	46	17	11
7a	74	18	8	20
7b	38	37	25	
8	52	29	9	11
9	13	48	22	12

*B denotes blank, NS denotes non-scorable

MATCH-BACK RULES

In order to create a single student record in the central student file, it was necessary to establish match-back rules to combine separate student records into one student record. Match-back rules were applied to link multiple-choice and constructed responses. They were also used to merge student responses captured on different subjects and to link test results with student demographic information.

DATA EXCHANGE, STORAGE, AND RECOVERY POLICIES

Data Exchange Procedures

The exchange of data between DRC, PDE, and other contractors is a critical and essential component in the success of the PSSA program. To support this process, DRC used the following data exchange procedures to ensure that all data files were successfully and accurately transferred.

- Files were posted to DRC's secure Pennsylvania FTP site with a standard and logical folder structure.
- Standard file naming conventions were established and used.
- The information necessary to perform these quality control procedures accompanied each data exchange.

Data Exchange Quality Control Procedures

- ***Record Count Check –*** Confirm the expected record count and provide the record count in files sent and received.
- ***File Count Check–*** Confirm that the number of files sent and received matches the number of files expected.
- ***Duplicate File Check –*** Verify that duplicate files were not sent or received.
- ***File Date*** – Verify that the version of the file received matches the file creation date.
- ***File Type Verification Check –*** Verify that data sent and received matches the format expected (e.g., Excel, CSV, PDF, Text file [delimited/fixed field length]).
- ***File Log –*** A log of files sent and received will be maintained.
- ***Data Validation –*** Data checking procedures will be used to verify that the data is in the specified file layout and matches the expected values.

Images

As part of the scanning process, the multi-page TIFF images were archived to tape before being separated into single page TIFFs and transmitted to the scoring centers. If any of the images were lost/deleted/corrupted at a scoring center, they could be restored from the archived multi-page TIFF images. In addition to archiving the images, the scoring center servers used RAID (Redundant Array of Independent Disks) 5 disk management technology to mirror the images to redundant disk drives. If a disk drive failed in a scoring center server, the images could be quickly restored from the redundant disk drive. In the event that the disk drive and the multi-page TIFF images could not be restored, the original documents would be rescanned. Images are stored for a PDE specified period.

Data

Once a reader submitted a score for a constructed-response item, the data was electronically transmitted to our SQL Servers. The log files documenting the changes were backed up hourly. Full back-ups were done nightly (Monday–Friday) and two additional full back-ups were run over the weekend on the handscoring SQL Servers with the backup tapes being rotated off-site. All data is stored for a PDE-specified period.

Storage

All physical servers are housed in secure server rooms in DRC's corporate headquarters in Maple Grove, or the Brooklyn Park or Woodbury locations. The server rooms are constructed of concrete floors, walls, and ceilings and designed to be fire and crush proof. They have fire suppression systems to minimize the effect of any fire started within the server room. Access to the server rooms is controlled through a card access system and is restricted to authorized technology support staff only. A log is maintained documenting each time a server room is entered, by whom, and for what purpose. In case of a disaster at any of the locations, another server can take over full operations.

DRC maintains backup servers that can be used to replace a failed server within 24 hours. Every server's configuration is documented in the event a rebuild is required. Each server has an assigned primary and secondary network analyst responsible for its operation.

The servers utilize load-sharing, redundant power supplies and implement RAID subsystems to minimize the effect of a failed disk. The server rooms all have Uninterruptible Power Supply (UPS) systems. For longer periods of power failure, an on-site diesel power generate will automatically start and supply needed power. The computing environment, both servers and communications hardware, will continue to function without interruption when the utility power is disrupted.

Two copies of complete system and data backup are created each weekend. One of these copies is stored in a secure room at the Maple Grove location. The second copy is stored in a secure room at the Woodbury location. These backups are stored indefinitely. Incremental backups of all files on the network are made each day. The incremental backups are kept for 6 weeks.

DRC utilizes a storage area network (SAN) for maximum speed, flexibility, and redundancy in our data storage solution. Servers are connected to the SAN via redundant connections to ensure minimum interruptions due to hardware failures. The SAN allows disk space to be reallocated with ease for availability to those applications or servers as needed. The SAN currently houses 13 Terabytes of storage and is expandable to 26 Terabytes.

Chapter Nine: Summary Demographic, Program, and Accommodation Data for the 2008 Science PSSA

ASSESSED STUDENTS

The total number of answer documents processed by grade level for the 2008 Science PSSA is presented on the first line of Table 9–1. This number pertains to the total number of records on the student file and is typically less than the "Used Science Booklets Scanned" column shown in Table 8–1. The reason for the difference is that completely blank answer booklets (no student name and no items responded to) get removed from the initial batch of materials scanned. See Chapter Eight for more details on processing. The second line shows the number and percent of students with a PSSA science score, followed by the number and percent not receiving a score. The final line gives the number of students contributing to state summary statistics, which is especially relevant for all tables following Table 9–2. (See the section of this chapter entitled, "Composition of Sample Used in Subsequent Tables" for additional explanation.)

Assessed students include those from public schools who are required to participate as well as those from a small number of non-public schools (numbering fewer than 500 per grade level) that elected to participate. Also included were home-schooled students, which numbered fewer than 100 per grade.

As may be observed from Table 9–1, not all students were assessed. Although there are a variety of reasons for this, the major ones pertain to

1. extended absence from school that continued beyond the assessment window

2. being absent without make-up for at least one section of the science assessment

3. a situation in which there was a non-attempt (failed to meet the criteria of having attempted one of more sections of the science test) on the part of the student and no exclusion code was marked by school personnel.

4. ELL students in the first year in U.S. schools

5. medical emergency

6. other reasons (includes parental request due to religious reasons, students who are court-agency placed, students with multiple reasons coded, and the category of "other").

The number of students without scores for these reasons is presented in Table 9–2.

Students in an assessed grade who met each of the following criteria were excused from the PSSA and should have participated in an alternate science assessment provided by the school entity:

1. had a significant cognitive disability,

2. required intensive instruction,

3. required adaptation and support to perform or participate meaningfully,

4. required substantial modification of the general education curriculum,

5. participation in the general education curriculum differed markedly in form and substance from that of other students (see *2008 PSSA Handbook for Assessment Coordinators and Administrators: Grades 4, 8, and 11 Science, PDE, 2008, p.7*).

A student must have completed five or more items in each of the science sections of the assessment to have been considered attempted.

Table 9–1. Students Assessed on the 2008 Science PSSA

	Grade 4		Grade 8		Grade 11	
	N	Pct	N	Pct	N	Pct
Number of answer documents processed	128,839		141,760		137,921	
Students with a science score	128,235	99.5	139,731	98.6	132,653	96.2
Number processed but not assessed (without a total score)	604	0.5	2,029	1.4	5,268	3.8
Students with a science score used in state summaries	126,426		137,790		131,157	

Table 9–2. Counts of Students without Scores on the 2008 Science PSSA

Reason for Non-Assessment	Grade 4		Grade 8		Grade 11	
	N	Pct	N	Pct	N	Pct
Extended Absence from School	97	16.1	527	26.0	1,375	26.1
Absent Without Make-up	105	17.4	354	17.4	952	18.1
Non-Attempt Science	276	45.7	740	36.5	2,065	39.2
Medical Emergency	49	8.1	153	7.5	243	4.6
Other Reasons	77	12.7	255	12.6	633	12.0
Total Count Not Assessed	604		2,029		5,268	

COMPOSITION OF SAMPLE USED IN SUBSEQUENT TABLES

Students included in the following demographic analyses were those who contributed to state summary statistics, using the individual student data file of September 24, 2008 (grade 4) or October 14, 2008 (grades 8 and 11). Consequently, these results may not reflect some data used in subsequent updates. Students not included in the present data were those who (1) enrolled in a Pennsylvania school after October 1, 2007, (2) were coded as ELL and enrolled after March 23, 2007, or (3) were home schooled.

Because some student file updates may occur subsequent to these analyses, there could be small differences in the counts along with slight percentage differences.

COLLECTION OF STUDENT DEMOGRAPHIC INFORMATION

Data for analyses involving demographic characteristics were obtained primarily from information initially supplied by school district personnel through the Pennsylvania Information Management System (PIMS) and subsequently transmitted to DRC. Later corrections and updates of demographic and attribution (for AYP) data were carried out through the DRC Student Precode and Corrections System. Some data such as accommodation information is marked directly on the student answer document at the time the PSSA is administered.

DEMOGRAPHIC CHARACTERISTICS

Frequency data for each category is presented in Table 9–3. Percentages are based on all students with a score in science as shown at the bottom of the table.

Table 9–3. Demographic Characteristics of 2008 Science PSSA

Demographic or Educational Characteristic	Grade 4		Grade 8		Grade 11	
	N	Pct	N	Pct	N	Pct
Gender						
Female	61,571	48.7	66,857	48.5	64,962	49.5
Male	64,669	51.2	70,667	51.3	65,937	50.3
Race/Ethnicity						
American Indian or Alaskan Native	200	0.2	239	0.2	205	0.2
Asian or Pacific Islander	3,716	2.9	3,576	2.6	3,312	2.5
Black/African American non-Hispanic	19,631	15.5	21,108	15.3	15,262	11.6
Latino/Hispanic	9,556	7.6	9,063	6.6	6,314	4.8
White non-Hispanic	92,148	72.9	102,757	74.6	105,228	80.2
Multi-Racial/Ethnic	950	0.8	698	0.5	482	0.4
Educational Category and Other Demographic Groups						
IEP (not gifted)	20,238	16.0	21,169	15.4	17,239	13.1
Student exited IEP in last 2 years	1,449	1.1	798	0.6	347	0.3
Gifted and has an IEP	5,250	4.2	8,032	5.8	7,129	5.4
504 Plan / Chapter 15	1,376	1.1	1,509	1.1	1,296	1.0
Title I	40,851	32.3	26,511	19.2	15,111	11.5
Title III - Served	2,841	1.7	1,979	1.4	1,210	0.9
Title III - Not Served	468	0.3	313	0.2	283	0.2
Migrant Student	127	0.1	116	0.1	83	0.1
ELL (enrolled after 3-23-07)	362	0.3	309	0.2	221	0.2
ELL (enrolled before 3-23-07)	2,947	2.3	1,983	1.4	1,272	1.0
Exited ESL/bilingual program – 1 yr	624	0.5	330	0.2	183	0.1
Exited ESL/bilingual program – 2 yr	640	0.5	525	0.4	300	0.2
Former ELL no longer monitored	884	0.7	1,230	0.9	844	0.6
Foreign Exchange Student	2	0.0	5	0.0	98	0.1
Economically Disadvantaged	47,988	38.0	46,698	33.9	31,773	24.2
Enrollment						
Current Enrollment in school of residence after Oct 1, 2007	4,345	3.4	4,429	3.2	3,225	2.5
Current Enrollment in district of residence after Oct 1, 2007	2,262	1.8	2,490	1.8	2,016	1.5
Current Enrollment as PA resident after Oct 1, 2007	0	0.0	0	0.0	0	0.0
Enrolled in school of residence after Oct 1, 2006 but on/before Oct 1, 2007	21,468	17.0	20,973	15.2	16,731	12.8
Enrolled in district of residence after Oct 1, 2006 but on/before Oct 1, 2007	12,523	9.9	13,252	9.6	10,918	8.3
Homeless as defined by McKinney-Vento Act	198	0.2	196	0.1	107	0.1
School Choice Provision	61	0.0	149	0.1	9	0.0
Total Number Scored	126,426		137,790		131,157	

EDUCATION IN NON-TRADITIONAL SETTINGS

For each category the number and percent are presented for all students with a score in science. Table 9–4 reveals an incidence of less than one percent for the majority of these settings. Also shown are home schooled students assessed by parental request.

Table 9–4. Participation in 2008 Science PSSA by Students in Non-Traditional Settings

Non-Traditional Educational Settings	Grade 4		Grade 8		Grade 11	
	N	Pct	N	Pct	N	Pct
Court/agency placed	110	0.1	623	0.5	846	0.6
Homebound instruction	1	0.0	2	0.0	16	0.0
Special education student with IEP placed in program operated by entity other than district/school of residence	868	0.7	1,228	0.9	1,152	0.9
Student placed in alternative Education for Disruptive Youth program conducted outside of regular classroom	62	0.0	773	0.6	718	0.5
Home schooled student assessed by parental request	0	0.0	0	0.0	0	0.0
ELL student tested outside district	63	0.0	42	0.0	1	0.0

PRIMARY DISABILITY OF IEP STUDENTS ASSESSED ON THE PSSA

School personnel supplied the primary disability information for those students who had an IEP (not gifted) through the Pennsylvania Information Management System (PIMS) in conjunction with the DRC Student Precode System. Beginning with the 2006 assessment, the disability categories were presented in a sequence matching a Department of Education numbering system and two previously separate categories were combined. In Table 9–5, for assessed students classified as IEP, the number and percent in each disability category are presented. For example, if 20,000 students statewide were categorized as IEP and 4,000 students were classified with a particular disability, the table entry would display 4,000 followed by 20 percent. Data is also supplied for the substantial percentage of students who were unclassified with respect to a specific disability. The last row presents the percent of assessed students classified as IEP. The most prominent finding in Table 9–5 is that specific learning disability is the category with the highest incidence of occurrence across all grade levels.

Table 9–5. Incidence of Primary Disabilities among IEP Students Assessed on the 2008 Science PSSA

Primary Disability of Students Having an IEP	Grade 4		Grade 8		Grade 11	
	N	Pct	N	Pct	N	Pct
Traumatic Brain Injury	17	0.1	50	0.2	35	0.2
Hearing Impairment including Deafness	166	0.8	155	0.7	137	0.8
Specific Learning Disability	9,896	48.9	13,479	63.7	11,090	64.3
Mental Retardation	430	2.1	757	3.6	714	4.1
Orthopedic Impairment	33	0.2	29	0.1	34	0.2
Emotional Disturbance	971	4.8	1,685	8.0	1,474	8.6
Speech or Language Impairment	4,843	23.9	704	3.3	145	0.8
Visual Impairment including Blindness	62	0.3	56	0.3	48	0.3
Deaf/Blind	3	0.0	6	0.0	13	0.1
Multiple Disabilities	30	0.1	38	0.2	33	0.2

Table 9–5 (continued). Incidence of Primary Disabilities among IEP Students Assessed on the 2008 Science PSSA

Primary Disability of Students Having an IEP	Grade 4		Grade 8		Grade 11	
	N	Pct	N	Pct	N	Pct
Autism	664	3.3	448	2.1	246	1.4
Other Health Impairment	1,340	6.6	1,353	6.4	886	5.1
IEP Students without a disability code	1,783	8.8	2,409	11.4	2,384	13.8
Number of IEP Students Assessed	20,238	100.0	21,169	100.0	17,239	100.0
Percent of Assessed Students Classified as IEP		16.0		15.4		13.1

TEST ACCOMMODATIONS PROVIDED

School personnel supplied information regarding accommodations that a student may have received while taking the PSSA. Accommodations are classified in terms of presentation, response, setting, and timing modifications to enable students to better manage disabilities that hinder their ability to learn and respond to assessments.

The frequency with which these accommodations were utilized is summarized in Tables 9–6 through 9–9. The values in the table are based on all students with a score in science. Please note that a glossary of accommodation terms as applied to the PSSA is provided in Table 9–12.

PRESENTATION ACCOMMODATIONS RECEIVED

Presentation Accommodations are those which provide alternate ways for students to access and process printed instructional material and assessments. These include auditory, tactile, visual, and combined auditory/visual modes of presentation. There were eleven categories of presentation modifications in the 2008 PSSA science assessment. As depicted in Table 9–6 the actual frequencies are quite low, generally representing less than two-tenths of one percent of assessed students statewide. Two exceptions were test directions read aloud, signed, or recorded, and test items/questions read aloud or signed in which more than 1,000 instances occurred at each grade level but one.

Table 9–6. Incidence of Presentation Accommodations Received on the Science 2008 PSSA

Type of Presentation Accommodation	Grade 4		Grade 8		Grade 11	
	N	Pct	N	Pct	N	Pct
Braille format	4	0.0	8	0.0	11	0.0
Large print format	66	0.1	69	0.1	49	0.0
Magnification device	9	0.0	7	0.0	9	0.0
Reading windows, reading guides	109	0.1	4	0.0	5	0.0
Sign language interpreter	11	0.0	11	0.0	11	0.0
Qualified interpreter for ELL student	155	0.1	97	0.1	121	0.1
Test directions read aloud, signed, or recorded *	7,715	6.1	3,308	2.4	1,585	1.2
Test items/questions read aloud or signed *	9,400	7.4	2,717	2.0	948	0.7
Test items/questions recorded	329	0.3	61	0.0	22	0.0
Electronic screen reader *	0	0.0	1	0.0	8	0.0
Other	170	0.1	126	0.1	68	0.1

*Not a standardized test format modification made available through PDE. See Table 9–12 for more explanation.

RESPONSE ACCOMMODATIONS RECEIVED

Response Accommodations permit students to complete assignments, tests, and activities in different ways to solve or organize problems using some type of assistive device or organizer. There were eleven categories of response accommodations on the 2008 PSSA science assessment. The frequency with which these accommodations were utilized is summarized in Table 9–7. The actual frequencies are quite low, most representing less than two-tenths of one percent of assessed students statewide.

Table 9–7. Incidence of Response Accommodations Received on the 2008 Science PSSA

Type of Response Accommodation	Grade 4		Grade 8		Grade 11	
	N	Pct	N	Pct	N	Pct
Brailler / Note taker	3	0.0	6	0.0	10	0.0
Test Administrator scribed open-ended responses at student's direction	715	0.6	105	0.1	42	0.0
Test administrator marked multiple-choice responses at student's direction	424	0.3	95	0.1	29	0.0
Test Administrator transcribed student responses	405	0.3	193	0.1	75	0.1
Augmentative communication device	2	0.2	1	0.0	1	0.0
Typewriter, word processor or computer	17	0.0	65	0.0	26	0.0
Audio recording of student responses	0	0.0	0	0.0	1	0.0
Manipulative (Cranmer Abacus, Number Line)	0	0.0	0	0.0	0	0.0
Translation dictionary for ELL student	73	0.1	150	0.1	210	0.2
Electronic Screen Reader *	0	0.0	0	0.0	10	0.0
Other (approved by PDE)	139	0.1	70	0.1	47	0.0

*Not a standardized test format modification made available through PDE. See Table 9–12 for more explanation.

SETTING ACCOMMODATIONS RECEIVED

Setting Accommodations permit a change in location in which a student receives instruction or participates in an assessment. There were three categories of setting accommodations on the 2008 PSSA science assessment. As depicted in Table 9–8, the most common accommodations were small group testing and testing in a separate setting. Typically the percentage of IEP students tested in separate or small group settings decreases from lower to higher grade levels.

Table 9–8. Incidence of Setting Accommodations Received on the 2008 Science PSSA

Type of Setting Accommodation	Grade 4		Grade 8		Grade 11	
	N	Pct	N	Pct	N	Pct
Hospital/home testing	21	0.0	62	0.0	96	0.1
Tested in separate setting	6,263	5.0	3,878	2.8	2,546	1.9
Small group testing	12,416	9.8	9,940	7.2	6,611	5.0

TIMING ACCOMMODATIONS RECEIVED

Timing Accommodations involve a change in the allowable length of time to complete assignments or assessments, including the way in which time is organized. There were four categories of timing accommodations on the 2008 PSSA science assessment. As depicted in

Table 9–9, the most common accommodations were scheduled extended time and requested extended time.

Table 9–9. Incidence of Timing Accommodations Received on the 2008 Science PSSA

Type of Timing Accommodation	Grade 4		Grade 8		Grade 11	
	N	Pct	N	Pct	N	Pct
Scheduled Extended Time	5,164	4.1	3,158	2.3	2,742	2.1
Requested Extended Time	1,474	1.2	1,934	1.4	1,540	1.2
Multiple Test Sessions	1,105	0.9	925	0.7	891	0.7
Changed Test Schedule	160	0.1	71	0.1	80	0.1

INCIDENCE OF ACCOMMODATIONS AND IEP AND ELL STATUS

It is reasonable to expect that students with an IEP would receive the majority of accommodations; however, certain accommodations are specific to particular disabilities or to students classified as English Language Learners (ELL). A cross-tabulation between each of the accommodations and IEP and ELL status revealed a much greater incidence for students with an IEP. As observed in Tables 9–6 through 9–9, the most frequently occurring accommodations for assessed students were:

- Test directions read aloud, signed, or recorded
- Test items/questions read aloud or signed
- tested in separate setting
- small group testing
- scheduled extended time
- requested extended time
- multiple test sessions.

Because the accommodations with the largest frequencies can potentially supply the most stable data when broken out for subgroup analysis, these were selected for display in Table 9–10.

Coding for IEP is dichotomous as students are classified as IEP and non-IEP. For purposes of this analysis, an English Language Learner (ELL) was a student classified as ELL and enrolled in a U.S. school on or before March 23, 2007. All other assessed students, including those who exited an ESL/bilingual program and in the first or second year of monitoring were regarded as "Not ELL." Students coded as ELL and enrolled in a U.S. school after March 23, 2007 are excluded from state summary statistics as stated earlier in this chapter.

Customarily, a considerably larger percent of IEP students receive a given accommodation than Non-IEP students. Likewise, certain accommodations tend to occur more frequently for ELL students than for non-ELL students. To separate out the effect of being classified IEP or ELL, four possible combinations are presented in Table 9–10. These include: (1) General education students (those who are neither IEP nor ELL), and students classified as (2) IEP but not ELL, (3) ELL but not IEP, and (4) both IEP and ELL. The bottom row for each grade provides the total number of students having a science score within each of the four classifications.

General findings from Table 9–10 include:

- General education (neither IEP nor ELL) students had a very low incidence of accommodations, as 17 of the 21 instances (across grades) were less than one percent and only one reached two percent.
- The IEP / not ELL students generally manifested the largest percentage of these accommodations (13 of 21 instances) and second largest in the remaining instances.
- The ELL / not IEP students received a larger percent of these accommodations than the General education students in all instances, but typically fewer than the IEP / not ELL group.
- The Both IEP and IEP group had the largest percent of accommodations in eight instances and exceeded the ELL / not IEP group in 18 of 21 instances.

Table 9–10. Incidence of IEP and ELL Students Receiving Selected Accommodations

	Classification of Students Regarding IEP and ELL							
Accommodation Received	**General education (Not IEP or ELL)**		**IEP and not ELL**		**ELL and not IEP**		**Both IEP and ELL**	
Grade 4	**N**	**Pct**	**N**	**Pct**	**N**	**Pct**	**N**	**Pct**
Test directions read aloud, signed, recorded	748	0.7	6,438	32.7	338	12.2	191	35.7
Test items/questions read aloud or signed	1,185	1.1	7,532	38.2	459	16.5	224	41.9
Tested in separate setting	1,096	1.1	4,648	23.6	383	13.8	136	25.4
Small group testing	2,142	2.1	9,367	47.5	655	23.6	252	47.1
Scheduled extended time	578	0.6	4,214	21.4	238	8.6	134	25.0
Requested extended time	837	0.8	535	2.7	69	2.5	33	6.2
Multiple test sessions	120	0.1	936	4.8	29	1.0	20	3.7
Column N for Grade 4	103,414		19,703		2,774		535	
Grade 8	**N**	**Pct**	**N**	**Pct**	**N**	**Pct**	**N**	**Pct**
Test directions read aloud, signed, recorded	162	0.1	3,015	14.1	93	4.6	38	13.9
Test items/questions read aloud or signed	152	0.1	2,411	11.5	113	5.6	41	15.0
Tested in separate setting	401	0.3	3,306	15.8	140	6.9	31	11.4
Small group testing	791	0.7	8,759	41.9	305	15.1	85	31.1
Scheduled extended time	383	0.3	2,642	12.6	103	5.1	30	11.0
Requested extended time	1,102	1.0	742	3.6	81	4.0	9	3.3
Multiple test sessions	101	0.1	768	15.2	46	2.3	10	3.7
Column N for Grade 8	114,602		20,896		2,019		273	
Grade 11	**N**	**Pct**	**N**	**Pct**	**N**	**Pct**	**N**	**Pct**
Test directions read aloud, signed, recorded	75	0.1	1,452	8.5	48	3.5	10	8.1
Test items/questions read aloud or signed	152	0.1	2,411	11.5	113	5.6	41	15.0
Tested in separate setting	294	0.3	2,103	12.3	139	10.1	10	8.1
Small group testing	515	0.5	5,836	34.1	224	16.4	36	29.3
Scheduled extended time	231	0.2	2,334	13.6	162	11.8	15	12.2
Requested extended time	1,050	0.9	430	2.5	53	3.9	7	5.7
Multiple test sessions	144	0.1	688	4.0	56	4.1	3	2.4
Column N for Grade 11	112,548		17,116		1,370		123	

INCIDENCE OF ACCOMMODATIONS AND PRIMARY DISABILITY CLASSIFICATION

To further delineate the use of commonly employed accommodations, a grade level breakdown by primary disability is presented in Table 9–11. A selection was made based on the more frequently occurring categories of disability and accommodations rather than displaying data for all of them. As may be seen from a perusal of Tables 9–6 through 9–9, the accommodations with the larger frequencies are those that involve a change in test timing, setting or that necessitate the reading of test directions. Selected for incorporation in Table 9–11 are seven accommodations with frequencies in excess of 1,000 in at least one grade level. These include two presentation, two setting, and three timing accommodations. Response accommodations tended to be infrequent and highly specific to particular and infrequent disability categories or to students classified as English Language Learner (ELL) and were not included in Table 9–11. Seven Primary Disability categories were selected that had a minimum of 100 students so classified at each grade level.

The entries for Table 9–11 represent the number and percent of students with a particular disability (columns) who received the listed accommodation (rows). For example, if 200 students out of 500 classified with a particular disability received scheduled extended time, the table entry would display 200 followed by 40%. The frequency of assessed students at each grade with a particular disability may be found in Table 9–5.

The most prominent and consistent findings from Table 9–11 are (1) the heavy use of test directions read aloud, signed, or recorded, test items/questions read aloud or signed, scheduled extended time, separate settings, and small group testing for all disability categories except speech and language impairment and that (2) the general tendency for the percent of grade 4 students receiving these particular accommodations to exceed that of grade 8 and 11 students.

Table 9-11 Incidence of Test Accommodations Received for Selected Primary Disability Classifications: 2008 Science PSSA

		Primary Disability of Assessed Student with an IEP: Number and Percent															
Type of Accommodation Received	Grade Level	Autism		Emotional Disturbance		Deafness / Hearing Impairment		Mental Retardation		Other Health Impairment		Specific Learning Disability		Speech or Language Impairment	
Test directions read aloud, signed, recorded	4	258	39%	304	31%	24	14%	176	41%	568	42%	4433	45%	198	4%
	8	55	12%	177	10%	17	11%	131	17%	188	14%	1993	15%	30	4%
	11	17	7%	114	8%	21	15%	98	14%	58	6%	834	7%	1	1%
Test items / questions read aloud or signed	4	288	43%	316	32%	30	18%	224	52%	673	50%	5274	53%	247	5%
	8	34	8%	102	6%	14	9%	115	15%	168	12%	1696	13%	17	2%
	11	13	5%	61	4%	7	5%	63	9%	39	4%	533	5%	1	1%
Separate setting	4	189	28%	217	22%	28	17%	128	30%	427	32%	3216	32%	210	4%
	8	66	15%	233	14%	12	8%	138	18%	255	19%	2234	17%	35	5%
	11	37	15%	148	10%	18	13%	125	17%	112	13%	1354	12%	8	5%
Small group testing	4	370	56%	498	51%	40	24%	250	58%	860	64%	6165	62%	389	8%
	8	211	47%	655	39%	34	22%	287	38%	602	44%	5758	43%	74	10%
	11	85	35%	524	35%	27	20%	260	36%	266	30%	3598	32%	21	14%
Scheduled extended time	4	180	27%	201	21%	15	9%	127	29%	355	26%	2807	28%	111	2%
	8	64	14%	215	13%	7	4%	107	14%	144	11%	1456	11%	18	3%
	11	42	17%	198	13%	20	15%	124	17%	83	9%	1282	12%	10	7%
Student-requested extended time	4	22	3%	37	4%	2	1%	15	3%	50	4%	317	3%	54	1%
	8	37	8%	62	4%	3	2%	28	4%	57	4%	469	3%	10	1%
	11	20	8%	32	2%	2	1%	29	4%	26	3%	269	2%	5	3%
Multiple test sessions	4	38	6%	80	8%	0	0%	21	5%	84	6%	541	5%	13	0%
	8	16	4%	96	6%	1	1%	27	4%	25	2%	260	2%	3	0%
	11	7	3%	78	5%	3	2%	35	5%	15	2%	199	2%	1	1%

Note: Results displayed are for most frequently occurring accommodations and disability classifications

GLOSSARY OF ACCOMMODATIONS TERMS

Table 9–12 provides a brief description of accommodations terms as used in the PSSA. School personnel identified the accommodations that a student received by marking a bubble in the student answer document as seen in the left column. The right column contains an explanation abstracted from the *2008 Accommodations Guidelines* (PDE, 2008, pages 24–46).

Table 9–12. Glossary of Accommodations Terms as Applied in the PSSA

Type of Testing Accommodation	**Explanation**
Student used the following Presentation Accommodations *(mark all that apply)*	
Braille format	Students may use a Braille edition of the test. Answers must then be transcribed into the answer booklet without alteration.
Large print format	Students with visual impairments may use a large print edition. Answers must then be transcribed into the answer booklet without alteration.
Magnification device	Devices to magnify print may be used for students with visual impairments and or print disabilities.
Reading windows, reading guides	Students with visual impairments may use reading windows and reading guides in all assessments.
Sign language interpreter	Deaf/hearing impaired students may receive test directions from a qualified interpreter. Signing is also permitted for: essay prompts (writing), items/questions (math, science only).
Qualified interpreter for ELL student	An interpreter may translate directions or clarify instructions for the assessments. They may translate, not define specific words or test questions on the mathematics and science tests. On the reading test interpreters may only translate directions and may not translate or define words in the passage or test questions.
Test directions read aloud, signed, or recorded	Directions for all PSSA tests may be read aloud, signed or presented by audio recording.
Test items/questions read aloud or signed	Students unable to decode text visually may have items / questions read aloud for mathematics and science only; however, words may not be defined.
Test prompts recorded	Writing essay prompts may be presented by audio recording.
Electronic screen reader (PDE approval required)	Students with a severe visual disability may use an electronic screen reader; however, PDE must approve the program and functions prior to the test window.
Other (PDE approval required)	Other presentation accommodations indicated in the *Accommodation Guidelines* may be provided; however, PDE approval is required prior to the test window.
Spanish mathematics version	This version may be taken by students whose first language is Spanish and who have been enrolled in U.S. schools for fewer than 3 years.

Table 9–12 (continued). Glossary of Accommodations Terms as Applied in the PSSA

Student used the following Response Accommodations *(mark all that apply)*	
Brailler / Note taker (per *Accommodations Guidelines*)	Students using this device as part of their regular instructional program may use it on the PSSA; however, without thesaurus, spell- or grammar- checker, etc.
Test administrator scribed open-ended responses at student's direction	A test administrator may record word-for-word exactly what a student dictated directly into the PSSA test booklet. This includes MC and OE responses for reading, math and science. For writing, MC items and scribing one's own recorded essay.
Test administrator marked multiple-choice responses at student's direction	A test administrator may mark an answer booklet at the direction of a student. (e.g., a student may point to a multiple-choice answer with the test administrator marking the response in the answer booklet).
Test administrator transcribed (copied) student responses. (per *Accommodations Guidelines*)	For writing prompts the test administrator may transcribe handwriting that is extremely difficult to read. On reading, mathematics, or science illegible handwriting may be transcribed for open-ended items only.
Augmentative communication device	Students with severe communication difficulties may use a special device to convey responses, which must be transcribed into the test booklet by the test administrator.
Typewriter, word processor or computer (per *Accommodations Guidelines*)	An allowable accommodation as a typing function only for students with identified need. Supports such as dictionaries, thesauri, spell checkers and grammar checkers must be turned off. Answers must then be transcribed into the answer booklet without alteration.
Audio recording of student responses (per *Accommodations Guidelines*)	An electronic recording device may be used to record responses, which must be transcribed into the test booklet by the test administrator. [Students who are unable to use a pencil or have illegible handwriting may answer reading, mathematics, and writing multiple-choice questions orally. Answers must be recorded in the answer booklet without alteration during the testing period.]
Manipulative (Cranmer Abacus, number line)	An adaptive calculator or a Cranmer Abacus may be used for the calculator portion of the test only. Eligible students are only those with blindness, low vision, or partial sight.
Translation dictionary for ELL student	A word-to-word dictionary that translates native language to English (or vice versa) without word definitions or pictures is allowed on any portion of the mathematics test and open-ended section of the reading test (but not for the reading passage or multiple-choice items). Cannot be used on any section of the writing test.
Electronic screen reader (PDE approval required)	Students with blindness or extremely low vision may use computer software that converts text to synthesized speech or Braille.
Other (per *Accommodations Guidelines* or PDE approval)	Other accommodations may be appropriate and available if they do not compromise the integrity of the assessment. Documentation must be provided to PDE.

Table 9–12 (continued). Glossary of Accommodations Terms as Applied in the PSSA

Student used the following Setting Accommodations *(mark all that apply)*	
Hospital/home testing	A student who is confined to a hospital or to home during the testing window may be tested in that environment.
Tested in a separate setting	A separate room may be used to reduce distraction.
Small group testing	Some students may require a test setting with fewer students or a setting apart from all other students.
Student used the following Timing Accommodations *(mark all that apply)*	
Scheduled extended time	Extended time may be allotted for each section of the test as a planned accommodation to enable students to finish.
Student-requested extended time	A student may request extended time if working productively.
Multiple test sessions	Multiple test sessions (breaks within a test section) may be scheduled for the completion of each test section; however, a test section must be completed within one school day.
Changed test schedule	Students whose disabilities prevent them from following a regular, planned test schedule may follow an individual schedule, enabling test completion.

Chapter Ten: Form Analysis and Item Calibration

SAMPLING

In order to expedite the analysis process, a target sample of students was selected for use in calibrating items. The target sample was aimed to cover roughly 50% of the student population while preserving ethnic representation. This was done using random sampling without replacement at the district level for approximately 85% of the sample and at the school level for Pittsburgh and Philadelphia districts for approximately 15% of the sample based on 2007–2008 enrollment counts. In any given year, additional students end up being available before the calibration is scheduled to begin. If possible, the additional students are included in the analyses, provided there is time to ensure the demographic profile is similar to the original target sample.

TEST FORM STATISTICS

Table 10–1 contains an overview of the test-level data based on the complete set of common items in each subject area. Test length in total number of points (Pts.), mean number of points received (Mean), standard deviations (SD), test reliabilities (R), and traditional standard errors of measurement (SEM) are shown by grade and content area. These statistics are based on the total test using both multiple-choice and open-ended tasks for the common sections of each form. Both item types (MC and OE) are included in the calculation of total test reliability. Detailed item-level statistics for the common items can be found in Appendices F through G.

Test reliability is discussed in more detail in Chapter Thirteen.

Table 10–1. 2008 Summary of Common Item Performances

Grade	Pts.	Mean	SD	R	SEM
4	66	45.80	11.04	0.91	3.36
8	66	38.25	11.71	0.91	3.59
11	72	36.11	12.46	0.91	3.78

The standard deviations shown in the table are for the standard deviations of observed scores. Assuming normally distributed scores, one would expect about two-thirds of the observations to be within one standard deviation of the mean. An estimate of the standard deviation of the true scores can be computed as $\hat{\sigma}_T = \sqrt{\hat{\sigma}_x^2 - \hat{\sigma}_x^2(1-\hat{\rho}_{xx})}$ .The conditional standard error of measurement (CSEM) also indicates the degree of measurement error in score units, but as a function of one's actual test score. Therefore, the CSEM may be especially useful in characterizing measurement precision in the neighborhood of a score level used for decision-making—such as cut scores for identifying students who meet a performance standard. The CSEMs are documented in Appendix H in the column labeled "Scaled Score SE."

TRADITIONAL ITEM STATISTICS

Although all items were previously reviewed for both content and statistical quality, a thorough item analysis was conducted in the spring to ensure that the items and forms performed as expected. With any psychometric model, an item analysis is a search for unexpected results. For example, *more able*[1] students are expected to pass easy items and *less able* students are expected to fail difficult items. If either of these situations does not occur, the item should be reviewed to determine the nature of the problem and the characteristics of the students affected.

The most familiar indices of item performance are *proportion correct* (P-Value) and item discrimination. Discrimination for dichotomous items is typically represented by the *point-biserial correlation* coefficient. The correlation will have a positive value when the mean score of the students answering correctly is higher than the mean score of the students answering incorrectly. This indicates that students who did well on the total test tended to do well on this item. The index will take its maximum theoretical value of 1.0 if *every* student who answered the item correctly scored better on the test than *any* student who answered incorrectly[2].

The P-Value indicates an item's difficulty for some specified group (e.g. grade level). 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.

Table 10–2 provides some distributional indices for the P-Value and corrected point-biserial correlation (Pt. Bis.) for the multiple-choice items on the common form in each grade and content area.

Table 10–2. Common Form Statistics by Grade for Multiple-Choice Items

	4		8		11	
	P-Value	Pt. Bis.	P-Value	Pt. Bis.	P-Value	Pt. Bis.
Minimum	0.425	0.219	0.256	0.103	0.267	0.132
Maximum	0.944	0.540	0.900	0.536	0.881	0.504
Mean	0.703	-	0.590	-	0.578	-
Median	0.734	0.401	0.606	0.399	0.559	0.345

[1] Following the Rasch literature, *ability* is used in this discussion as a generic term for the construct that is being measured by the exam. *Competence*, *achievement*, *learning* and *status* are among the alternatives that are sometimes used, but are all subject to some degree of misinterpretation.

[2] It is legitimate to view the point biserial correlations as standardized means. A positive value means students who chose that response had a higher mean score than the average student; a negative value means students who chose that response had a lower than average mean score.

RASCH ITEM STATISTICS AND SCALING

WINSTEPS© software, implementing the Rasch model, was used to obtain estimates of logit difficulties for both dichotomously- and polytomously-scored items. The parameters estimated for polytomous items are the *step difficulties* associated with the Masters Partial Credit model. This software is capable of handling all the item types currently in use with the PSSA. WINSTEPS© version 3.54 was used for all calibrations. See Wright and Masters (1982) and Rasch (1960) for further information about the models used for these analyses.

The Rasch model expresses item difficulty (and student ability) in units referred to as *logits*, rather than in percent correct. In the simplest case, a logit is a transformed P-Value with the average P-Value becoming a logit of zero. In this form, logits resemble z-scores or standard normal deviates; a very difficult item might have a logit of +4 and a very easy item might have a logit of –4. However, they have no formal relationship to the normal distribution.

The logit metric has several mathematical advantages over P-Values. It is an interval scale, meaning that two items with logits of zero and one (respectively) are the *same distance* apart as items with logits of +3 and +4. Logits are not dependent on the ability level of the students. For example, a form can have a mean logit of zero, whether the average P-Value of the sample is 0.8 or 0.3.

The standard Rasch calibration procedure arbitrarily sets the mean difficulty of the items on any form at zero. Under normal circumstances where all students are administered a common set of items, any item with a P-Value lower than the average item on the form receives a positive logit difficulty and any item with a P-Value higher than the average receives a negative logit. Consequently, the logits for any calibration, whether it is a grade 3 reading test or a high school science test, relate to an arbitrary origin defined by the center of items on that form. The average grade 3 reading item will have a logit of zero; the average high school science item will have a logit of zero. Logits for both item difficulties and student abilities are placed on the same scale and relate to the same mean item difficulty.

There are any number of other arbitrary choices that could be made for centering the item difficulties. Rather than using all the items, the origin could be defined by a subset. For the PSSA, all test forms in a particular grade and content area share a common block of items. The items on all forms can then be easily adjusted to a single (but still arbitrary) origin by defining the origin as the mean of the **common** items. With this done, the origins for all the forms will be statistically equal. For example, items on any two forms that are equally difficult will now have *statistically* equal logit difficulties.

During PSSA administration, test forms were spiraled within classrooms. In effect, students are administered the same set of common items but different field test or matrix sets. As a result, there are cross checks that are made to ensure the calibrations and links are reasonable across forms. The goal of spiraling is to achieve randomly equivalent samples of students across forms with equal standard deviations and arbitrary means. Any differences in performance observed among the groups should be due only to differences in form difficulty. After linking, the mean of the logit abilities should be statistically equal for each sample of students.

Because of the equivalent samples, common items should have the same P-Values regardless of which form and sample is being considered. Also, for all items, both common and matrix, a plot of the relationship between the P-Value and the logit should fall along a single, curved line. Figures 10–1 through 10–3 plot this relationship for common items. The curves are nearly linear

in the center, but curve towards asymptotes of one and zero, respectively, on the left and right. The graphs show that items with low P-Values (indicating a more difficult item that fewer students answered correctly) also showed higher logit difficulty, and items with high P-Values had lower logit difficulties (i.e., the two scales are inversely related). The spread of the graph points is indicative of the dispersion of item difficulties in the common items. Common OE items are also graphed in Figures 10–1 to 10–3. These items appear with triangle-shaped markers. The OE items generally fall on the same curve as the MC items but subtle differences can occur. The OE items were placed on the MC item difficulty (P-Value) scale—which ranges from 0.00 to 1.00—by dividing the mean OE item score by the maximum OE score possible. Also, the MC items were calibrated concurrently. The OE items were placed on the MC scale in a separate step (i.e., MC items were concurrently calibrated, then anchored by programmatically fixing their values when the difficulties of OE items were estimated).

Figure 10–1. Grade 4

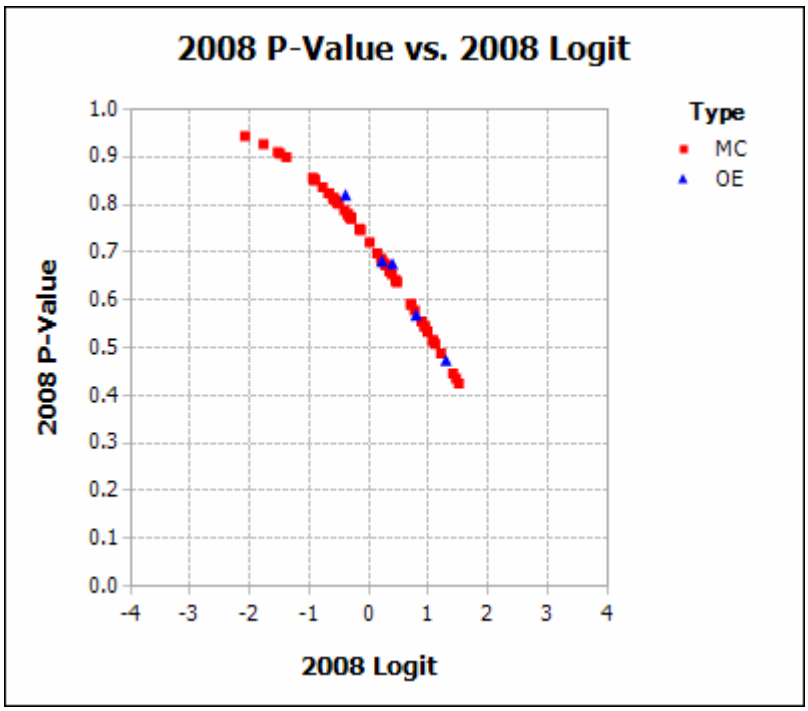


Figure 10–2. Grade 8

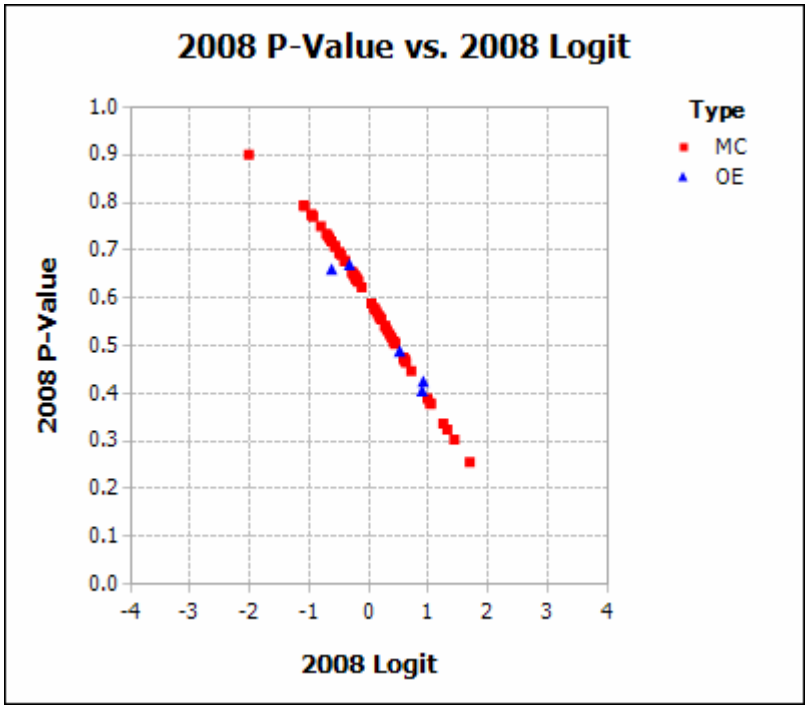
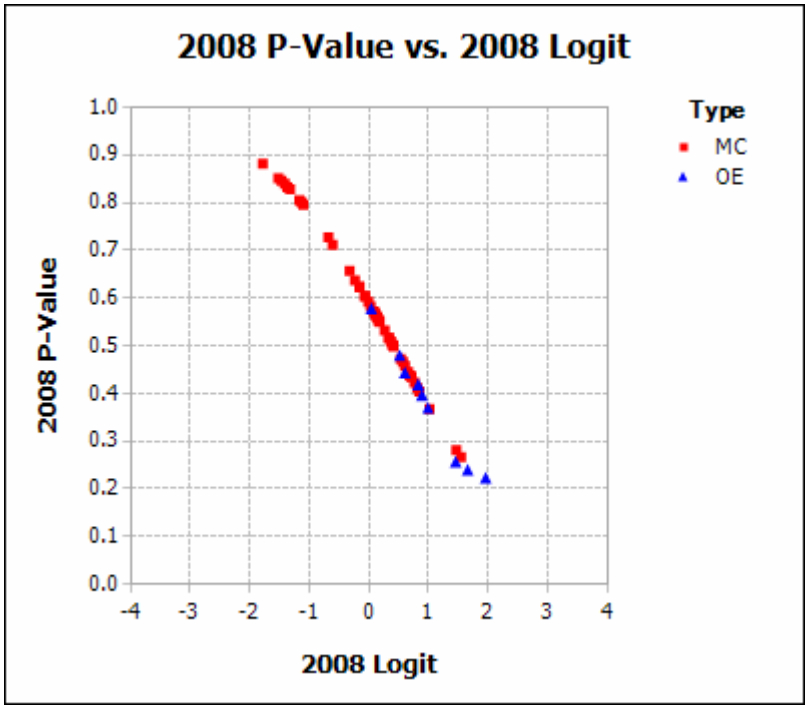


Figure 10–3. Grade 11



Below are the mean raw scores by form for the common items. In addition, the number of students (N), test length in total points (Pts.), minimum observed score (Min), maximum observed score (Max), median score (Median), and standard deviation (SD) are shown by form, grade, and content area. The extent to which the mean raw scores across forms are similar indicates the extent to which the student populations taking each form are of approximately equal ability. This equivalence of ability distributions across forms is the desired outcome of spiraling and allows for optimum analysis of the embedded field test items.

Table 10–3. 2008 Summary Raw-Score Statistics for Common Items by Form

	Grade 4						
Form	**N**	**Pts.**	**Min**	**Max**	**Mean**	**Median**	**SD**
All	126426	66	5	66	45.80	48.0	11.04
1[3]	7163	66	6	66	45.31	47.0	11.20
2	7027	66	8	66	45.65	48.0	11.13
3	7001	66	6	66	45.71	48.0	11.10
4	7048	66	8	65	45.68	48.0	11.01
5	7030	66	7	66	45.76	48.0	10.94
6	7046	66	6	66	45.73	48.0	11.06
7	7012	66	8	66	45.85	48.0	11.01
8	7018	66	8	66	45.88	48.0	10.93
9	7022	66	6	66	45.86	48.0	10.99
10	7024	66	9	66	46.11	48.0	11.03
11	7036	66	6	66	45.92	48.0	11.05
12	7011	66	10	66	45.82	48.0	11.03
13	6981	66	5	66	45.98	48.0	10.94
14	6996	66	7	66	45.99	48.0	11.01
15	7049	66	7	66	45.97	48.0	11.05
16	6981	66	7	66	45.83	48.0	11.17
17	6978	66	8	66	45.47	48.0	11.04
18	7003	66	9	66	45.82	48.0	10.97

[3] In all grades, Form 1 was used to generate modified versions (e.g., Large Print and Braille) of the common form; thus, the mean P-Values for these forms are somewhat lower.

Table 10–3 (continued). 2008 Summary Raw-Score Statistics for Common Item by Form

	Grade 8						
Form	**N**	**Pts.**	**Min**	**Max**	**Mean**	**Median**	**SD**
All	137790	66	1	65	38.25	39.0	11.71
1[3]	7746	66	6	65	37.97	39.0	11.73
2	7660	66	6	65	38.05	39.0	11.84
3	7621	66	4	65	38.29	40.0	11.72
4	7680	66	5	65	38.38	40.0	11.65
5	7657	66	4	64	38.20	39.0	11.81
6	7628	66	4	65	38.11	39.0	11.71
7	7644	66	3	65	38.19	39.0	11.61
8	7635	66	7	64	38.32	39.0	11.61
9	7667	66	6	65	38.38	40.0	11.81
10	7637	66	3	64	38.21	39.0	11.77
11	7657	66	6	65	38.33	40.0	11.70
12	7658	66	7	65	38.31	39.0	11.70
13	7632	66	1	63	38.44	40.0	11.65
14	7641	66	7	64	38.28	39.0	11.76
15	7649	66	6	65	38.32	39.0	11.58
16	7649	66	1	64	38.31	39.0	11.68
17	7660	66	7	65	37.99	39.0	11.69
18	7669	66	8	64	38.35	40.0	11.73

	Grade 11						
Form	**N**	**Pts.**	**Min**	**Max**	**Mean**	**Median**	**SD**
All	131157	72	3	70	36.11	36.0	12.46
1[3]	11042	72	6	68	35.78	36.0	12.37
2	10933	72	3	68	36.11	36.0	12.50
3	10925	72	5	67	36.15	36.0	12.50
4	10927	72	3	69	36.15	36.0	12.35
5	10938	72	5	69	36.04	36.0	12.58
6	10877	72	5	69	36.13	37.0	12.52
7	10877	72	5	69	36.15	36.0	12.36
8	10965	72	3	68	36.18	36.0	12.48
9	10944	72	5	68	36.03	36.0	12.46
10	10945	72	5	68	36.10	36.0	12.50
11	10936	72	5	67	36.21	36.0	12.39
12	10848	72	4	70	36.25	37.0	12.49

Chapter Eleven: Linking

Because the science assessment was newly operational this year, no prior-year linking took place. Linking procedures will be reported in the 2009 technical report after the next operational administration.

Chapter Twelve: Scaled Scores and Performance Levels

Content area total scores for students and schools were based exclusively on the common sections. For each grade level PSSA science test, the Basic and Proficient cut scores were set to 1150 and 1275, respectively. The lowest obtainable scaled scores (LOSS) were set to 1050 at grades 4 and 11, and 925 at grade 8. The LOSS values occur near to what would be considered 'chance' performance (25% of the possible points off of only MC items). The highest obtainable scaled scores (HOSS) will 'float' during each PSSA administration (that is, there is no fixed maximum value). The values of 1150 and 1275 for the Basic and Proficient Cut Scores were selected because they were somewhat similar to the corresponding cuts on other PSSA tests. The resulting multiplicative constants (slopes) of the transformation formulas were also similar to those for the other PSSA tests. Table 12–1 gives the linear transformations that were used to convert 2008 logits (X) into the scaled scores.

Table 12–1. Transformation to Scaled Scores

4	Y = 176.75X + 1225.65
8	Y = 191.54X + 1196.64
11	Y = 101.81X + 1194.69

COMMON ITEMS AND MATRIX SAMPLED ITEMS

As noted earlier, student-level scores were based on the common items only. This ensures that any decision made about students will be done in the most equitable manner. School-level scaled scores for the content areas are based on the mean of the student-level scaled scores. This ensures that the scaled scores used for school accountability directly reflect the student-level results. It is a simple matter to aggregate up to the school, district, and state levels.

For the purpose of providing school-level results at the content standard (academic standards category) level, all items on all matrix forms plus the common items are utilized. This ensures that decisions about potential school-level strengths and weaknesses are based on broad sampling of the curriculum.

SCORES FOR CONTENT STANDARDS

School results are presented as the percent of total points achieved as compared to district and state level results.

Interpreting Scaled Scores and Performance Levels

A *Scaled Score*, in the simplest sense, is a transformed number-correct score[4]. When all students take the same items, as in the common sections of the PSSA, the more points the student earns, the higher the associated scaled score. The value of switching to the more abstract scaled score metric lies in the achievement of a more general and equitable result.

To illustrate, a raw score of 30 is meaningless unless the reader is also told how many points were possible. The same score has quite different meanings if it is based on a thirty-item test as

[4] This is done in two steps. First, there is a nonlinear transformation that converts number correct scores to logits, and then a linear transformation to convert logits to scaled scores.

opposed to a sixty-item test. *Number-correct scores 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 quite different from the same score based on sixty *easy* items. *Number-correct scores are transformed to scaled scores to remove the effects of test length and item difficulty*. As a result, scaled scores lend themselves to interpretations at what is referred to as an interval level, while raw scores do not. Interval-level scales, in the testing industry, allow one to interpret a scaled score difference of 5 points the same whether the scores are 1295 vs. 1300 or 1445 vs. 1450. Raw score differences, in this context, cannot be interpreted in this manner and are thus neither generalizable nor equitable.

If school scores are approximately normally distributed with a given mean and standard deviation, then a scaled score one standard deviation above the base year mean indicates the school did better than about 5/6th of the schools in the base year. About two thirds of the schools will have scaled scores between one SD below the mean and one SD above the mean. About 16% of the schools will be less than one SD below the mean. Scaled scores three standard deviations and more from the mean are considered extreme scores.

Scores such as 1300, 1200, etc. are fairly arbitrary; they could have been called zero and one, or 100 and 110, or any other ordered pair without affecting any of the relationships among schools, years, students, or items. Changing the scale would simply be changing the labels on the axis of a graph without moving any of the points. Like the temperature scales of Fahrenheit and Celsius, the new scale will acquire meaning to users only with experience.

A scaled score of 1300—or any other value for a particular grade and content area test, like grade 4 science—should have the same absolute meaning in the current year as it had in previous years, when test scores are properly equated across years. More importantly, an increase in the scaled score for grade 4 science from one year to another means that student performance improved[5]; it does not say anything about whether the this year's exam is easier or harder than last year's exam. To make these interpretations requires no information about the length or the difficulty of the test in either year, although these variables are essential for the process of deriving the scaled scores.

As noted above, PSSA scaled scores should be interpreted only within each grade and content area test (grade 4 science in the examples above). A student with scaled scores of 1450 in grade 4 science and 1400 in grade 4 mathematics does not necessarily imply that the student performed better in science than mathematics. Further, no across-grade comparisons or growth statements for a student is appropriate. So, a 1400 in grade 4 science and a 1400 in grade 8 science does not indicate a student had no achievement growth from grade 4 to grade 8 in science.

Raw to scaled score tables for the Spring 2008 assessment can be found in Appendix H.

PSSA Performance Levels for Science

Performance levels are another way to attach meaning to the scaled score metric. They 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.

[5] This example is not an endorsement of conducting a trend analysis with just two years of results. Further, small differences may not be statistically or practically significant.

The Pennsylvania General Performance Level Descriptors, as developed by PDE and teacher panels, are given below.

- **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 been reached. 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.

Chapter Thirteen: Test Score Reliability and Validity Evidence

RELIABILITY

This chapter provides reliability indices and standard errors of measurement (SEM). For the Rasch model (assuming adequate model/data fit) raw scores are sufficient statistics for abilities and scaled scores; performance levels set on scaled scores are identical to those based on raw scores.

Reliability Indices

Reliability refers to the expected consistency of test scores. As indicated below, the reliability coefficient expresses the consistency of test scores as the ratio of true score variance to total score variance (true score variance plus error variance). If all test score variance were true, the index would equal 1.0. Conversely, the index will be 0.0 if none of the test score variance were true. Clearly, a larger coefficient is better as it indicates the test scores are influenced less by random sources of error. Generally speaking, reliabilities go up with an increase in test length and population heterogeneity and go down with shorter tests and more homogeneous populations.

$$R = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2} = \frac{\sigma_O^2 - \sigma_e^2}{\sigma_O^2}$$

Although a number of reliability indices exist, a frequently reported index for achievement tests is Coefficient Alpha. Consequently, this index is the one reported for the PSSA. Alpha indicates the internal consistency over the responses to a set of items measuring an underlying trait, in this case science achievement. From this perspective, Alpha can be thought of as the correlation between scores if the students could be tested twice with the same instrument without the second testing being affected by the first. It can also be conceptualized as the extent to which an exchangeable set of items from the same domain would result in similar ordering of students.

While sensitive to random errors associated with content sampling variability, the index is not sensitive to other types of errors that can affect test scores, such as temporal stability or variability in performance that might occur across testing occasions. It is also not sensitive to rater error. Consequently, this index might be positively biased by these factors. On the other hand, there are also factors that might negatively bias this estimate. These include tests that are comprised of mixed item types (e.g., MC and OE items) and tests that include strata (sub-domains) that are homogeneous enough for the average covariance within strata to exceed the average covariance between strata. The reliability coefficient is a unit-free indicator that reflects the degree to which scores are free of measurement error. The *standard error of measurement* (*SEM*) is another indicator of precision. If everyone being tested had the same true score[6], there would still be some variation in observed scores due to imperfections in the measurement process, such as random differences in attention during instruction or concentration during testing. The standard error is defined as the standard deviation[7] of the distribution of observed

[6] True score is the score the person would receive if the measurement process were perfect.

[7] The standard deviation of a distribution is a measure of the dispersion of the observations. For the normal distribution about 16% of the observations are more than one standard deviation above the mean and the same percentage is more than one standard deviation below the mean.

scores for students with identical true scores. Because the SEM is an index of the random variability in test scores in actual score units, it represents important information for test score users.

The Cronbach's Alpha reliability indices (Cronbach, 1951) for the PSSA were calculated. Tables 13–1 through 13–3 provide reliability information on the science tests by strand for the total student population and for students in each gender and ethnicity group. Other groups such as English language learner (ELL), individualized education plan (IEP), and economically disadvantaged (ED) were also included for reliability estimation. The contents of the table include number of points possible (K), number of students tested (N), mean points received (X), standard deviation (SD), mean P-Value (P), reliability (R), traditional standard errors of measurement (SEM), and item type (Types).

As noted earlier, reliabilities tend to go up in value with an increase in test length and population heterogeneity and go down in value with shorter tests and more homogeneous populations. Across the grades and subjects tabled below, reliabilities for the sub-strands tended to follow these same trends. That is, strands with more items tended to show higher reliability coefficients. Also, groups exhibiting more variability in test scores tended to have higher reliability coefficients. Some strand score reliabilities may be too low warrant interpretation at the individual student level. There is no firm guideline regarding how low is too low. The lower a given reliability coefficient, the greater the potential for over-interpretation. As a point of reference, a reliability coefficient of 0.50 would suggest that there is as much error variance as true-score variance in the scores. It should be noted that the reliability of group mean scores (e.g., school or district means) tends to be higher than that of individual scores, meaning interpretation of strand scores at these aggregate levels is likely very reasonable.

Table 13–1. Grade 4

Overall

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	66	126426	45.796	11.038	0.698	0.907	3.363	MC/OE
A	*All*	33	126426	23.151	6.020	0.706	0.849	2.342	MC/OE
B	*All*	11	126426	7.599	2.304	0.678	0.645	1.372	MC/OE
C	*All*	11	126426	7.975	2.031	0.730	0.530	1.392	MC/OE
D	*All*	11	126426	7.072	2.210	0.660	0.595	1.406	MC/OE

Gender

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*Male*	66	64669	46.103	11.307	0.704	0.912	3.346	MC/OE
	Female	66	61571	45.495	10.727	0.692	0.901	3.375	MC/OE
A	*Male*	33	64669	23.153	6.171	0.708	0.857	2.337	MC/OE
	Female	33	61571	23.159	5.851	0.705	0.840	2.342	MC/OE
B	*Male*	11	64669	7.731	2.308	0.692	0.652	1.361	MC/OE
	Female	11	61571	7.464	2.291	0.664	0.636	1.382	MC/OE
C	*Male*	11	64669	7.985	2.048	0.732	0.543	1.384	MC/OE
	Female	11	61571	7.967	2.010	0.727	0.517	1.397	MC/OE
D	*Male*	11	64669	7.234	2.254	0.675	0.619	1.390	MC/OE
	Female	11	61571	6.905	2.150	0.644	0.564	1.420	MC/OE

Table 13–1 (continued). Grade 4

Ethnicity

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*White*	66	92148	48.328	9.615	0.736	0.885	3.257	MC/OE
	Black	66	19631	37.213	11.186	0.567	0.892	3.673	MC/OE
	Hispanic	66	9556	38.293	11.364	0.584	0.898	3.632	MC/OE
	Asian	66	3716	48.796	10.145	0.744	0.900	3.209	MC/OE
	Am. Indian	66	200	43.900	11.132	0.670	0.903	3.471	MC/OE
	Multi	66	950	43.167	11.620	0.658	0.910	3.478	MC/OE
A	*White*	33	92148	24.441	5.340	0.746	0.820	2.265	MC/OE
	Black	33	19631	18.711	6.100	0.571	0.822	2.573	MC/OE
	Hispanic	33	9556	19.365	6.166	0.591	0.830	2.543	MC/OE
	Asian	33	3716	24.957	5.495	0.761	0.839	2.205	MC/OE
	Am. Indian	33	200	22.030	5.870	0.675	0.829	2.427	MC/OE
	Multi	33	950	21.809	6.305	0.666	0.852	2.424	MC/OE
B	*White*	11	92148	8.061	2.064	0.719	0.594	1.315	MC/OE
	Black	11	19631	6.078	2.391	0.542	0.595	1.522	MC/OE
	Hispanic	11	9556	6.222	2.431	0.555	0.617	1.505	MC/OE
	Asian	11	3716	7.926	2.205	0.706	0.644	1.315	MC/OE
	Am. Indian	11	200	7.315	2.422	0.653	0.662	1.408	MC/OE
	Multi	11	950	7.099	2.389	0.634	0.633	1.448	MC/OE
C	*White*	11	92148	8.322	1.832	0.761	0.462	1.344	MC/OE
	Black	11	19631	6.846	2.233	0.627	0.529	1.532	MC/OE
	Hispanic	11	9556	6.868	2.196	0.632	0.528	1.508	MC/OE
	Asian	11	3716	8.339	1.866	0.768	0.475	1.352	MC/OE
	Am. Indian	11	200	7.785	2.210	0.713	0.594	1.408	MC/OE
	Multi	11	950	7.625	2.155	0.700	0.549	1.446	MC/OE
D	*White*	11	92148	7.505	2.018	0.700	0.532	1.380	MC/OE
	Black	11	19631	5.578	2.192	0.522	0.545	1.478	MC/OE
	Hispanic	11	9556	5.837	2.249	0.546	0.573	1.469	MC/OE
	Asian	11	3716	7.575	2.055	0.705	0.558	1.366	MC/OE
	Am. Indian	11	200	6.770	2.272	0.630	0.589	1.456	MC/OE
	Multi	11	950	6.634	2.240	0.619	0.593	1.430	MC/OE

ELL

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	66	3309	33.458	10.794	0.512	0.881	3.719	MC/OE
A	*All*	33	3309	16.974	5.944	0.518	0.808	2.605	MC/OE
B	*All*	11	3309	5.265	2.333	0.472	0.554	1.557	MC/OE
C	*All*	11	3309	6.113	2.198	0.565	0.503	1.549	MC/OE
D	*All*	11	3309	5.107	2.122	0.479	0.508	1.487	MC/OE

IEP

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	66	20238	38.936	12.052	0.594	0.910	3.606	MC/OE
A	*All*	33	20238	19.322	6.520	0.590	0.849	2.530	MC/OE
B	*All*	11	20238	6.558	2.539	0.587	0.657	1.487	MC/OE
C	*All*	11	20238	7.046	2.244	0.645	0.559	1.491	MC/OE
D	*All*	11	20238	6.010	2.332	0.563	0.607	1.462	MC/OE

Table 13–1 (continued). Grade 4

ED

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	66	47988	40.357	11.347	0.615	0.900	3.582	MC/OE
A	*All*	33	47988	20.304	6.159	0.620	0.834	2.507	MC/OE
B	*All*	11	47988	6.679	2.419	0.595	0.629	1.474	MC/OE
C	*All*	11	47988	7.223	2.165	0.662	0.528	1.487	MC/OE
D	*All*	11	47988	6.151	2.247	0.575	0.577	1.461	MC/OE

Table 13–2. Grade 8

Overall

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	66	137790	38.246	11.708	0.584	0.906	3.594	MC/OE
A	*All*	33	137790	19.596	6.195	0.597	0.832	2.538	MC/OE
B	*All*	11	137790	6.438	2.638	0.595	0.695	1.457	MC/OE
C	*All*	11	137790	5.901	2.349	0.550	0.577	1.528	MC/OE
D	*All*	11	137790	6.310	2.042	0.565	0.519	1.416	MC/OE

Gender

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*Male*	66	70667	38.505	12.206	0.588	0.914	3.574	MC/OE
	Female	66	66857	38.015	11.134	0.579	0.895	3.606	MC/OE
A	*Male*	33	70667	19.583	6.447	0.597	0.846	2.529	MC/OE
	Female	33	66857	19.632	5.909	0.597	0.815	2.541	MC/OE
B	*Male*	11	70667	6.432	2.681	0.594	0.707	1.451	MC/OE
	Female	11	66857	6.452	2.588	0.597	0.681	1.461	MC/OE
C	*Male*	11	70667	6.037	2.427	0.563	0.608	1.519	MC/OE
	Female	11	66857	5.765	2.252	0.537	0.535	1.535	MC/OE
D	*Male*	11	70667	6.452	2.089	0.581	0.552	1.398	MC/OE
	Female	11	66857	6.166	1.977	0.549	0.478	1.429	MC/OE

Table 13–2 (continued). Grade 8

Ethnicity

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*White*	66	102757	40.786	10.745	0.622	0.891	3.547	MC/OE
	Black	66	21108	29.042	10.216	0.444	0.868	3.712	MC/OE
	Hispanic	66	9063	30.188	10.724	0.461	0.880	3.708	MC/OE
	Asian	66	3576	42.238	11.102	0.645	0.901	3.497	MC/OE
	Am. Indian	66	239	36.444	11.998	0.556	0.908	3.633	MC/OE
	Multi	66	698	33.317	11.885	0.509	0.905	3.664	MC/OE
A	*White*	33	102757	20.833	5.739	0.635	0.810	2.499	MC/OE
	Black	33	21108	15.099	5.601	0.458	0.775	2.659	MC/OE
	Hispanic	33	9063	15.605	5.788	0.473	0.790	2.651	MC/OE
	Asian	33	3576	21.859	5.893	0.666	0.826	2.461	MC/OE
	Am. Indian	33	239	18.573	6.398	0.565	0.839	2.571	MC/OE
	Multi	33	698	17.129	6.355	0.521	0.831	2.611	MC/OE
B	*White*	11	102757	6.939	2.490	0.639	0.668	1.434	MC/OE
	Black	11	21108	4.638	2.334	0.437	0.588	1.499	MC/OE
	Hispanic	11	9063	4.859	2.443	0.456	0.624	1.498	MC/OE
	Asian	11	3576	7.101	2.531	0.659	0.693	1.404	MC/OE
	Am. Indian	11	239	6.146	2.519	0.569	0.645	1.500	MC/OE
	Multi	11	698	5.500	2.660	0.512	0.687	1.489	MC/OE
C	*White*	11	102757	6.336	2.217	0.589	0.530	1.520	MC/OE
	Black	11	21108	4.333	2.120	0.410	0.484	1.523	MC/OE
	Hispanic	11	9063	4.528	2.183	0.427	0.508	1.530	MC/OE
	Asian	11	3576	6.505	2.261	0.601	0.553	1.511	MC/OE
	Am. Indian	11	239	5.678	2.403	0.529	0.592	1.534	MC/OE
	Multi	11	698	5.076	2.330	0.476	0.568	1.531	MC/OE
D	*White*	11	102757	6.678	1.914	0.599	0.468	1.396	MC/OE
	Black	11	21108	4.972	1.931	0.443	0.424	1.466	MC/OE
	Hispanic	11	9063	5.196	1.982	0.463	0.452	1.466	MC/OE
	Asian	11	3576	6.773	1.928	0.608	0.486	1.381	MC/OE
	Am. Indian	11	239	6.046	2.144	0.540	0.564	1.417	MC/OE
	Multi	11	698	5.612	2.020	0.501	0.494	1.437	MC/OE

ELL

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	66	2292	24.846	8.923	0.382	0.828	3.705	MC/OE
A	*All*	33	2292	12.997	5.082	0.395	0.724	2.669	MC/OE
B	*All*	11	2292	3.727	2.007	0.357	0.461	1.474	MC/OE
C	*All*	11	2292	3.711	1.935	0.355	0.400	1.498	MC/OE
D	*All*	11	2292	4.411	1.850	0.392	0.355	1.485	MC/OE

Table 13–2 (continued). Grade 8

IEP

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	66	21169	27.860	10.484	0.426	0.875	3.713	MC/OE
A	*All*	33	21169	14.203	5.632	0.431	0.778	2.654	MC/OE
B	*All*	11	21169	4.452	2.359	0.416	0.590	1.510	MC/OE
C	*All*	11	21169	4.212	2.168	0.398	0.507	1.522	MC/OE
D	*All*	11	21169	4.993	2.022	0.446	0.469	1.474	MC/OE

ED

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	66	46698	31.935	11.030	0.488	0.888	3.699	MC/OE
A	*All*	33	46698	16.475	5.938	0.500	0.803	2.634	MC/OE
B	*All*	11	46698	5.197	2.503	0.485	0.639	1.503	MC/OE
C	*All*	11	46698	4.834	2.243	0.455	0.530	1.537	MC/OE
D	*All*	11	46698	5.427	2.000	0.484	0.469	1.458	MC/OE

Table 13–3. Grade 11

Overall

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	72	131157	36.107	12.459	0.531	0.908	3.779	MC/OE
A	*All*	36	131157	18.417	6.183	0.547	0.822	2.609	MC/OE
B	*All*	10	131157	5.139	2.045	0.525	0.516	1.423	MC/OE
C	*All*	12	131157	6.168	2.774	0.540	0.660	1.618	MC/OE
D	*All*	14	131157	6.382	2.948	0.485	0.671	1.690	MC/OE

Gender

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*Male*	72	65937	36.621	13.124	0.538	0.918	3.760	MC/OE
	Female	72	64922	35.627	11.712	0.525	0.896	3.786	MC/OE
A	*Male*	36	65937	18.537	6.472	0.550	0.839	2.594	MC/OE
	Female	36	64922	18.317	5.867	0.545	0.801	2.615	MC/OE
B	*Male*	10	65937	5.196	2.110	0.531	0.552	1.412	MC/OE
	Female	10	64922	5.086	1.973	0.519	0.473	1.432	MC/OE
C	*Male*	12	65937	6.293	2.873	0.551	0.688	1.604	MC/OE
	Female	12	64922	6.049	2.663	0.529	0.626	1.628	MC/OE
D	*Male*	14	65937	6.595	3.079	0.499	0.698	1.691	MC/OE
	Female	14	64922	6.175	2.791	0.472	0.638	1.680	MC/OE

Table 13–3 (continued). Grade 11

Ethnicity

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*White*	72	105228	38.102	11.801	0.558	0.897	3.782	MC/OE
	Black	72	15262	25.677	10.109	0.394	0.870	3.647	MC/OE
	Hispanic	72	6314	27.431	11.056	0.417	0.889	3.688	MC/OE
	Asian	72	3312	39.360	13.333	0.577	0.921	3.749	MC/OE
	Am. Indian	72	205	33.024	12.760	0.490	0.913	3.758	MC/OE
	Multi	72	482	29.996	11.965	0.449	0.903	3.721	MC/OE
A	*White*	36	105228	19.372	5.850	0.573	0.801	2.607	MC/OE
	Black	36	15262	13.437	5.312	0.414	0.770	2.548	MC/OE
	Hispanic	36	6314	14.296	5.673	0.436	0.794	2.574	MC/OE
	Asian	36	3312	19.865	6.537	0.589	0.842	2.597	MC/OE
	Am. Indian	36	205	17.141	6.305	0.511	0.828	2.617	MC/OE
	Multi	36	482	15.521	6.102	0.467	0.821	2.583	MC/OE
B	*White*	10	105228	5.380	1.982	0.548	0.489	1.416	MC/OE
	Black	10	15262	3.892	1.856	0.405	0.410	1.426	MC/OE
	Hispanic	10	6314	4.058	1.949	0.421	0.469	1.420	MC/OE
	Asian	10	3312	5.545	2.162	0.567	0.576	1.408	MC/OE
	Am. Indian	10	205	4.517	2.255	0.464	0.627	1.377	MC/OE
	Multi	10	482	4.461	2.084	0.460	0.535	1.422	MC/OE
C	*White*	12	105228	6.546	2.683	0.570	0.634	1.624	MC/OE
	Black	12	15262	4.161	2.288	0.383	0.555	1.527	MC/OE
	Hispanic	12	6314	4.476	2.456	0.407	0.599	1.555	MC/OE
	Asian	12	3312	7.012	2.942	0.608	0.715	1.570	MC/OE
	Am. Indian	12	205	5.527	2.575	0.493	0.601	1.627	MC/OE
	Multi	12	482	5.064	2.587	0.455	0.629	1.576	MC/OE
D	*White*	14	105228	6.804	2.855	0.514	0.646	1.698	MC/OE
	Black	14	15262	4.186	2.368	0.339	0.560	1.571	MC/OE
	Hispanic	14	6314	4.601	2.569	0.367	0.610	1.605	MC/OE
	Asian	14	3312	6.937	3.084	0.523	0.698	1.694	MC/OE
	Am. Indian	14	205	5.839	3.011	0.450	0.693	1.669	MC/OE
	Multi	14	482	4.950	2.724	0.386	0.637	1.641	MC/OE

ELL

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	72	1493	22.279	8.882	0.347	0.839	3.561	MC/OE
A	*All*	36	1493	11.562	4.714	0.363	0.722	2.484	MC/OE
B	*All*	10	1493	3.298	1.756	0.350	0.378	1.385	MC/OE
C	*All*	12	1493	3.670	2.207	0.340	0.543	1.492	MC/OE
D	*All*	14	1493	3.748	2.045	0.309	0.431	1.543	MC/OE

IEP

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	72	17239	24.930	10.204	0.380	0.873	3.638	MC/OE
A	*All*	36	17239	12.978	5.340	0.396	0.770	2.559	MC/OE
B	*All*	10	17239	3.720	1.845	0.386	0.408	1.419	MC/OE
C	*All*	12	17239	3.982	2.258	0.368	0.564	1.490	MC/OE
D	*All*	14	17239	4.250	2.432	0.342	0.580	1.576	MC/OE

Table 13–3 (continued). Grade 11

ED

Strand	Group	K	N	X	SD	P	R	SEM	Types
Overall	*All*	72	31773	29.245	11.364	0.440	0.892	3.727	MC/OE
A	*All*	36	31773	15.148	5.770	0.459	0.798	2.591	MC/OE
B	*All*	10	31773	4.268	1.960	0.440	0.468	1.430	MC/OE
C	*All*	12	31773	4.824	2.524	0.435	0.610	1.577	MC/OE
D	*All*	14	31773	5.005	2.704	0.393	0.634	1.635	MC/OE

Decision Consistency

In a standards-based testing program there is also interest in knowing how accurately students are classified into the various performance categories. Classification consistency refers to the degree with which the achievement level for each student can be replicated upon retesting using the same form or an equivalent form (Huynh, 1976). Since it is not feasible to repeat PSSA testing in order to estimate the proportion of students who would be reclassified in the same performance levels, a statistical model needs to be imposed on the data in order to project the consistency of classifications solely using data from the available administration (Hambleton and Novick, 1973). Although a number of procedures are available, two well known methods were developed by Hanson and Brennan (1990) and Livingston and Lewis (1995) utilizing specific True Score Models.

Hanson and Brennan (1990) utilized a four parameter beta binomial and a four parameter beta compound binomial model for estimating single administration estimates of classification consistency. The models are given by:

$$\Pr(X = i) = \int_{l}^{u} \Pr(X = i \mid \tau, k) g(\tau \mid \alpha, \beta, l, u) d\tau ,$$

where l and u are, respectively, the lower and upper bounds of the distribution. If $k \le 0$, then the conditional error distribution is binomial. If $k > 0$, then the conditional error distribution is compound binomial using Lord's (1965) two-term approximation to the compound binomial distribution. Parameters for the true score density are estimated using the method of moments.

In order to use this method, the test must consist of purely dichotomous items. A simple way to satisfy this requirement is to dichotomize the assessment. For any polytomous item with a maximum score of u, create a set S of u dichotomous items to replace it. For example, $S_{1(x)} = \{u_1, u_2, u_3, u_4\}$ is a set created for a polytomous item with a maximum score of 4 for examinee x. Then, for an examinee y with a score of 3, $S_{1(y)} = \{1,1,1,0\}$. Local independence of these newly created dichotomous items within the set is sacrificed due to the fact that to get the 3rd item in the set correct, you must get the 1st and 2nd items in the same set correct (the same goes for getting an item incorrect). Artificial local independence cannot be manufactured for these items within each set.

To solve the problem of a complex assessment, Livingston and Lewis (1995) proposed an effective test length,

$$n = \frac{(\mu_x - X_{\min})(X_{\max} - \mu_x) - r\sigma_x^2}{\sigma_x^2(1 - r)} ,$$

which transforms the original raw score random variable from $X = 0,...,K$ into a new random variable $X' = 0,...,n$, where n is the number of dichotomous, locally independent, equally difficult items required to produce a raw score of the same reliability. Then, using the transformed observed distribution X', parameters are estimated for a four parameter beta-binomial model where the conditional error distribution is assumed to be binomial. The X' distribution is then converted back onto the original X scale using interpolation. This method is designed only to estimate a contingency table, not a full bivariate distribution.

Stearns and Smith (2007) found that results from the Hanson and Brennan (1990) method on a dichotomized version of a complex assessment yields similar results to the Livingston and Lewis (1995) method. The results of the consistency analyses by cut score and overall are presented in Table 13–4. The results—derived using the program *BB-Class* (Brennan, 2004)—showed that the consistency index values across methods were very similar. It should be noted that consistency indices for the four performance levels should be lower than those based on two categories, as seen below. This is not surprising since classification using four levels would allow more opportunity to change the achievement levels. Hence there would be more classification errors in the four achievement levels, resulting in lower consistency indices.

Table 13–4. Decision Consistency

Grade	Method	Overall	BB/B	B/P	P/A
4	*H.B.(1990)*	0.751	0.963	0.920	0.865
	L.L.(1995)	0.747	0.962	0.918	0.862
8	*H.B.(1990)*	0.684	0.904	0.878	0.898
	L.L.(1995)	0.683	0.904	0.878	0.897
11	*H.B.(1990)*	0.719	0.909	0.879	0.929
	L.L.(1995)	0.719	0.908	0.879	0.929

Rater Agreement

Rater Agreement information is provided in Chapter Eight.

VALIDITY

As noted in the *Standards for Educational and Psychological Testing*, "validity refers to the degree to which evidence and theory support the interpretation of test scores entailed by the proposed uses of the tests" (AERA, APA, & NCME, 1999, p. 9). Thus, the validity of the PSSA must be judged in relation to its primary purposes as delineated in Chapter One. Multiple sources of evidence should be considered, and in fact, virtually all of this technical report's content is part of this body of evidence (test developmental steps, item writing procedures, item review, item analysis, bias analysis, etc.). Below, validity evidence specifically related to test content is presented in terms of how the PSSA assessments were assembled to reflect the state content standards. Very significant supporting information on the item development processes is presented in Chapter Three.

Content Validity Evidence

PDE commissioned Achieve, Inc. to conduct two reviews (2003 and 2005) during the period in which PDE was in the process of developing and refining the Assessment Anchor Content

Standards (Assessment Anchors) and Eligible Content for science. Through an iterative process of successive refinement in which each version underwent review and modification, final documents for science emerged. Similarly, PDE submitted sets of items designed to measure these anchors (see Chapter Two for additional details on science anchors). The item development process also benefited from an evaluation of how well test items aligned with the Assessment Anchors and Eligible Content. The reviews conducted by Achieve (2007) focused on:

- Assessment Anchors and Eligible Content for science.
- Alignment of assessments to the Assessment Anchors and Eligible Content and, subsequently, in developing items tailored toward these anchors.

Achieve, Inc. (2007 draft). *Report for Pennsylvania: Grade 4, 8, & 11—Alignment of Test Items to Assessment Anchors and Eligible Content Statements*.

Internal Validity Evidence

Information regarding the PSSA's internal structure is an important source of construct-related evidence of validity. Correlations by reporting category (A–D) were calculated using Pearson's Correlation Coefficient. These correlations are presented in Tables 13–5 through 13–7.

Table 13–5. Grade 4 Correlations (Reliabilities on Diagonal)

S	0.907				
SA	0.959	0.849			
SB	0.829	0.723	0.645		
SC	0.767	0.657	0.559	0.530	
SD	0.814	0.707	0.617	0.540	0.595
	S	**SA**	**SB**	**SC**	**SD**

Table 13–6. Grade 8 Correlations (Reliabilities on Diagonal)

S	0.906				
SA	0.958	0.832			
SB	0.862	0.763	0.695		
SC	0.819	0.709	0.640	0.577	
SD	0.772	0.660	0.600	0.567	0.519
	S	**SA**	**SB**	**SC**	**SD**

Table 13–7. Grade 11 Correlations (Reliabilities on Diagonal)

S	0.908				
SA	0.955	0.822			
SB	0.780	0.673	0.516		
SC	0.865	0.761	0.612	0.660	
SD	0.868	0.757	0.614	0.695	0.671
	S	**SA**	**SB**	**SC**	**SD**

VALIDITY EVIDENCE SUMMARY

As noted earlier, "validity refers to the degree to which evidence and theory support the interpretation of test scores entailed by the proposed uses of the tests" (AERA, APA, & NCME, 1999, p. 9). Further, it was stated that validity considerations must be judged in relation to the PSSA's primary purposes (overviewed in Chapter One). While this entire technical document provides information that can (and should) be weighed when establishing the body of validity

evidence, this summary is structured based primarily on two of the five sources of validity evidence identified in the *Standards for Educational and Psychological Testing*, that is, evidence based on test content and internal structure.

Validity evidence related to test content was presented earlier in this chapter. Further, details about how the 2008 PSSA assessments were assembled to reflect the state content standards and information regarding educator reviews (including content and bias and sensitivity reviews) was presented in Chapter Three. On the whole, the early chapters of this technical report show that a strong link can be established between each PSSA item and its associated eligible content.

Strand score intercorrelations were also presented in this chapter. In general one finds that the science strands were strongly correlated with each other. This is provides some favorable evidence regarding the internal relationships among the test's components.

Validity of score inferences is bolstered when test scores are consistent. Here, the reliability of the total test scores were very good.

Reported in Chapter Five, differential item functioning (DIF) with respect to gender and ethnicity helps address construct-irrelevant variance, which represents an important threat to the validity of inferences made from achievement test scores. As noted in that chapter, field test items are screened and reviewed for DIF. Only items approved by teacher committees are eligible for operational use.

References

Achieve, Inc. (2007 draft). *Report for Pennsylvania: Grade 4, 8, & 11—Alignment of Test Items to Assessment Anchors and Eligible Content Statements*.

AERA, APA, NCME (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education). (1999). *Standards for Educational and Psychological Tests*. Washington, DC: American Educational Research Association.

Allman, C. (2004). Test access: Making tests accessible for students with visual impairments – A guide for test publishers, test developers, and state assessment personnel (2nd edition). Louisville, KY: American Printing House for the Blind. Available from http://www.aph.org.

Brennan, R. 2004. *BB-Class (Version 1.0)*. CASMA: education.uiowa.edu/casma. Computer program.

Cronbach, L. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 16, 297-334.

Data Recognition Corporation. (2000). *Item Viewer and Authoring Network (IVAN): Informational Guide*. Maple Grove, MN: DRC.

Data Recognition Corporation. (2003–2007). *Fairnes*s *in Testing: Training Manual for Issues of Bias, Fairness, and Sensitivity*. Maple Grove, MN: DRC.

Data Recognition Corporation. (2004–2007). *Pennsylvania System of School Assessment (PSSA) Style Guide.* Maple Grove, MN: DRC.

Data Recognition Corporation. (2005, December). *Technical Report for the PSSA 2005 Reading and Mathematics*. Maple Grove, MN: DRC.

Data Recognition Corporation. (2007, May). *Technical Report for the PSSA 2006 Reading and Mathematics: Grades 4, 6, and 7*. Maple Grove, MN: DRC.

Data Recognition Corporation. (2007, May). *Technical Report for the PSSA 2006 Writing: Grades 5, 8, and 11*. Maple Grove, MN: DRC.

Data Recognition Corporation. (2007). *Preliminary Technical Report for 2008 PSSA Science*. Maple Grove, MN: DRC.

Data Recognition Corporation. (2008, February). *Technical Report for the PSSA 2007 Reading and Mathematics: Grades 3, 4, 5, 6, 7, 8, and 11*. Maple Grove, MN: DRC.

Data Recognition Corporation (2008, February) *Preliminary Technical Report for 2008 PSSA Science,* Maple Grove, MN : DRC.

Dorans, N., Schmitt, Al, & Bleistein, C. (1992). The standardization approach to assessing comprehensive differential item functioning. *Journal of Educational Measurement*, 29, 309–319.

Hambleton, R. & Novick, M. (1973). "Toward an Integration of Theory and Method for Criterion-Referenced Tests." *Journal of Educational Measurement*, 10, 159–170.

Hanson, B. A., & Brennan, R. L. (1990). An Investigation of Classification Consistency Indexes Estimated Under Alternative Strong True Score Theory Models. *Journal of Educational Measurement, 27(4),* 345–359.

Huynh, H. (1976). "On the Reliability of Decisions in Domain-Referenced Testing." *Journal of Educational Measurement*, 13, 253–264.

Lewis, D.M., Mitzel, H.C. & Green, D.R. (1996). *Standard Setting: A Bookmark Approach*. Symposium presented at the Council of Chief State School Officers National Conference on Large-Scale Assessment: Phoenix, AZ.

Livingston, S. & Lewis, C. (1995). "Estimating the Consistency and Accuracy of Classifications Based on Test Scores." *Journal of Educational Measurement* 32, 179–197.

Mantel, N. & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719–748.

No Child Left Behind Act of 2001, Pub. L. No. 107–110, 115 Stat. 1425 (2002).

Pennsylvania State Board of Education. (1999, January). *Chapter 4. Academic Standards and Assessment*. Harrisburg, PA: Pennsylvania State Board of Education. Retrieved November 8, 2004 from http://www.pde.state.pa. Also available from http://www.pacode.com/secure/data/022/Chapter4/s4.51.html.

Pennsylvania Department of Education. (2004). *Mathematics Item and Scoring Sampler*. Harrisburg, PA: PDE. Posted separately by grade level. Retrieved December 13, 2004 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2004). *Reading Item and Scoring Sampler*. Harrisburg, PA: PDE. Posted separately by grade level. Retrieved December 13, 2004 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2004, April). *Assessment Anchors and Eligible Content*. Harrisburg, PA: PDE. Retrieved December 13, 2004 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2004, November). *Mathematics Assessment Handbook*. Harrisburg, PA: PDE. Retrieved December 13, 2004 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2004, November). *Reading Assessment Handbook*. Harrisburg, PA: PDE. Retrieved December 13, 2004 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2005, December). *2005–2006 Mathematics Assessment Handbook*. Harrisburg, PA: PDE. Retrieved January 30, 2006 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2005, December). *2005–2006 Reading Assessment Handbook*. Harrisburg, PA: PDE. Retrieved January 30, 2006 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2005). *2005–2006 Mathematics Item and Scoring Sampler*. Harrisburg, PA: PDE. Posted separately by grade level. Retrieved January 30, 2006 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2005). *2005–2006 Reading Item and Scoring Sampler*. Harrisburg, PA: PDE. Posted separately by grade level. Retrieved January 30, 2006 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2005, December). *2005–2006 Writing Assessment Handbook*. Harrisburg, PA: PDE. Retrieved January 30, 2006 from http://www.pde.state.pa.http://www.pde.state.pa.

Pennsylvania Department of Education. (2005). *2005–2006 Writing Item and Scoring Sampler*. Harrisburg, PA: PDE. Posted separately by grade level. Retrieved September 14, 2005 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2006). *2006–2007 Mathematics Item and Scoring Sampler*. Harrisburg, PA: PDE. Posted separately by grade level. Retrieved January 30, 2007 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2006). *2006–2007 Reading Item and Scoring Sampler*. Harrisburg, PA: PDE. Posted separately by grade level. Retrieved January 30, 2007 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2006). *2006–2007 Writing Item and Scoring Sampler*. Harrisburg, PA: PDE. Posted separately by grade level. Retrieved January 30, 2007 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2006, December). *2006–2007 Writing Assessment Handbook*. Harrisburg, PA: PDE. Retrieved January 30, 2006 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2006). *2006–2007 Science Item and Scoring Sampler*. Harrisburg, PA: PDE. Posted separately by grade level. Retrieved March 15, 2007 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2006, November). Science *Assessment Handbook*. Harrisburg, PA: PDE. Retrieved March 15, 2007 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2007, January). *2006–2007 Mathematics Assessment Handbook*. Harrisburg, PA: PDE. Retrieved January 30, 2007 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2007, January). *2006–2007 Reading Assessment Handbook*. Harrisburg, PA: PDE. Retrieved January 30, 2007 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2007, January). *2007 Accommodations Guidelines for Students with IEPs, Students with 504 Plans, English Language Learners, and All Students*. Harrisburg, PA: PDE. Retrieved January 30, 2007 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2007). *PSSA 2007 Handbook for Assessment Coordinators and Administrators: Grades 3–8 and 11 Reading and Mathematics*. Harrisburg, PA: PDE. Retrieved January 30, 2007 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2007, March). *PSSA Reading and Mathematics Directions for Administration Manual*. Harrisburg, PA: PDE. Posted separately by grade level. Retrieved April 2, 2007 from http://www.pde.state.pa.

Pennsylvania Department of Education. (2007). *2008 PSSA Accommodations Guidelines for Students with IEPs and Students with 504 Plans*. Harrisburg, PA: PDE. Retrieved March 4, 2008, from http://www.pde.state.pa.us.

Pennsylvania Department of Education. (2008). *PSSA 2008 Handbook for Assessment Coordinators and Administrators: Grades 4, 8, and 11 Science*. Harrisburg, PA: Retrieved March 4, 2008, from http://www.pde.state.pa.us.

Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen: Danish Institute for Educational Research.

Stearns, M., & Smith R. M. (2007). *Estimation of classification consistency indices for complex assessments: Model based approaches*. Paper presented at the 2007 Annual Convention of the American Educational Research Association. Chicago, IL.

Thacker, A. A. & Dickinson, E. R. (2004). *Item Content and Difficulty Mapping by Form and Item Type for the 2001-2003 Pennsylvania System of School Assessment (PSSA)*. Alexandria, VA: Human Resources Research Organization.

Thacker, A. A., Dickinson, E. R. & Koger, M. E. (2004). *Relationships Among the Pennsylvania System of School Assessment (PSSA) and Other Commonly Administered Assessments*. Alexandria, VA: Human Resources Research Organization.

Thompson, S., Johnstone, C. J. & Thurlow, M. L. (2002). *Universal Design Applied to Large Scale Assessments* (Synthesis Report 44), Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Webb, N. L. (1999). *Research Monograph No. 18: Alignment of Science and Mathematics Standards and Assessments in Four States.* Madison, WI: National Institute for Science Education.

WINSTEPS (2000). *WINSTEPS® Rasch Measurement.* Copyright John M. Linacre.

Wright, B. & Masters, G. (1982). *Rating Scale Analysis*. Chicago: MESA Press.

Appendix A:

PSSA Science Scoring Guidelines

DESCRIPTION OF SCORING GUIDELINES FOR 4-POINT OPEN-ENDED ITEMS:

General Description of Science Scoring Guidelines:

4 – The response demonstrates a *thorough* understanding of the scientific content, concepts, and procedures required by the task/s.

The response provides a clear, complete, and correct response as required by the task/s. Response may contain a minor blemish (e.g., misspelled words) or omission in work or explanation that does not detract from demonstrating a thorough understanding.

3 – The response demonstrates a general understanding of the scientific content, concepts, and procedures required by the task/s.

The responses, as required by the task, are mostly complete and correct. The response may have minor errors or omissions that do not detract from demonstrating a general understanding.

2 – The response demonstrates a partial understanding of the scientific content, concepts, and procedures required by the task/s.

The response is somewhat correct with partial understanding of the required scientific content, concepts, and/or procedures demonstrated and/or explained. The response may contain some work that is incomplete or unclear.

1 – The response demonstrates a *minimal* understanding of the scientific content, concepts, and procedures as required by the task/s.

0 – The response provides insufficient evidence to demonstrate any understanding of the scientific content, concepts, and procedures as required by the task/s for that grade level.

Response may show only information copied or rephrased from the question or insufficient correct information to receive a score of 1.

Special Categories within zero reported separately:
BLK – Blank, entirely erased or written refusal to respond
OT – Off Task
IL – Illegible
LOE – Response in a language other than English

DESCRIPTION OF SCORING GUIDELINES FOR 2-POINT OPEN-ENDED ITEMS:

General Description of Science Scoring Guidelines:

2 – The response demonstrates a *thorough* understanding of the scientific content, concepts, and procedures required by the task/s.

The response provides a clear, complete, and correct response as required by the task/s. Response may contain a minor blemish (e.g., misspelled words) or omission in work or explanation that does not detract from demonstrating a thorough understanding.

1 – The response demonstrates a partial understanding of the scientific content, concepts, and procedures required by the task/s.

The response is somewhat correct with partial understanding of the required scientific content, concepts, and/or procedures demonstrated and/or explained. The response may contain some work that is incomplete or unclear.

0 – The response provides insufficient evidence to demonstrate any understanding of the scientific content, concepts, and procedures as required by the task/s for that grade level.

Response may show only information copied or rephrased from the question or insufficient correct information to receive a score of 1.

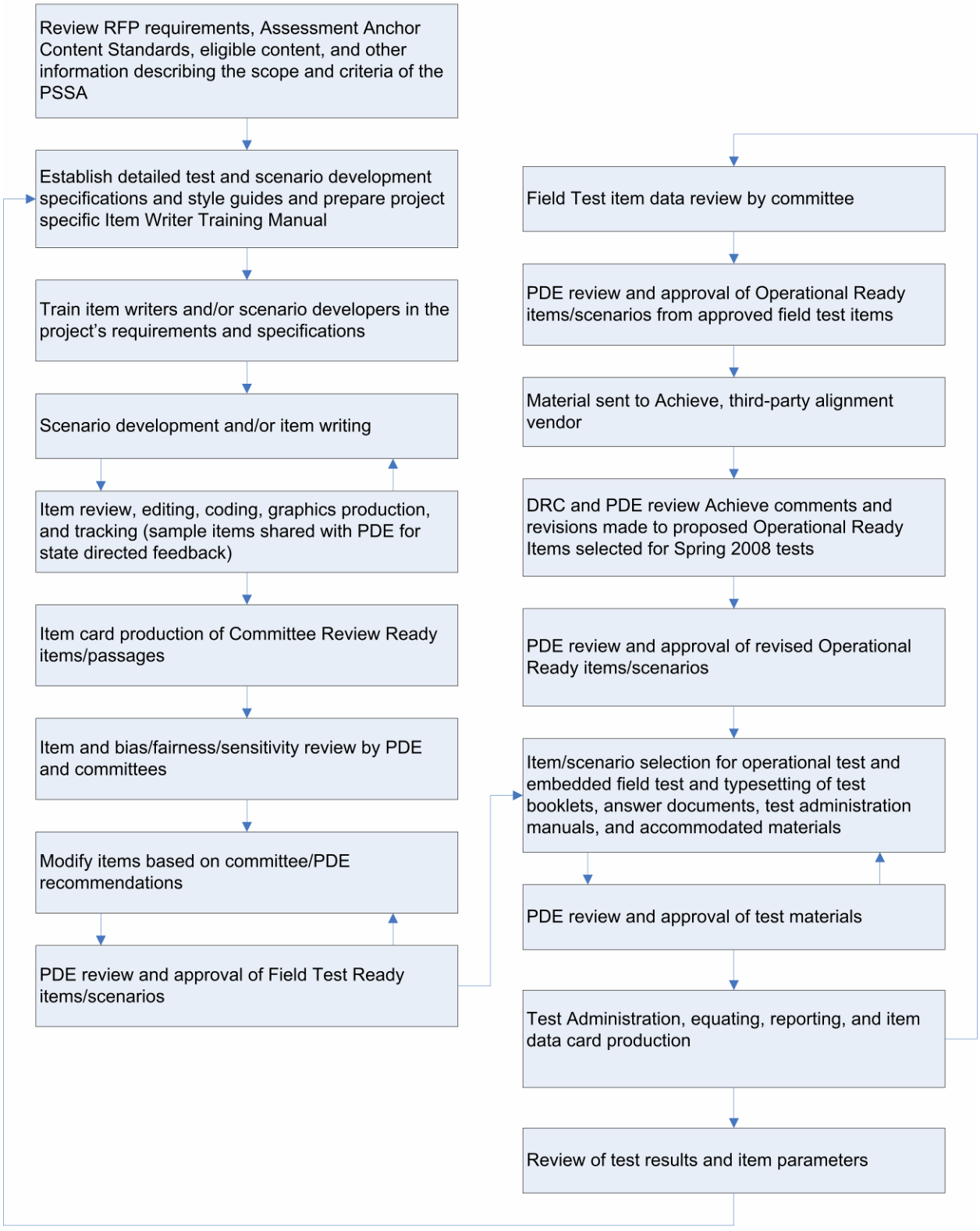
Special Categories within zero reported separately:
BLK – Blank, entirely erased or written refusal to respond
OT – Off Task
IL – Illegible
LOE – Response in a language other than English

Appendix B:

Item and Test Development Processes

DRC Item and Test Development Process

Science



Appendix C:

2008□Item Review Form

Item Rating Sheet

Reviewer Signature: _____

Content Area: _____ **Grade:** _____

July 2006 Committee Grantville, PA Unique ID number	Content Alignment	Rigor Level Alignment				Technical Design			Universal Design		STATUS
	Standards	Grade	Difficulty	Depth of Knowledge	Source of Challenge	Correct Answer	Distracters	Graphics	Language Demand	Bias	Acceptance Status
	—Higher —Lower —None	—Above —At —Below	—Easy —Medium —Hard	—Recall —Application —Strategic Thinking	—Yes —No	—Yes —No	—Yes —No	—Yes —No	—Yes —No	—Yes —No	— Approved as is. — Accepted with suggested revisions.

Appendix D:

2008 Item Review Cards and IVAN Card

Client:	50	Item ID: *(for internal use only)*

Item Writing Form - Pennsylvania

Item Writer Information		Stimulus	
Item Writer Name		**Stimulus Used**	
Item Writer #		**Stimulus ID**	
Email Address		**Stimulus Title**	
Submission Date		**Stimulus Type**	
Item Information		**Delivery**	Atch ☐ Fax ☐ Mail ☐
Item Type		**Title**	
Grade		**Author**	
Subject		**Publisher**	
Goal 1		**Date Published**	
Goal 2		**Source Page**	
Goal 3		**URL**	
Goal 4		**Permission Needed**	
Taxonomy Level			
Depth of Knowledge		**Passage Title #1**	
Difficulty		**Temp Passage ID #1**	
Focus			
Graphics		**Passage Title #2**	
Calculator		**Temp Passage ID #2**	
Points			

Comment	

Prompt / Stem

Answer Options	
Key:	
Option A. Rationale:	
Option B. Rationale:	
Option C. Rationale:	
Option D. Rationale:	

Rubric

Client:	50	Item ID: *(for internal use only)*

Passage Writing Form - Pennsylvania

Passage Writer Information		**Reference**	
Passage Writer Name		**Delivery**	Atch ☐ Fax ☐ Mail ☐
Passage Writer #		**Title**	
Email Address		**Publisher**	
Submission Date		**Date Published**	
Item Information		**Source Page**	
Subject		**URL**	
Grade		**Permission Needed**	
Type			
Category		**Paired Passage**	
Temp. Passage ID			
Passage Title			
Passage Author			

Comment	

Passage

IVAN Item Card

DATA RECOGNITION DRC CORPORATION

Released: **Item Status:**

Item Name	Item Type	Key	Grade	Subject	Report Category	Asmt Anchor	Sub - Anchor	Eligible Content	Content Difficulty	DRP	Item Calculator

Depth of Knowledge:

Administration

Form Grade	Form Subject	Form Name	Sequence	Form Type	Month	Year	Report Category	Asmt Anchor	Sub-Anchor	Eligible Content	Day	Session	Calculator

Statistics Detail

Label	P-Value	Pt. Bis. Corr.
0		
1		
2		
3		
4		
Omits		
Mean		

Label	Value
N	
Outfit t	
Logit	

DIF Analysis	Value
White/Black	
Male/Female	

Notes:

Appendix E:

2008 Performance Level Descriptors

Pennsylvania Department of Education - Grade 4 Science Performance Level Descriptors			
Below Basic	**Basic**	**Proficient**	**Advanced**
A fourth-grade student performing at Below Basic Level demonstrates a limited conceptual understanding of science content and an ineffective application of processes in the four Pennsylvania Science Reporting Categories.	A fourth-grade student performing at the Basic Level demonstrates partial conceptual understanding of science content and the application of processes in the four Pennsylvania Science Reporting Categories. A student performing at the Basic Level A. conducts simple investigations following explicit directions; identifies the processes and tools of science; recognizes changes to natural and human-made systems; recognizes models; and identifies patterns. B. recognizes the needs of organisms; describes variations in populations; distinguishes between living and nonliving; identifies environmental changes; and describes human uses of the environment. C. uses basic physical properties to describe matter; identifies forms of energy; identifies different forces; and describes position and motion of objects. D. names major Earth features; identifies natural products and resources; recognizes water systems; describes changes in weather; and connects diurnal and seasonal changes to Earth movements.	A fourth-grade student performing at the Proficient Level demonstrates a general conceptual understanding of science content and the application of processes in the four Pennsylvania Science Reporting Categories. A student performing at the Proficient Level A. conducts simple investigations and makes measurements; applies the processes of science; observes, records, and describes changes to systems; uses models; distinguishes between scientific fact and opinion; and connects changes to results. B. describes structure, function, and needs of organisms; explains the impact of adaptations; identifies basic genetic characteristics; describes ecosystem components and interactions; and identifies causes of environmental changes. C. compares multiple physical properties of matter; describes energy forms and flow through a system; and describes relative motion and position of objects and forces causing changes in motion of objects. D. describes the formation of Earth features; describes major types of resources and patterns of use; identifies soil components; identifies phase changes of water and compares water systems; identifies components of weather; and describes Earth, Sun, and Moon motions.	A fourth-grade student performing at the Advanced Level demonstrates a thorough conceptual understanding of science content and the application of processes in the four Pennsylvania Science Reporting Categories. A student performing at the Advanced Level A. designs simple investigations based on testable questions; utilizes appropriate tools of science; categorizes systems; uses models in explanations; compares patterns of change in organisms and objects; and states conclusions. B. compares the structure, function, and needs of organisms; relates adaptations to survival; describes genetic variations; compares components of ecological systems; predicts environmental changes; and describes basic competition for resources. C. categorizes matter using multiple characteristics and physical properties; illustrates energy flow and conversions through systems; and explains changes in motion of objects caused by forces. D. compares the formation of Earth features and structures; contrasts types and uses of resources; describes the composition of soil; explains the water cycle and the role of water systems; uses patterns and instruments to predict weather; and explains time through Earth, Sun, and Moon motions.

Pennsylvania Department of Education

Grade 8 Science Performance Level Descriptors

Below Basic	Basic	Proficient	Advanced
An eighth-grade student performing at Below Basic Level demonstrates a limited conceptual understanding of science content and an ineffective application of processes in the four Pennsylvania Science Reporting Categories.	An eighth-grade student performing at the Basic Level demonstrates partial conceptual understanding of science content and the application of processes in the four Pennsylvania Science Reporting Categories. A student performing at the Basic Level A. compares scientific theory to opinion; identifies inferences, descriptions, conclusions, explanations, predictions, results, and models; applies scientific process skills, appropriate measurements, and tools to solve problems; and describes systems. B. compares structural functions of organisms; identifies levels of organization; identifies adaptation, mutation, and biotechnology as variables of biological changes; identifies traits by type; and identifies major biomes and components and factors affecting community change. C. identifies elements, compounds, and mixtures; categorizes matter by properties; identifies kinetic and potential energy; describes balanced, unbalanced, frictional and gravitational forces on objects; identifies simple machines; and distinguishes between renewable and nonrenewable energy sources. D. describes basic rock types; identifies changes in Earth's surface; identifies soil types; describes the water cycle; identifies characteristics of water systems; recognizes weather patterns and distinguishes between climate types; and identifies the relationships between and among the objects of our solar system.	An eighth-grade student performing at the Proficient Level demonstrates a general conceptual understanding of science content and the application of processes in the four Pennsylvania Science Reporting Categories. A student performing at the Proficient Level A. contrasts scientific theory and opinion; develops inferences, descriptions, explanations, predictions, models, and critiques based on evidence; designs experiments to solve problems; describes system components; and communicates conclusions. B. categorizes organisms; relates structure to function; describes relationships between gene mutations, adaptations, natural selection, and biotechnology; compares gene dominance in the expression of traits; and identifies relationships within ecosystems and human impacts on the environment. C. differentiates between elements, compounds, and mixtures; describes components of simple chemical reactions; explains heat transfers and conversions; describes inertia and momentum; describes the function of simple machines; and distinguishes between forms and sources of energy. D. explains the rock cycle; compares changes in Earth's surface; explains the formation of soils and fossils; describes the physical processes in the water cycle; compares water systems; describes factors affecting regional weather or climate; describes the relationships between and among the objects of our solar system; and describes impacts of technological processes.	An eighth-grade student performing at the Advanced Level demonstrates a thorough conceptual understanding of science content and the application of processes in the four Pennsylvania Science Reporting Categories. A student performing at the Advanced Level A. explains the development and evolution of scientific theories; analyzes inferences, descriptions, explanations, predictions, results, models, and critiques; analyzes the effectiveness of experimental designs; evaluates system changes and components; and evaluates conclusions. B. explains the relationship between structure and function within organisms; connects concepts of natural selection to survival needs; makes inferences about how selection processes effect changes in human and natural systems; describes the impact of genetic changes over time on populations; and uses evidence to describe relationships within ecosystems and human impacts on the environment. C. explains the structures and the physical and chemical properties of matter; uses properties to distinguish one substance from another; uses heat transformation and conversions to evaluate forms of energy; compares the effect of forces on objects; explains mechanical advantage of simple machines; and explains environmental impacts of energy choices. D. classifies rocks by type; relates changes in Earth's surface to rock types, soil formation, and fossil formation; analyzes physical characteristics of water resources; explains global weather patterns and their impact on local weather and climate; explains the relationships between and among the objects of our solar system; and evaluates the impact of human-made processes on resources.

Pennsylvania Department of Education

Grade 11 Science Performance Level Descriptors

Below Basic	Basic	Proficient	Advanced
An eleventh-grade student performing at Below Basic Level demonstrates a limited conceptual understanding of science content and an ineffective application of processes in the four Pennsylvania Science Reporting Categories.	An eleventh-grade student performing at the Basic Level demonstrates partial conceptual understanding of science content and the application of processes in the four Pennsylvania Science Reporting Categories. A student performing at the Basic Level A. identifies appropriate technologies; compares system parts; describes models; and recognizes patterns and processes. B. describes structure and function at multiple levels of organization; recognizes the mechanisms of evolution; describes genetic information; recognizes the characteristics of and interactions within ecosystems; describes patterns of change; and identifies the impact of human-made systems. C. recognizes the relationship between structure and property of matter; describes energy sources, transfer, and conversion; recognizes environmental impacts of energy use; and describes principles of force and motion. D. describes lithospheric forces; describes functions, relationships, and factors relating to water resources; uses weather data and climate influences to make predictions; and describes the composition and structure of the universe.	An eleventh-grade student performing at the Proficient Level demonstrates a general conceptual understanding of science content and the application of processes in the four Pennsylvania Science Reporting Categories. A student performing at the Proficient Level A. compares appropriate technologies; explains systems; compares models; analyzes patterns and processes; and describes experimental or design processes. B. explains structure and function at multiple levels of organization; explains the mechanisms of evolution; describes the expression and inheritance of genetic information; explains the characteristics of and interactions within ecosystems; interprets patterns of change; and explains the impact of human-made systems. C. explains the relationship between structure and property of matter; compares energy sources, transfer, and conversion; explains environmental impacts of energy use; and uses principles of force and motion to solve problems. D. explains lithospheric forces; explains significance and contributions of water resources; analyzes weather and climate influences; and explains the composition and structure of the universe.	An eleventh-grade student performing at the Advanced Level demonstrates a thorough conceptual understanding of science content and the application of processes in the four Pennsylvania Science Reporting Categories. A student performing at the Advanced Level A. evaluates appropriate technologies; analyzes systems; evaluates models and processes; predicts patterns; and critiques aspects of experimental or design processes. B. evaluates structure and function at multiple levels of organization; explains the theory of evolution; interprets genetic information; evaluates the characteristics of and interactions within ecosystems; analyzes patterns of change; and evaluates the impact of human-made systems. C. analyzes the relationship between structure and property of matter; evaluates energy sources, transfer, and conversion; analyzes environmental impacts of energy use; and evaluates principles of force and motion. D. analyzes lithospheric forces; analyzes water resources; evaluates weather and climate influences; and explains the formation of the solar system and the composition, structure, and evolution of the universe.

Appendix F:

2008 Grades 4, 8, and 11 Multiple-Choice Statistics

Table Legend

Column	Description
Cont.	Tested Content
Grade	Tested Grade
ID	Item identification number
Pub. ID	Public item identification number
Form	Form on which the item appeared 'All' = common form
Seq.	Sequence number of the item
Status	Status of the item at the time of administration 'OP' = common, 'MX' = matrix 'FT' = field test
Key	Answer key
R.C.	Reporting category
D.O.K.	Depth of knowledge
N	Number of students
P-Val.	P-Value
Proportion [A-D, Omit]	Proportion selecting the response
Pt. Bis.	Item total correlation
Correlation [A-D]	Point biserial of the response
Logit	Rasch logit difficulty
Logit SE	Rasch logit difficulty standard error
Fit	Rasch outfit statistic

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	4	554137	0001	All	1	OP	C	A	2	126117	0.854	0.019	0.082	0.854	0.045	0.000	0.333	-0.170	-0.340	0.333	-0.101	-0.928	0.008	0.0
Science	4	553956	0002	All	2	OP	B	A	2	126117	0.780	0.061	0.780	0.098	0.061	0.000	0.464	-0.391	0.464	-0.268	-0.398	-0.355	0.007	-9.9
Science	4	557900	0003	All	3	OP	D	B	1	126117	0.534	0.164	0.100	0.201	0.534	0.001	0.342	-0.354	-0.288	-0.254	0.342	0.980	0.006	9.9
Science	4	554278	0004	All	4	OP	D	D	2	126117	0.591	0.148	0.079	0.181	0.591	0.001	0.460	-0.324	-0.440	-0.428	0.460	0.704	0.006	-9.9
Science	4	554238	0005	All	5	OP	A	A	2	126117	0.773	0.773	0.049	0.099	0.079	0.000	0.479	0.479	-0.333	-0.346	-0.384	-0.306	0.007	-9.9
Science	4	554233	0006	All	6	OP	C	A	2	126117	0.555	0.213	0.055	0.555	0.176	0.001	0.306	-0.314	-0.379	0.306	-0.150	0.878	0.006	9.9
Science	4	553950	0007	All	7	OP	A	A	1	126117	0.533	0.533	0.190	0.098	0.178	0.001	0.219	0.219	-0.115	-0.274	-0.202	0.983	0.006	9.9
Science	4	554106	0008	All	8	OP	C	C	2	126117	0.802	0.034	0.044	0.802	0.119	0.000	0.395	-0.307	-0.329	0.395	-0.255	-0.508	0.008	-8.7
Science	4	554156	0009	All	9	OP	B	D	2	126117	0.508	0.279	0.508	0.113	0.100	0.000	0.358	-0.284	0.358	-0.338	-0.372	1.110	0.006	9.9
Science	4	554124	0010	All	10	OP	A	C	2	126117	0.748	0.748	0.082	0.022	0.146	0.001	0.320	0.320	-0.246	-0.239	-0.237	-0.152	0.007	9.9
Science	4	554132	0011	All	11	OP	C	B	2	126117	0.899	0.028	0.048	0.899	0.025	0.001	0.446	-0.309	-0.321	0.446	-0.251	-1.390	0.010	-9.9
Science	4	554026	0012	All	12	OP	B	D	2	126117	0.687	0.122	0.687	0.058	0.132	0.001	0.389	-0.272	0.389	-0.293	-0.347	0.201	0.007	6.8
Science	4	554042	0013	All	13	OP	B	B	2	126117	0.578	0.032	0.578	0.302	0.087	0.001	0.297	-0.272	0.297	-0.245	-0.266	0.767	0.006	9.9
Science	4	557901	0014	All	14	OP	C	A	2	126117	0.812	0.042	0.028	0.812	0.116	0.001	0.344	-0.265	-0.273	0.344	-0.223	-0.589	0.008	1.2
Science	4	554230	0015	All	15	OP	D	D	2	126117	0.661	0.125	0.162	0.050	0.661	0.001	0.326	-0.317	-0.172	-0.354	0.326	0.341	0.007	9.9
Science	4	554142	0016	All	16	OP	C	B	2	126117	0.640	0.111	0.133	0.640	0.115	0.001	0.540	-0.425	-0.452	0.540	-0.475	0.457	0.006	-9.9
Science	4	553976	0017	All	17	OP	C	A	2	126117	0.788	0.059	0.056	0.788	0.097	0.001	0.420	-0.265	-0.323	0.420	-0.324	-0.414	0.007	-9.9
Science	4	554085	0018	All	18	OP	B	B	2	126117	0.591	0.120	0.591	0.056	0.232	0.001	0.431	-0.436	0.431	-0.325	-0.339	0.704	0.006	-9.9
Science	4	554066	0019	All	19	OP	A	A	1	126117	0.698	0.698	0.039	0.076	0.185	0.001	0.424	0.424	-0.383	-0.424	-0.271	0.137	0.007	-9.9
Science	4	554254	0020	All	20	OP	D	B	2	126117	0.773	0.035	0.073	0.118	0.773	0.001	0.468	-0.362	-0.321	-0.366	0.468	-0.309	0.007	-9.9
Science	4	553960	0021	All	21	OP	D	D	2	126117	0.824	0.123	0.025	0.026	0.824	0.001	0.296	-0.179	-0.259	-0.257	0.296	-0.680	0.008	9.9
Science	4	554119	0022	All	22	OP	A	A	2	126117	0.808	0.808	0.045	0.058	0.086	0.001	0.502	0.502	-0.364	-0.353	-0.377	-0.554	0.008	-9.9
Science	4	554067	0023	All	23	OP	D	A	2	126117	0.927	0.015	0.023	0.033	0.927	0.001	0.371	-0.237	-0.237	-0.241	0.371	-1.769	0.011	-9.9
Science	4	554090	0024	All	24	OP	C	A	1	126117	0.535	0.140	0.209	0.535	0.114	0.002	0.353	-0.336	-0.306	0.353	-0.276	0.980	0.006	9.9
Science	4	553946	0025	All	25	OP	C	D	2	126117	0.780	0.054	0.043	0.780	0.121	0.001	0.385	-0.351	-0.284	0.385	-0.237	-0.355	0.007	-2.7
Science	4	553991	0026	All	26	OP	B	A	2	126117	0.803	0.108	0.803	0.045	0.041	0.001	0.403	-0.275	0.403	-0.300	-0.304	-0.523	0.008	-9.9
Science	4	554263	0027	All	27	OP	D	A	2	126117	0.749	0.106	0.060	0.083	0.749	0.001	0.531	-0.384	-0.411	-0.428	0.531	-0.157	0.007	-9.9
Science	4	554043	0028	All	28	OP	D	A	1	126117	0.676	0.029	0.241	0.052	0.676	0.002	0.462	-0.374	-0.373	-0.408	0.462	0.262	0.007	-9.9
Science	4	553957	0029	All	36	OP	D	A	2	126117	0.944	0.014	0.007	0.035	0.944	0.000	0.251	-0.182	-0.166	-0.148	0.251	-2.081	0.013	2.6
Science	4	554193	0030	All	37	OP	A	B	2	126117	0.851	0.851	0.068	0.042	0.038	0.000	0.316	0.316	-0.213	-0.221	-0.201	-0.910	0.008	1.5
Science	4	554107	0031	All	38	OP	D	B	1	126117	0.425	0.135	0.153	0.286	0.425	0.001	0.247	-0.273	-0.242	-0.216	0.247	1.507	0.006	9.9
Science	4	554252	0032	All	39	OP	A	D	2	126117	0.683	0.683	0.162	0.043	0.111	0.001	0.455	0.455	-0.397	-0.367	-0.315	0.227	0.007	-9.9
Science	4	553969	0033	All	40	OP	C	A	2	126117	0.642	0.132	0.140	0.642	0.085	0.001	0.410	-0.260	-0.381	0.410	-0.351	0.447	0.006	-7.8
Science	4	554005	0034	All	41	OP	A	A	2	126117	0.771	0.771	0.040	0.098	0.091	0.001	0.466	0.466	-0.318	-0.335	-0.383	-0.296	0.007	-9.9
Science	4	553970	0035	All	42	OP	B	A	2	126117	0.857	0.034	0.857	0.047	0.062	0.001	0.486	-0.328	0.486	-0.314	-0.367	-0.944	0.009	-9.9
Science	4	554173	0036	All	43	OP	B	A	3	126117	0.515	0.201	0.515	0.164	0.119	0.001	0.322	-0.261	0.322	-0.293	-0.299	1.076	0.006	9.9
Science	4	554135	0037	All	44	OP	A	D	2	126117	0.852	0.852	0.105	0.020	0.022	0.000	0.238	0.238	-0.140	-0.210	-0.193	-0.913	0.008	9.9
Science	4	553973	0038	All	45	OP	B	A	2	126117	0.721	0.089	0.721	0.156	0.033	0.001	0.493	-0.328	0.493	-0.472	-0.256	0.011	0.007	-9.9
Science	4	554003	0039	All	46	OP	C	B	2	126117	0.673	0.108	0.133	0.673	0.085	0.001	0.444	-0.313	-0.369	0.444	-0.380	0.280	0.007	-9.9
Science	4	554204	0040	All	47	OP	B	A	2	126117	0.815	0.079	0.815	0.054	0.051	0.001	0.482	-0.367	0.482	-0.328	-0.334	-0.608	0.008	-9.9
Science	4	554061	0041	All	48	OP	C	A	2	126117	0.446	0.132	0.157	0.446	0.264	0.001	0.359	-0.386	-0.401	0.359	-0.276	1.410	0.006	9.9
Science	4	554027	0042	All	49	OP	B	A	2	126117	0.824	0.065	0.824	0.060	0.050	0.001	0.435	-0.284	0.435	-0.292	-0.348	-0.675	0.008	-9.9
Science	4	554037	0043	All	50	OP	C	A	2	126117	0.637	0.167	0.056	0.637	0.138	0.001	0.394	-0.378	-0.396	0.394	-0.216	0.469	0.006	4.1

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	4	554012	0044	All	51	OP	B	C	2	126117	0.436	0.290	0.436	0.080	0.194	0.001	0.296	-0.243	0.296	-0.418	-0.269	1.460	0.006	9.9
Science	4	554196	0045	All	52	OP	A	C	2	126117	0.836	0.836	0.073	0.060	0.031	0.000	0.315	0.315	-0.260	-0.140	-0.249	-0.780	0.008	6.4
Science	4	554224	0046	All	53	OP	B	C	2	126117	0.661	0.064	0.661	0.077	0.197	0.001	0.431	-0.362	0.431	-0.390	-0.318	0.346	0.007	-9.9
Science	4	554000	0047	All	54	OP	C	A	2	126117	0.747	0.064	0.075	0.747	0.112	0.001	0.418	-0.321	-0.353	0.418	-0.274	-0.144	0.007	-9.9
Science	4	554076	0048	All	55	OP	D	A	1	126117	0.654	0.163	0.096	0.086	0.654	0.001	0.426	-0.340	-0.357	-0.335	0.426	0.384	0.006	-9.9
Science	4	554028	0049	All	56	OP	A	A	2	126117	0.545	0.545	0.113	0.216	0.124	0.001	0.439	0.439	-0.452	-0.393	-0.312	0.932	0.006	-9.9
Science	4	554148	0050	All	57	OP	C	D	2	126117	0.545	0.269	0.080	0.545	0.105	0.001	0.438	-0.381	-0.440	0.438	-0.350	0.933	0.006	-9.9
Science	4	554209	0051	All	58	OP	B	C	2	126117	0.516	0.158	0.516	0.129	0.196	0.001	0.236	-0.209	0.236	-0.301	-0.140	1.072	0.006	9.9
Science	4	554147	0052	All	59	OP	D	A	2	126117	0.774	0.114	0.052	0.058	0.774	0.001	0.507	-0.373	-0.396	-0.378	0.507	-0.317	0.007	-9.9
Science	4	554242	0053	All	60	OP	A	A	2	126117	0.488	0.488	0.184	0.098	0.229	0.001	0.437	0.437	-0.386	-0.432	-0.397	1.210	0.006	-9.9
Science	4	554172	0054	All	61	OP	A	C	2	126117	0.907	0.907	0.029	0.035	0.028	0.001	0.399	0.399	-0.260	-0.262	-0.252	-1.492	0.010	-9.9
Science	4	554218	0055	All	62	OP	C	C	2	126117	0.910	0.028	0.032	0.910	0.029	0.001	0.355	-0.238	-0.233	0.355	-0.208	-1.531	0.010	-9.9
Science	4	554167	0056	All	63	OP	D	C	2	126117	0.812	0.042	0.097	0.047	0.812	0.001	0.385	-0.241	-0.267	-0.311	0.385	-0.588	0.008	-5.1
Science	4	553958	0057	1	29	MX	A	A	3	7141	0.351	0.351	0.165	0.313	0.169	0.001	0.132	0.132	-0.138	-0.119	-0.154	1.814	0.027	9.9
Science	4	554072	0058	1	30	MX	A	D	2	7141	0.280	0.280	0.279	0.280	0.158	0.002	0.152	0.152	-0.083	-0.187	-0.275	2.190	0.028	9.9
Science	4	558197	0059	1	31	FT	D	C		7141	0.565	0.146	0.214	0.070	0.565	0.004	0.375	-0.287	-0.346	-0.340	0.375	0.783	0.026	-1.8
Science	4	558200	0060	1	64	FT	D	A		7141	0.517	0.313	0.097	0.071	0.517	0.001	0.291	-0.160	-0.361	-0.404	0.291	1.009	0.026	5.3
Science	4	558198	0061	1	65	FT	C	C		7141	0.465	0.213	0.164	0.465	0.156	0.001	0.201	-0.234	-0.220	0.201	-0.081	1.256	0.026	9.9
Science	4	558201	0062	1	66	FT	C	C		7141	0.452	0.170	0.162	0.452	0.212	0.003	0.274	-0.289	-0.280	0.274	-0.215	1.318	0.026	7.3
Science	4	554018	0063	2	29	MX	D	C	2	7006	0.767	0.065	0.074	0.091	0.767	0.001	0.334	-0.254	-0.216	-0.241	0.334	-0.288	0.031	-0.9
Science	4	554104	0064	2	30	MX	B	A	2	7006	0.560	0.246	0.560	0.086	0.106	0.002	0.414	-0.347	0.414	-0.375	-0.372	0.847	0.026	-3.6
Science	4	558204	0065	2	31	FT	A	D		7006	0.743	0.743	0.118	0.065	0.071	0.003	0.443	0.443	-0.292	-0.378	-0.346	-0.135	0.030	-6.6
Science	4	558202	0066	2	64	FT	C	B		7006	0.771	0.092	0.066	0.771	0.070	0.001	0.480	-0.343	-0.362	0.480	-0.357	-0.310	0.031	-9.3
Science	4	558203	0067	2	65	FT	B	A		7006	0.706	0.065	0.706	0.052	0.174	0.002	0.345	-0.302	0.345	-0.296	-0.235	0.083	0.029	1.2
Science	4	558205	0068	2	66	FT	B	C		7006	0.912	0.032	0.912	0.026	0.026	0.004	0.398	-0.283	0.398	-0.235	-0.265	-1.580	0.044	-6.0
Science	4	554009	0069	3	29	MX	C	B	2	6978	0.702	0.074	0.089	0.702	0.135	0.001	0.421	-0.387	-0.356	0.421	-0.265	0.115	0.028	-4.9
Science	4	554098	0070	3	30	MX	B	D	1	6978	0.805	0.054	0.805	0.071	0.068	0.002	0.356	-0.267	0.356	-0.233	-0.258	-0.532	0.032	-1.1
Science	4	558206	0071	3	31	FT	D	B		6978	0.689	0.035	0.192	0.079	0.689	0.005	0.363	-0.312	-0.225	-0.389	0.363	0.188	0.028	-0.4
Science	4	558207	0072	3	64	FT	A	C		6978	0.591	0.591	0.107	0.087	0.214	0.001	0.208	0.208	-0.216	-0.210	-0.119	0.694	0.027	9.9
Science	4	558209	0073	3	65	FT	B	A		6978	0.907	0.030	0.907	0.041	0.019	0.002	0.421	-0.297	0.421	-0.274	-0.259	-1.491	0.043	-7.6
Science	4	558211	0074	3	66	FT	A	D		6978	0.357	0.357	0.046	0.338	0.255	0.004	0.137	0.137	-0.315	-0.072	-0.172	1.832	0.027	9.9
Science	4	554248	0075	4	29	MX	C	C	2	7022	0.561	0.206	0.127	0.561	0.105	0.001	0.229	-0.158	-0.220	0.229	-0.209	0.834	0.026	9.3
Science	4	554251	0076	4	30	MX	A	C	2	7022	0.763	0.763	0.039	0.079	0.116	0.002	0.256	0.256	-0.314	-0.166	-0.134	-0.248	0.030	6.9
Science	4	558212	0077	4	31	FT	B	A		7022	0.421	0.159	0.421	0.199	0.218	0.003	0.282	-0.333	0.282	-0.145	-0.346	1.494	0.026	6.6
Science	4	558214	0078	4	64	FT	B	B		7022	0.577	0.128	0.577	0.081	0.214	0.001	0.251	-0.334	0.251	-0.258	-0.100	0.756	0.026	9.6
Science	4	558215	0079	4	65	FT	A	C		7022	0.408	0.408	0.183	0.195	0.212	0.002	0.180	0.180	-0.167	-0.150	-0.194	1.559	0.026	9.9
Science	4	558213	0080	4	66	FT	B	D		7022	0.258	0.208	0.258	0.260	0.269	0.005	0.130	-0.155	0.130	-0.105	-0.186	2.347	0.029	9.9
Science	4	553940	0081	5	29	MX	B	A	2	7005	0.808	0.079	0.808	0.078	0.034	0.001	0.394	-0.269	0.394	-0.280	-0.307	-0.546	0.032	-3.9
Science	4	554102	0082	5	30	MX	D	D	2	7005	0.384	0.201	0.213	0.201	0.384	0.002	0.191	-0.114	-0.086	-0.367	0.191	1.687	0.027	9.9
Science	4	558217	0083	5	31	FT	C	A		7005	0.443	0.347	0.068	0.443	0.139	0.003	0.223	-0.126	-0.325	0.223	-0.319	1.401	0.026	9.9
Science	4	558218	0084	5	64	FT	D	C		7005	0.543	0.132	0.156	0.167	0.543	0.001	0.271	-0.176	-0.257	-0.256	0.271	0.927	0.026	6.0
Science	4	558220	0085	5	65	FT	A	C		7005	0.756	0.756	0.051	0.120	0.071	0.002	0.337	0.337	-0.297	-0.192	-0.287	-0.195	0.030	-0.4
Science	4	558219	0086	5	66	FT	C	D		7005	0.280	0.189	0.149	0.280	0.378	0.003	0.097	-0.172	-0.243	0.097	-0.010	2.233	0.028	9.9

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	4	553971	0087	6	29	MX	B	D	2	7021	0.726	0.081	0.726	0.109	0.083	0.001	0.482	-0.374	0.482	-0.375	-0.355	-0.025	0.029	-9.2
Science	4	554029	0088	6	30	MX	D	A	1	7021	0.936	0.019	0.023	0.018	0.936	0.003	0.387	-0.275	-0.236	-0.216	0.387	-1.939	0.050	-7.6
Science	4	558223	0089	6	31	FT	C	A		7021	0.297	0.217	0.183	0.297	0.299	0.003	0.077	-0.068	-0.197	0.077	-0.013	2.159	0.028	9.9
Science	4	558222	0090	6	64	FT	C	A		7021	0.631	0.126	0.163	0.631	0.078	0.001	0.315	-0.190	-0.259	0.315	-0.322	0.493	0.027	2.0
Science	4	558224	0091	6	65	FT	A	B		7021	0.453	0.453	0.281	0.085	0.178	0.002	0.356	0.356	-0.320	-0.360	-0.331	1.364	0.026	1.1
Science	4	558225	0092	6	66	FT	D	C		7021	0.585	0.129	0.141	0.140	0.585	0.004	0.263	-0.296	-0.186	-0.173	0.263	0.725	0.026	9.0
Science	4	553996	0093	7	29	MX	C	A	2	6991	0.738	0.080	0.149	0.738	0.032	0.001	0.417	-0.288	-0.344	0.417	-0.323	-0.077	0.029	-5.1
Science	4	554243	0094	7	30	MX	C	B	2	6991	0.749	0.120	0.092	0.749	0.037	0.001	0.363	-0.266	-0.236	0.363	-0.341	-0.147	0.030	-1.6
Science	4	558227	0095	7	31	FT	B	D		6991	0.447	0.290	0.447	0.219	0.041	0.003	0.331	-0.287	0.331	-0.318	-0.440	1.392	0.026	2.6
Science	4	558229	0096	7	64	FT	D	D		6991	0.662	0.076	0.223	0.036	0.662	0.001	0.246	-0.184	-0.169	-0.294	0.246	0.343	0.027	8.0
Science	4	558228	0097	7	65	FT	D	B		6991	0.343	0.203	0.127	0.325	0.343	0.002	0.140	-0.060	-0.391	-0.087	0.140	1.902	0.027	9.9
Science	4	558230	0098	7	66	FT	B	C		6991	0.219	0.322	0.219	0.201	0.253	0.004	-0.013	0.003	-0.013	-0.002	0.049	2.613	0.031	9.9
Science	4	554192	0099	8	29	MX	C	A	2	7003	0.512	0.188	0.134	0.512	0.163	0.002	0.357	-0.307	-0.352	0.357	-0.291	1.087	0.026	-0.2
Science	4	554277	0100	8	30	MX	C	D	2	7003	0.434	0.294	0.188	0.434	0.082	0.002	0.309	-0.290	-0.294	0.309	-0.317	1.465	0.026	4.4
Science	4	558232	0101	8	31	FT	C	A		7003	0.683	0.054	0.113	0.683	0.147	0.003	0.462	-0.336	-0.336	0.462	-0.412	0.234	0.028	-8.1
Science	4	558233	0102	8	64	FT	B	C		7003	0.630	0.135	0.630	0.128	0.106	0.002	0.454	-0.352	0.454	-0.393	-0.367	0.513	0.027	-9.5
Science	4	558234	0103	8	65	FT	A	D		7003	0.377	0.377	0.170	0.200	0.251	0.002	0.177	0.177	-0.268	-0.217	-0.072	1.745	0.027	9.9
Science	4	558235	0104	8	66	FT	A	B		7003	0.436	0.436	0.223	0.132	0.205	0.003	0.252	0.252	-0.272	-0.256	-0.187	1.453	0.026	9.9
Science	4	554164	0105	9	29	MX	D	B	2	7006	0.232	0.272	0.298	0.195	0.232	0.002	0.118	-0.093	-0.127	-0.206	0.118	2.559	0.030	9.9
Science	4	554129	0106	9	30	MX	B	A	2	7006	0.804	0.069	0.804	0.071	0.054	0.002	0.444	-0.313	0.444	-0.329	-0.311	-0.516	0.032	-6.3
Science	4	558238	0107	9	31	FT	D	B		7006	0.498	0.294	0.112	0.092	0.498	0.004	0.290	-0.231	-0.274	-0.292	0.290	1.154	0.026	7.5
Science	4	558240	0108	9	64	FT	C	A		7006	0.701	0.107	0.115	0.701	0.075	0.002	0.453	-0.323	-0.401	0.453	-0.330	0.132	0.028	-6.5
Science	4	558237	0109	9	65	FT	A	C		7006	0.721	0.721	0.074	0.106	0.098	0.001	0.380	0.380	-0.269	-0.270	-0.313	0.020	0.029	-3.6
Science	4	558239	0110	9	66	FT	C	D		7006	0.495	0.141	0.248	0.495	0.114	0.002	0.224	-0.201	-0.172	0.224	-0.250	1.171	0.026	9.9
Science	4	549576	0111	10	29	MX	B	C	1	7011	0.705	0.123	0.705	0.083	0.089	0.001	0.448	-0.350	0.448	-0.343	-0.352	0.137	0.028	-6.4
Science	4	554038	0112	10	30	MX	B	D	1	7011	0.506	0.193	0.506	0.119	0.181	0.001	0.249	-0.233	0.249	-0.237	-0.204	1.142	0.026	9.9
Science	4	558242	0113	10	31	FT	A	C		7011	0.846	0.846	0.061	0.044	0.047	0.002	0.474	0.474	-0.330	-0.321	-0.339	-0.828	0.035	-8.7
Science	4	558244	0114	10	64	FT	D	D		7011	0.524	0.117	0.225	0.132	0.524	0.001	0.218	-0.228	-0.165	-0.194	0.218	1.056	0.026	9.9
Science	4	558241	0115	10	65	FT	C	B		7011	0.859	0.055	0.054	0.859	0.031	0.002	0.468	-0.313	-0.348	0.468	-0.303	-0.941	0.036	-8.6
Science	4	558248	0116	10	66	FT	D	D		7011	0.431	0.152	0.280	0.134	0.431	0.003	0.332	-0.404	-0.296	-0.245	0.332	1.505	0.026	3.9
Science	4	553974	0117	11	29	MX	A	D	2	7023	0.532	0.532	0.255	0.136	0.075	0.002	0.119	0.119	-0.100	-0.052	-0.174	1.004	0.026	9.9
Science	4	554143	0118	11	30	MX	B	B	2	7023	0.772	0.142	0.772	0.028	0.057	0.002	0.337	-0.265	0.337	-0.244	-0.218	-0.289	0.031	-1.1
Science	4	558247	0119	11	31	FT	A	A		7023	0.886	0.886	0.018	0.023	0.070	0.003	0.330	0.330	-0.175	-0.222	-0.250	-1.236	0.039	-2.6
Science	4	558246	0120	11	64	FT	B	B		7023	0.748	0.075	0.748	0.078	0.096	0.002	0.372	-0.272	0.372	-0.255	-0.287	-0.139	0.030	-2.0
Science	4	558250	0121	11	65	FT	D	C		7023	0.738	0.105	0.082	0.072	0.738	0.003	0.427	-0.309	-0.302	-0.360	0.427	-0.077	0.029	-5.6
Science	4	558252	0122	11	66	FT	C	D		7023	0.727	0.196	0.034	0.727	0.039	0.004	0.308	-0.180	-0.336	0.308	-0.325	-0.011	0.029	4.8
Science	4	554179	0123	12	29	MX	A	C	2	7000	0.443	0.443	0.287	0.086	0.182	0.001	0.249	0.249	-0.183	-0.333	-0.261	1.404	0.026	9.2
Science	4	554241	0124	12	30	MX	D	B	2	7000	0.762	0.086	0.074	0.077	0.762	0.002	0.465	-0.292	-0.384	-0.368	0.465	-0.227	0.030	-8.1
Science	4	558251	0125	12	31	FT	A	B		7000	0.236	0.236	0.516	0.116	0.127	0.005	0.126	0.126	-0.005	-0.329	-0.435	2.502	0.030	9.9
Science	4	558253	0126	12	64	FT	B	A		7000	0.336	0.416	0.336	0.127	0.120	0.001	0.208	-0.148	0.208	-0.283	-0.313	1.937	0.027	9.9
Science	4	558256	0127	12	65	FT	A	C		7000	0.750	0.750	0.023	0.170	0.055	0.001	0.184	0.184	-0.263	-0.055	-0.221	-0.155	0.030	9.9
Science	4	558254	0128	12	66	FT	B	D		7000	0.387	0.260	0.387	0.207	0.143	0.004	0.025	-0.023	0.025	-0.007	-0.047	1.679	0.027	9.9
Science	4	554086	0129	13	29	MX	A	C	2	6969	0.596	0.596	0.300	0.051	0.051	0.002	0.207	0.207	-0.084	-0.301	-0.334	0.694	0.027	9.9

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	4	554207	0130	13	30	MX	A	B	2	6969	0.868	0.868	0.052	0.043	0.034	0.002	0.472	0.472	-0.330	-0.339	-0.302	-1.042	0.037	-9.4
Science	4	558257	0131	13	31	FT	B	A		6969	0.828	0.057	0.828	0.082	0.030	0.004	0.409	-0.283	0.409	-0.280	-0.317	-0.685	0.034	-4.9
Science	4	558258	0132	13	64	FT	D	B		6969	0.592	0.065	0.236	0.106	0.592	0.001	0.334	-0.332	-0.300	-0.204	0.334	0.712	0.027	2.2
Science	4	558260	0133	13	65	FT	D	C		6969	0.507	0.077	0.075	0.339	0.507	0.002	0.293	-0.373	-0.382	-0.185	0.293	1.125	0.026	7.6
Science	4	558259	0134	13	66	FT	A	D		6969	0.404	0.404	0.123	0.405	0.065	0.003	0.380	0.380	-0.409	-0.369	-0.322	1.620	0.027	-1.1
Science	4	554091	0135	14	29	MX	D	D	2	6985	0.522	0.083	0.098	0.296	0.522	0.001	0.354	-0.353	-0.228	-0.344	0.354	1.055	0.026	0.7
Science	4	553986	0136	14	30	MX	B	C	2	6985	0.777	0.107	0.777	0.042	0.072	0.001	0.418	-0.318	0.418	-0.280	-0.317	-0.315	0.031	-4.5
Science	4	558265	0137	14	31	FT	D	D		6985	0.904	0.043	0.027	0.021	0.904	0.004	0.327	-0.218	-0.199	-0.226	0.327	-1.436	0.042	-0.9
Science	4	558262	0138	14	64	FT	B	A		6985	0.365	0.266	0.365	0.172	0.196	0.002	0.194	-0.164	0.194	-0.251	-0.182	1.819	0.027	9.9
Science	4	558263	0139	14	65	FT	C	C		6985	0.611	0.063	0.098	0.611	0.227	0.001	0.400	-0.394	-0.421	0.400	-0.268	0.621	0.027	-3.0
Science	4	558264	0140	14	66	FT	C	B		6985	0.544	0.190	0.151	0.544	0.111	0.003	0.319	-0.302	-0.256	0.319	-0.271	0.949	0.026	4.6
Science	4	554189	0141	15	29	MX	B	C	2	7034	0.743	0.103	0.743	0.081	0.071	0.001	0.430	-0.356	0.430	-0.335	-0.267	-0.101	0.029	-5.2
Science	4	554068	0142	15	30	MX	D	D	2	7034	0.479	0.112	0.202	0.206	0.479	0.002	0.184	-0.217	-0.115	-0.181	0.184	1.258	0.026	9.9
Science	4	558268	0143	15	31	FT	D	B		7034	0.360	0.188	0.227	0.223	0.360	0.003	0.127	-0.215	-0.032	-0.145	0.127	1.842	0.027	9.9
Science	4	558267	0144	15	64	FT	A	A		7034	0.658	0.658	0.122	0.150	0.069	0.001	0.404	0.404	-0.399	-0.239	-0.357	0.379	0.027	-3.3
Science	4	558269	0145	15	65	FT	B	D		7034	0.681	0.107	0.681	0.084	0.126	0.002	0.397	-0.281	0.397	-0.340	-0.315	0.257	0.028	-3.0
Science	4	558271	0146	15	66	FT	C	C		7034	0.390	0.284	0.114	0.390	0.208	0.003	0.339	-0.291	-0.395	0.339	-0.353	1.687	0.027	3.0
Science	4	554089	0147	16	29	MX	B	C	2	6964	0.567	0.079	0.567	0.137	0.215	0.002	0.347	-0.348	0.347	-0.227	-0.320	0.827	0.027	2.5
Science	4	554211	0148	16	30	MX	C	A	2	6964	0.796	0.096	0.052	0.796	0.052	0.005	0.519	-0.396	-0.388	0.519	-0.355	-0.463	0.032	-9.9
Science	4	558272	0149	16	31	FT	A	A		6964	0.382	0.382	0.358	0.068	0.185	0.006	0.280	0.280	-0.223	-0.357	-0.328	1.727	0.027	9.1
Science	4	558273	0150	16	64	FT	A	B		6964	0.825	0.825	0.055	0.046	0.074	0.001	0.479	0.479	-0.314	-0.340	-0.361	-0.681	0.034	-8.8
Science	4	558274	0151	16	65	FT	C	C		6964	0.515	0.077	0.296	0.515	0.110	0.001	0.279	-0.219	-0.239	0.279	-0.277	1.081	0.026	8.4
Science	4	558275	0152	16	66	FT	A	D		6964	0.438	0.438	0.067	0.245	0.248	0.003	0.271	0.271	-0.359	-0.193	-0.278	1.451	0.026	9.9
Science	4	554044	0153	17	29	MX	C	D	2	6967	0.551	0.222	0.157	0.551	0.067	0.001	0.378	-0.355	-0.268	0.378	-0.360	0.865	0.026	-2.2
Science	4	554229	0154	17	30	MX	D	B	1	6967	0.789	0.021	0.069	0.118	0.789	0.003	0.460	-0.291	-0.338	-0.372	0.460	-0.450	0.031	-7.2
Science	4	558279	0155	17	31	FT	D	B		6967	0.513	0.149	0.163	0.170	0.513	0.004	0.356	-0.254	-0.326	-0.370	0.356	1.048	0.026	1.7
Science	4	558277	0156	17	64	FT	C	A		6967	0.724	0.045	0.136	0.724	0.093	0.001	0.423	-0.315	-0.323	0.423	-0.325	-0.036	0.029	-6.1
Science	4	558278	0157	17	65	FT	A	C		6967	0.904	0.904	0.057	0.022	0.016	0.001	0.407	0.407	-0.310	-0.243	-0.237	-1.485	0.042	-6.7
Science	4	558280	0158	17	66	FT	B	D		6967	0.369	0.229	0.369	0.276	0.122	0.003	0.292	-0.210	0.292	-0.284	-0.451	1.753	0.027	6.8
Science	4	553980	0159	18	29	MX	B	A	2	6991	0.835	0.070	0.835	0.057	0.036	0.001	0.382	-0.179	0.382	-0.320	-0.315	-0.759	0.034	-2.3
Science	4	554249	0160	18	30	MX	B	C	2	6991	0.723	0.064	0.723	0.090	0.122	0.001	0.496	-0.368	0.496	-0.385	-0.386	0.003	0.029	-9.9
Science	4	558284	0161	18	31	FT	C	D		6991	0.570	0.253	0.074	0.570	0.100	0.004	0.380	-0.278	-0.343	0.380	-0.390	0.806	0.026	-2.0
Science	4	558282	0162	18	64	FT	A	A		6991	0.551	0.551	0.061	0.250	0.137	0.001	0.360	0.360	-0.366	-0.283	-0.328	0.901	0.026	0.3
Science	4	558283	0163	18	65	FT	D	B		6991	0.256	0.239	0.342	0.161	0.256	0.001	0.117	-0.173	-0.048	-0.241	0.117	2.409	0.029	9.9
Science	4	558286	0164	18	66	FT	C	C		6991	0.722	0.097	0.076	0.722	0.103	0.002	0.404	-0.380	-0.265	0.404	-0.279	0.011	0.029	-2.6
Science	8	549631	0165	All	1	OP	D	D	2	136648	0.900	0.034	0.043	0.022	0.900	0.000	0.336	-0.158	-0.276	-0.203	0.336	-2.017	0.009	-9.9
Science	8	549777	0166	All	2	OP	A	C	2	136648	0.708	0.708	0.100	0.133	0.058	0.001	0.443	0.443	-0.360	-0.371	-0.241	-0.567	0.006	-9.9
Science	8	549665	0167	All	3	OP	C	D	2	136648	0.641	0.148	0.147	0.641	0.064	0.001	0.395	-0.385	-0.261	0.395	-0.271	-0.218	0.006	-9.9
Science	8	549682	0168	All	4	OP	D	A	2	136648	0.516	0.039	0.055	0.389	0.516	0.001	0.306	-0.309	-0.373	-0.238	0.306	0.382	0.006	9.9
Science	8	549741	0169	All	5	OP	C	A	2	136648	0.695	0.081	0.118	0.695	0.105	0.001	0.499	-0.390	-0.388	0.499	-0.369	-0.497	0.006	-9.9
Science	8	549688	0170	All	6	OP	B	A	2	136648	0.640	0.068	0.640	0.090	0.201	0.001	0.217	-0.272	0.217	-0.284	-0.052	-0.222	0.006	9.9
Science	8	549771	0171	All	7	OP	B	A	2	136648	0.652	0.146	0.652	0.154	0.047	0.001	0.326	-0.227	0.326	-0.260	-0.312	-0.277	0.006	9.9
Science	8	549807	0172	All	8	OP	B	D	1	136648	0.256	0.172	0.256	0.396	0.174	0.001	0.103	-0.203	0.103	-0.061	-0.144	1.690	0.007	9.9

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	8	549621	0173	All	9	OP	C	A	2	136648	0.623	0.169	0.139	0.623	0.068	0.001	0.442	-0.346	-0.368	0.442	-0.357	-0.125	0.006	-9.9
Science	8	549808	0174	All	10	OP	C	C	2	136648	0.645	0.051	0.064	0.645	0.239	0.001	0.311	-0.356	-0.342	0.311	-0.173	-0.240	0.006	9.9
Science	8	549757	0175	All	11	OP	B	A	2	136648	0.795	0.132	0.795	0.049	0.023	0.001	0.436	-0.367	0.436	-0.260	-0.275	-1.095	0.007	-9.9
Science	8	549879	0176	All	12	OP	A	D	2	136648	0.558	0.558	0.126	0.166	0.148	0.001	0.472	0.472	-0.378	-0.409	-0.416	0.186	0.006	-9.9
Science	8	549809	0177	All	13	OP	A	D	2	136648	0.709	0.709	0.216	0.043	0.031	0.001	0.386	0.386	-0.281	-0.347	-0.329	-0.578	0.006	-3.2
Science	8	549854	0178	All	14	OP	B	C	2	136648	0.540	0.219	0.540	0.107	0.133	0.001	0.527	-0.503	0.527	-0.395	-0.446	0.273	0.006	-9.9
Science	8	549694	0179	All	15	OP	B	C	2	136648	0.389	0.226	0.389	0.199	0.185	0.001	0.168	-0.086	0.168	-0.214	-0.197	0.982	0.006	9.9
Science	8	549750	0180	All	16	OP	D	A	2	136648	0.474	0.113	0.172	0.239	0.474	0.002	0.359	-0.371	-0.333	-0.295	0.359	0.581	0.006	9.9
Science	8	549579	0181	All	17	OP	C	D	2	136648	0.447	0.112	0.251	0.447	0.189	0.001	0.306	-0.349	-0.233	0.306	-0.297	0.707	0.006	9.9
Science	8	549798	0182	All	18	OP	A	A	1	136648	0.719	0.719	0.063	0.118	0.098	0.001	0.426	0.426	-0.372	-0.314	-0.286	-0.630	0.006	-9.9
Science	8	549636	0183	All	19	OP	B	A	2	136648	0.770	0.076	0.770	0.050	0.102	0.001	0.385	-0.294	0.385	-0.336	-0.222	-0.937	0.007	-9.0
Science	8	549841	0184	All	20	OP	C	C	2	136648	0.378	0.091	0.403	0.378	0.126	0.001	0.185	-0.292	-0.144	0.185	-0.180	1.038	0.006	9.9
Science	8	549783	0185	All	21	OP	A	A	2	136648	0.588	0.588	0.089	0.147	0.174	0.001	0.352	0.352	-0.292	-0.307	-0.278	0.040	0.006	9.9
Science	8	549774	0186	All	22	OP	B	A	2	136648	0.576	0.198	0.576	0.119	0.106	0.001	0.355	-0.271	0.355	-0.380	-0.232	0.095	0.006	9.9
Science	8	549751	0187	All	23	OP	C	B	1	136648	0.524	0.089	0.277	0.524	0.109	0.002	0.405	-0.377	-0.330	0.405	-0.365	0.347	0.006	-9.6
Science	8	549870	0188	All	24	OP	B	A	1	136648	0.640	0.106	0.640	0.119	0.132	0.002	0.485	-0.366	0.485	-0.414	-0.371	-0.213	0.006	-9.9
Science	8	549772	0189	All	25	OP	B	A	2	136648	0.380	0.127	0.380	0.097	0.394	0.002	0.301	-0.369	0.301	-0.450	-0.205	1.027	0.006	9.9
Science	8	549724	0190	All	38	OP	C	A	2	136648	0.690	0.066	0.083	0.690	0.160	0.000	0.433	-0.363	-0.372	0.433	-0.289	-0.469	0.006	-9.9
Science	8	549697	0191	All	39	OP	D	D	2	136648	0.303	0.120	0.189	0.388	0.303	0.000	0.251	-0.343	-0.317	-0.213	0.251	1.431	0.006	9.9
Science	8	549596	0192	All	40	OP	A	A	2	136648	0.647	0.647	0.104	0.140	0.107	0.001	0.429	0.429	-0.390	-0.311	-0.331	-0.247	0.006	-9.9
Science	8	549735	0193	All	41	OP	A	A	2	136648	0.776	0.776	0.045	0.066	0.112	0.000	0.426	0.426	-0.322	-0.299	-0.306	-0.969	0.007	-9.9
Science	8	549603	0194	All	42	OP	C	B	2	136648	0.793	0.042	0.093	0.793	0.072	0.000	0.402	-0.327	-0.263	0.402	-0.278	-1.085	0.007	-9.9
Science	8	549693	0195	All	43	OP	A	C	2	136648	0.505	0.505	0.237	0.142	0.115	0.001	0.408	0.408	-0.369	-0.348	-0.361	0.435	0.006	-9.9
Science	8	549664	0196	All	44	OP	B	B	2	136648	0.634	0.211	0.634	0.094	0.060	0.001	0.375	-0.262	0.375	-0.344	-0.337	-0.183	0.006	2.1
Science	8	549714	0197	All	45	OP	C	A	2	136648	0.679	0.122	0.117	0.679	0.081	0.001	0.490	-0.365	-0.388	0.490	-0.389	-0.410	0.006	-9.9
Science	8	549624	0198	All	46	OP	A	A	2	136648	0.469	0.469	0.136	0.236	0.158	0.001	0.227	0.227	-0.262	-0.180	-0.189	0.598	0.006	9.9
Science	8	549642	0199	All	47	OP	D	C	2	136648	0.735	0.072	0.064	0.128	0.735	0.001	0.512	-0.378	-0.403	-0.382	0.512	-0.716	0.007	-9.9
Science	8	549627	0200	All	48	OP	D	B	2	136648	0.580	0.098	0.144	0.177	0.580	0.001	0.527	-0.424	-0.467	-0.444	0.527	0.086	0.006	-9.9
Science	8	549787	0201	All	49	OP	A	A	2	136648	0.655	0.655	0.141	0.100	0.102	0.001	0.446	0.446	-0.345	-0.368	-0.339	-0.284	0.006	-9.9
Science	8	549712	0202	All	50	OP	D	C	2	136648	0.532	0.138	0.142	0.187	0.532	0.001	0.283	-0.108	-0.359	-0.260	0.283	0.307	0.006	9.9
Science	8	549756	0203	All	51	OP	A	A	2	136648	0.636	0.636	0.164	0.110	0.090	0.001	0.492	0.492	-0.418	-0.398	-0.352	-0.187	0.006	-9.9
Science	8	549685	0204	All	52	OP	C	A	2	136648	0.751	0.096	0.101	0.751	0.051	0.001	0.499	-0.360	-0.382	0.499	-0.364	-0.809	0.007	-9.9
Science	8	549720	0205	All	53	OP	D	A	2	136648	0.696	0.118	0.084	0.100	0.696	0.001	0.532	-0.377	-0.434	-0.428	0.532	-0.500	0.006	-9.9
Science	8	549668	0206	All	54	OP	D	A	2	136648	0.509	0.094	0.284	0.112	0.509	0.001	0.406	-0.427	-0.285	-0.439	0.406	0.419	0.006	-9.9
Science	8	549638	0207	All	55	OP	D	A	2	136648	0.324	0.159	0.359	0.155	0.324	0.002	0.304	-0.407	-0.224	-0.396	0.304	1.314	0.006	9.9
Science	8	549832	0208	All	56	OP	A	D	2	136648	0.726	0.726	0.091	0.131	0.051	0.002	0.510	0.510	-0.387	-0.396	-0.361	-0.664	0.007	-9.9
Science	8	549749	0209	All	57	OP	B	A	2	136648	0.569	0.116	0.569	0.184	0.128	0.002	0.359	-0.257	0.359	-0.299	-0.357	0.133	0.006	9.9
Science	8	549846	0210	All	58	OP	C	B	1	136648	0.774	0.084	0.073	0.774	0.067	0.001	0.536	-0.409	-0.386	0.536	-0.375	-0.953	0.007	-9.9
Science	8	549824	0211	All	59	OP	D	D	1	136648	0.465	0.323	0.130	0.080	0.465	0.001	0.198	-0.102	-0.229	-0.328	0.198	0.620	0.006	9.9
Science	8	549613	0212	All	60	OP	B	A	2	136648	0.555	0.179	0.555	0.134	0.130	0.001	0.447	-0.370	0.447	-0.361	-0.405	0.203	0.006	-9.9
Science	8	549761	0213	All	61	OP	D	A	2	136648	0.542	0.093	0.265	0.098	0.542	0.001	0.392	-0.416	-0.268	-0.415	0.392	0.265	0.006	-1.5
Science	8	549641	0214	All	62	OP	B	B	1	136648	0.730	0.093	0.730	0.099	0.076	0.001	0.469	-0.304	0.469	-0.426	-0.304	-0.692	0.007	-9.9
Science	8	549667	0215	All	63	OP	B	B	1	136648	0.504	0.288	0.504	0.081	0.125	0.002	0.351	-0.304	0.351	-0.346	-0.288	0.436	0.006	9.9

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	8	549782	0216	All	64	OP	C	C	1	136648	0.677	0.109	0.100	0.677	0.111	0.002	0.393	-0.356	-0.299	0.393	-0.245	-0.405	0.006	-6.1
Science	8	549672	0217	All	65	OP	B	A	2	136648	0.562	0.135	0.562	0.148	0.152	0.002	0.420	-0.339	0.420	-0.416	-0.292	0.166	0.006	-9.9
Science	8	549673	0218	All	66	OP	D	A	2	136648	0.337	0.244	0.189	0.227	0.337	0.002	0.307	-0.266	-0.339	-0.341	0.307	1.247	0.006	9.9
Science	8	549645	0219	All	67	OP	A	B	2	136648	0.470	0.470	0.208	0.193	0.127	0.002	0.432	0.432	-0.389	-0.376	-0.409	0.604	0.006	-9.9
Science	8	549644	0220	All	68	OP	C	B	2	136648	0.474	0.190	0.130	0.474	0.203	0.002	0.379	-0.343	-0.327	0.379	-0.347	0.577	0.006	2.8
Science	8	558287	0221	1	26	FT	B	A		7664	0.634	0.154	0.634	0.116	0.094	0.002	0.370	-0.292	0.370	-0.313	-0.252	-0.198	0.026	-4.4
Science	8	558289	0222	1	27	FT	C	D		7664	0.252	0.217	0.253	0.252	0.274	0.003	-0.015	0.058	-0.063	-0.015	0.065	1.694	0.028	9.9
Science	8	558293	0223	1	28	FT	A	A		7664	0.648	0.648	0.106	0.109	0.135	0.002	0.441	0.441	-0.345	-0.389	-0.309	-0.269	0.026	-8.8
Science	8	558295	0224	1	29	FT	A	D		7664	0.350	0.350	0.227	0.222	0.198	0.002	0.258	0.258	-0.276	-0.289	-0.209	1.159	0.026	8.0
Science	8	558288	0225	1	30	FT	A	D		7664	0.286	0.286	0.311	0.249	0.152	0.002	0.118	0.118	-0.143	-0.100	-0.130	1.499	0.027	9.9
Science	8	558292	0226	1	31	FT	C	A		7664	0.560	0.142	0.147	0.560	0.148	0.002	0.362	-0.293	-0.321	0.362	-0.294	0.156	0.025	-0.8
Science	8	558294	0227	1	32	FT	A	A		7664	0.302	0.302	0.221	0.327	0.147	0.003	0.110	0.110	-0.209	0.019	-0.276	1.410	0.027	9.9
Science	8	558290	0228	1	33	FT	A	C		7664	0.565	0.565	0.146	0.164	0.121	0.003	0.416	0.416	-0.318	-0.378	-0.347	0.133	0.025	-7.1
Science	8	549857	0229	1	69	MX	B	C	2	7664	0.389	0.236	0.389	0.248	0.124	0.002	0.225	-0.281	0.225	-0.231	-0.085	0.965	0.025	9.9
Science	8	549861	0230	1	70	MX	D	A	2	7664	0.551	0.120	0.195	0.132	0.551	0.002	0.391	-0.377	-0.276	-0.345	0.391	0.203	0.025	-5.3
Science	8	558298	0231	2	26	FT	A	A		7582	0.483	0.483	0.270	0.128	0.116	0.002	0.365	0.365	-0.339	-0.360	-0.267	0.527	0.025	-1.2
Science	8	558300	0232	2	27	FT	C	A		7582	0.602	0.125	0.101	0.602	0.170	0.002	0.350	-0.273	-0.392	0.350	-0.207	-0.040	0.025	-0.4
Science	8	558302	0233	2	28	FT	D	C		7582	0.550	0.145	0.221	0.081	0.550	0.003	0.403	-0.389	-0.271	-0.406	0.403	0.213	0.025	-4.9
Science	8	558304	0234	2	29	FT	D	C		7582	0.208	0.374	0.145	0.271	0.208	0.003	0.089	0.041	-0.313	-0.186	0.089	1.996	0.030	9.9
Science	8	558296	0235	2	30	FT	C	D		7582	0.337	0.249	0.260	0.337	0.150	0.004	0.184	-0.313	-0.035	0.184	-0.248	1.241	0.026	9.9
Science	8	558297	0236	2	31	FT	A	B		7582	0.483	0.483	0.237	0.242	0.035	0.002	0.205	0.205	-0.207	-0.139	-0.259	0.529	0.025	9.9
Science	8	558299	0237	2	32	FT	C	A		7582	0.793	0.079	0.055	0.793	0.071	0.003	0.442	-0.278	-0.339	0.442	-0.312	-1.106	0.030	-9.4
Science	8	558301	0238	2	33	FT	B	C		7582	0.747	0.066	0.747	0.140	0.042	0.004	0.298	-0.241	0.298	-0.177	-0.262	-0.814	0.028	1.5
Science	8	549670	0239	2	69	MX	A	D	1	7582	0.377	0.377	0.328	0.128	0.165	0.002	0.276	0.276	-0.217	-0.346	-0.303	1.039	0.026	7.1
Science	8	549878	0240	2	70	MX	B	A	2	7582	0.580	0.119	0.580	0.155	0.144	0.002	0.463	-0.356	0.463	-0.431	-0.359	0.070	0.025	-9.9
Science	8	558306	0241	3	26	FT	B	A		7544	0.746	0.089	0.746	0.094	0.069	0.002	0.408	-0.338	0.408	-0.320	-0.210	-0.786	0.028	-6.2
Science	8	558309	0242	3	27	FT	B	A		7544	0.469	0.150	0.469	0.220	0.157	0.003	0.320	-0.262	0.320	-0.332	-0.246	0.609	0.025	4.0
Science	8	558311	0243	3	28	FT	B	C		7544	0.346	0.334	0.346	0.209	0.109	0.002	0.091	-0.064	0.091	-0.069	-0.189	1.208	0.026	9.9
Science	8	558314	0244	3	29	FT	A	C		7544	0.663	0.663	0.113	0.103	0.118	0.002	0.495	0.495	-0.361	-0.411	-0.387	-0.324	0.026	-9.9
Science	8	558305	0245	3	30	FT	C	A		7544	0.304	0.317	0.114	0.304	0.261	0.003	0.101	-0.062	-0.275	0.101	-0.068	1.426	0.027	9.9
Science	8	558307	0246	3	31	FT	A	D		7544	0.359	0.359	0.071	0.473	0.093	0.003	0.273	0.273	-0.459	-0.183	-0.418	1.142	0.026	6.5
Science	8	558308	0247	3	32	FT	A	C		7544	0.691	0.691	0.085	0.111	0.111	0.003	0.382	0.382	-0.384	-0.250	-0.243	-0.472	0.027	-4.1
Science	8	558312	0248	3	33	FT	B	B		7544	0.570	0.089	0.570	0.099	0.239	0.003	0.413	-0.394	0.413	-0.423	-0.280	0.135	0.025	-6.5
Science	8	549834	0249	3	69	MX	D	C	2	7544	0.535	0.119	0.117	0.227	0.535	0.002	0.338	-0.345	-0.342	-0.215	0.338	0.301	0.025	1.2
Science	8	549655	0250	3	70	MX	D	B	2	7544	0.400	0.139	0.277	0.181	0.400	0.002	0.253	-0.305	-0.214	-0.229	0.253	0.941	0.025	8.9
Science	8	558315	0251	4	26	FT	D	A		7605	0.601	0.052	0.127	0.217	0.601	0.002	0.282	-0.351	-0.324	-0.125	0.282	-0.005	0.025	7.1
Science	8	558318	0252	4	27	FT	A	A		7605	0.476	0.476	0.238	0.170	0.112	0.003	0.251	0.251	-0.160	-0.316	-0.220	0.583	0.025	8.9
Science	8	558321	0253	4	28	FT	A	C		7605	0.394	0.394	0.216	0.127	0.260	0.002	0.289	0.289	-0.289	-0.354	-0.217	0.973	0.025	6.3
Science	8	558324	0254	4	29	FT	C	C		7605	0.513	0.168	0.187	0.513	0.128	0.003	0.345	-0.296	-0.309	0.345	-0.281	0.411	0.025	-0.6
Science	8	558313	0255	4	30	FT	C	D		7605	0.475	0.158	0.224	0.475	0.140	0.003	0.200	-0.188	-0.211	0.200	-0.121	0.588	0.025	9.9
Science	8	558316	0256	4	31	FT	D	B		7605	0.436	0.152	0.159	0.250	0.436	0.003	0.391	-0.419	-0.349	-0.340	0.391	0.772	0.025	-4.7
Science	8	558317	0257	4	32	FT	B	C		7605	0.330	0.183	0.330	0.140	0.344	0.004	0.022	-0.017	0.022	-0.151	0.041	1.291	0.026	9.9
Science	8	558319	0258	4	33	FT	C	A		7605	0.752	0.067	0.073	0.752	0.104	0.004	0.360	-0.284	-0.294	0.360	-0.210	-0.810	0.028	-2.2

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	8	549743	0259	4	69	MX	A	D	2	7605	0.581	0.581	0.140	0.161	0.115	0.002	0.460	0.460	-0.429	-0.397	-0.308	0.088	0.025	-9.9
Science	8	549806	0260	4	70	MX	D	A	2	7605	0.390	0.360	0.114	0.133	0.390	0.003	0.277	-0.132	-0.430	-0.418	0.277	0.992	0.025	5.8
Science	8	558323	0261	5	26	FT	A	A		7581	0.453	0.453	0.123	0.219	0.203	0.002	0.402	0.402	-0.473	-0.407	-0.255	0.685	0.025	-3.8
Science	8	558329	0262	5	27	FT	D	A		7581	0.225	0.296	0.327	0.151	0.225	0.002	0.209	-0.239	-0.168	-0.376	0.209	1.903	0.029	9.9
Science	8	558331	0263	5	28	FT	A	D		7581	0.560	0.560	0.174	0.123	0.140	0.003	0.294	0.294	-0.245	-0.332	-0.180	0.177	0.025	9.6
Science	8	558333	0264	5	29	FT	D	D		7581	0.735	0.059	0.104	0.099	0.735	0.003	0.498	-0.373	-0.394	-0.345	0.498	-0.735	0.028	-9.9
Science	8	558322	0265	5	30	FT	B	B		7581	0.515	0.113	0.515	0.228	0.140	0.003	0.406	-0.436	0.406	-0.312	-0.346	0.388	0.025	-4.7
Science	8	558325	0266	5	31	FT	D	C		7581	0.371	0.303	0.205	0.118	0.371	0.003	0.252	-0.170	-0.303	-0.326	0.252	1.082	0.026	8.0
Science	8	558326	0267	5	32	FT	B	A		7581	0.495	0.173	0.495	0.139	0.191	0.002	0.324	-0.256	0.324	-0.368	-0.245	0.482	0.025	3.4
Science	8	558327	0268	5	33	FT	A	D		7581	0.486	0.486	0.140	0.126	0.244	0.004	0.392	0.392	-0.369	-0.416	-0.295	0.528	0.025	-3.3
Science	8	549855	0269	5	69	MX	A	B	2	7581	0.565	0.565	0.165	0.156	0.112	0.002	0.482	0.482	-0.398	-0.426	-0.380	0.150	0.025	-9.9
Science	8	549811	0270	5	70	MX	A	C	2	7581	0.404	0.404	0.113	0.217	0.262	0.003	0.279	0.279	-0.435	-0.291	-0.149	0.917	0.025	8.2
Science	8	558335	0271	6	26	FT	D	A		7561	0.489	0.208	0.223	0.078	0.489	0.002	0.372	-0.373	-0.233	-0.478	0.372	0.503	0.025	-2.6
Science	8	558338	0272	6	27	FT	B	A		7561	0.350	0.123	0.350	0.128	0.397	0.002	0.151	-0.311	0.151	-0.407	0.014	1.173	0.026	9.9
Science	8	558340	0273	6	28	FT	A	D		7561	0.403	0.403	0.063	0.472	0.060	0.002	0.202	0.202	-0.353	-0.136	-0.278	0.910	0.025	9.9
Science	8	558341	0274	6	29	FT	B	D		7561	0.676	0.105	0.676	0.125	0.092	0.002	0.429	-0.322	0.429	-0.308	-0.366	-0.405	0.026	-6.9
Science	8	558330	0275	6	30	FT	D	A		7561	0.140	0.127	0.094	0.637	0.140	0.002	0.101	-0.351	-0.359	-0.041	0.101	2.512	0.035	9.9
Science	8	558332	0276	6	31	FT	D	D		7561	0.237	0.391	0.141	0.228	0.237	0.002	0.179	-0.091	-0.382	-0.263	0.179	1.799	0.029	9.9
Science	8	558334	0277	6	32	FT	C	B		7561	0.617	0.096	0.171	0.617	0.113	0.003	0.461	-0.383	-0.350	0.461	-0.390	-0.104	0.026	-9.8
Science	8	558337	0278	6	33	FT	B	C		7561	0.213	0.204	0.213	0.164	0.416	0.003	0.019	-0.113	0.019	-0.094	0.055	1.955	0.030	9.9
Science	8	549823	0279	6	69	MX	B	C	2	7561	0.620	0.104	0.620	0.105	0.168	0.003	0.383	-0.370	0.383	-0.375	-0.201	-0.117	0.026	-3.7
Science	8	549822	0280	6	70	MX	C	A	2	7561	0.569	0.098	0.188	0.569	0.142	0.004	0.379	-0.403	-0.232	0.379	-0.353	0.130	0.025	-2.3
Science	8	558345	0281	7	26	FT	D	D		7579	0.259	0.583	0.069	0.086	0.259	0.002	0.191	-0.135	-0.412	-0.381	0.191	1.659	0.028	9.9
Science	8	558342	0282	7	27	FT	C	A		7579	0.440	0.321	0.117	0.440	0.119	0.003	0.200	-0.030	-0.379	0.200	-0.300	0.730	0.025	9.9
Science	8	558347	0283	7	28	FT	D	A		7579	0.315	0.194	0.265	0.221	0.315	0.003	0.187	-0.228	-0.220	-0.145	0.187	1.350	0.026	9.9
Science	8	558351	0284	7	29	FT	C	C		7579	0.522	0.185	0.147	0.522	0.143	0.002	0.237	-0.184	-0.203	0.237	-0.215	0.350	0.025	8.8
Science	8	558339	0285	7	30	FT	C	A		7579	0.641	0.138	0.190	0.641	0.027	0.003	0.311	-0.350	-0.156	0.311	-0.283	-0.220	0.026	2.3
Science	8	558343	0286	7	31	FT	A	B		7579	0.288	0.288	0.431	0.116	0.161	0.003	0.175	0.175	-0.136	-0.399	-0.121	1.497	0.027	9.9
Science	8	558344	0287	7	32	FT	C	D		7579	0.418	0.210	0.176	0.418	0.192	0.003	0.193	-0.133	-0.231	0.193	-0.181	0.835	0.025	9.9
Science	8	558348	0288	7	33	FT	D	C		7579	0.505	0.201	0.186	0.104	0.505	0.004	0.283	-0.190	-0.233	-0.361	0.283	0.429	0.025	5.9
Science	8	549882	0289	7	69	MX	A	B	2	7579	0.476	0.476	0.222	0.138	0.162	0.002	0.322	0.322	-0.280	-0.372	-0.216	0.564	0.025	1.4
Science	8	549804	0290	7	70	MX	A	C	3	7579	0.302	0.302	0.138	0.384	0.174	0.003	0.207	0.207	-0.306	-0.148	-0.275	1.420	0.027	9.9
Science	8	558352	0291	8	26	FT	B	A		7575	0.308	0.211	0.308	0.412	0.066	0.003	0.124	-0.166	0.124	-0.061	-0.309	1.401	0.027	9.9
Science	8	558355	0292	8	27	FT	A	C		7575	0.328	0.328	0.273	0.113	0.283	0.003	0.177	0.177	-0.206	-0.248	-0.116	1.294	0.026	9.9
Science	8	558358	0293	8	28	FT	B	A		7575	0.477	0.148	0.477	0.201	0.172	0.003	0.366	-0.286	0.366	-0.372	-0.304	0.573	0.025	-2.2
Science	8	558359	0294	8	29	FT	B	C		7575	0.497	0.211	0.497	0.185	0.104	0.002	0.144	-0.041	0.144	-0.159	-0.215	0.480	0.025	9.9
Science	8	558349	0295	8	30	FT	D	C		7575	0.679	0.100	0.097	0.120	0.679	0.003	0.353	-0.256	-0.307	-0.232	0.353	-0.406	0.026	-4.0
Science	8	558350	0296	8	31	FT	A	D		7575	0.527	0.527	0.186	0.115	0.169	0.003	0.450	0.450	-0.421	-0.455	-0.289	0.337	0.025	-9.9
Science	8	558354	0297	8	32	FT	C	A		7575	0.381	0.169	0.333	0.381	0.113	0.003	0.176	-0.240	-0.082	0.176	-0.271	1.027	0.025	9.9
Science	8	558353	0298	8	33	FT	B	B		7575	0.565	0.091	0.565	0.113	0.228	0.003	0.408	-0.415	0.408	-0.396	-0.265	0.159	0.025	-7.6
Science	8	549856	0299	8	69	MX	C	D	2	7575	0.297	0.260	0.232	0.297	0.209	0.002	0.114	-0.112	-0.204	0.114	-0.037	1.460	0.027	9.9
Science	8	554284	0300	8	70	MX	B	C	2	7575	0.386	0.187	0.386	0.243	0.182	0.003	0.236	-0.309	0.236	-0.176	-0.209	1.005	0.025	8.6
Science	8	558360	0301	9	26	FT	A	A		7605	0.587	0.587	0.111	0.133	0.167	0.002	0.417	0.417	-0.402	-0.420	-0.231	0.060	0.025	-6.7

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	8	558363	0302	9	27	FT	B	A		7605	0.566	0.171	0.566	0.184	0.077	0.002	0.305	-0.174	0.305	-0.290	-0.341	0.158	0.025	6.7
Science	8	558365	0303	9	28	FT	D	D		7605	0.484	0.232	0.193	0.088	0.484	0.002	0.326	-0.296	-0.257	-0.336	0.326	0.549	0.025	2.2
Science	8	558369	0304	9	29	FT	B	D		7605	0.422	0.221	0.422	0.133	0.222	0.002	0.266	-0.235	0.266	-0.357	-0.182	0.845	0.025	8.8
Science	8	558357	0305	9	30	FT	D	C		7605	0.698	0.121	0.105	0.074	0.698	0.002	0.518	-0.414	-0.409	-0.367	0.518	-0.507	0.027	-9.9
Science	8	558361	0306	9	31	FT	B	D		7605	0.305	0.298	0.305	0.250	0.144	0.002	0.177	-0.191	0.177	-0.133	-0.251	1.437	0.027	9.9
Science	8	558362	0307	9	32	FT	A	A		7605	0.769	0.769	0.066	0.110	0.053	0.003	0.476	0.476	-0.359	-0.333	-0.342	-0.925	0.029	-9.9
Science	8	558364	0308	9	33	FT	A	B		7605	0.585	0.585	0.180	0.131	0.101	0.003	0.430	0.430	-0.349	-0.371	-0.350	0.068	0.025	-7.9
Science	8	549833	0309	9	69	MX	C	A	2	7605	0.498	0.229	0.104	0.498	0.167	0.002	0.341	-0.254	-0.390	0.341	-0.286	0.484	0.025	0.6
Science	8	549867	0310	9	70	MX	A	B	2	7605	0.505	0.505	0.161	0.166	0.166	0.002	0.311	0.311	-0.241	-0.267	-0.306	0.448	0.025	4.0
Science	8	558370	0311	10	26	FT	C	A		7583	0.692	0.101	0.089	0.692	0.115	0.003	0.424	-0.321	-0.420	0.424	-0.237	-0.491	0.027	-6.7
Science	8	558373	0312	10	27	FT	A	A		7583	0.461	0.461	0.208	0.120	0.208	0.002	0.298	0.298	-0.279	-0.390	-0.187	0.641	0.025	5.2
Science	8	558376	0313	10	28	FT	A	D		7583	0.542	0.542	0.196	0.095	0.164	0.003	0.290	0.290	-0.227	-0.363	-0.185	0.261	0.025	6.9
Science	8	558377	0314	10	29	FT	B	D		7583	0.638	0.160	0.638	0.093	0.106	0.002	0.438	-0.276	0.438	-0.418	-0.367	-0.205	0.026	-8.9
Science	8	558367	0315	10	30	FT	A	D		7583	0.254	0.254	0.161	0.181	0.401	0.002	0.005	0.005	-0.150	-0.269	0.176	1.710	0.028	9.9
Science	8	558368	0316	10	31	FT	B	B		7583	0.465	0.104	0.465	0.374	0.054	0.003	0.353	-0.394	0.353	-0.269	-0.411	0.622	0.025	0.5
Science	8	558372	0317	10	32	FT	D	A		7583	0.326	0.396	0.121	0.153	0.326	0.003	0.201	-0.132	-0.305	-0.286	0.201	1.304	0.026	9.9
Science	8	558371	0318	10	33	FT	C	C		7583	0.394	0.155	0.140	0.394	0.308	0.004	0.264	-0.281	-0.289	0.264	-0.218	0.964	0.025	7.7
Science	8	549831	0319	10	69	MX	D	A	1	7583	0.582	0.147	0.138	0.131	0.582	0.003	0.532	-0.441	-0.483	-0.414	0.532	0.070	0.025	-9.9
Science	8	549686	0320	10	70	MX	C	C	2	7583	0.595	0.098	0.201	0.595	0.103	0.002	0.440	-0.399	-0.327	0.440	-0.383	0.006	0.025	-8.6
Science	8	558378	0321	11	26	FT	B	A		7597	0.430	0.214	0.430	0.166	0.188	0.002	0.378	-0.393	0.378	-0.378	-0.269	0.794	0.025	-2.4
Science	8	558380	0322	11	27	FT	A	A		7597	0.569	0.569	0.140	0.118	0.170	0.002	0.337	0.337	-0.227	-0.356	-0.271	0.140	0.025	1.5
Science	8	558384	0323	11	28	FT	D	B		7597	0.608	0.189	0.074	0.127	0.608	0.002	0.299	-0.130	-0.341	-0.320	0.299	-0.046	0.025	4.5
Science	8	558385	0324	11	29	FT	B	B		7597	0.319	0.348	0.319	0.271	0.060	0.002	0.158	-0.229	0.158	-0.034	-0.309	1.339	0.026	9.9
Science	8	558375	0325	11	30	FT	B	C		7597	0.727	0.100	0.727	0.138	0.033	0.001	0.129	-0.194	0.129	0.014	-0.143	-0.669	0.027	9.9
Science	8	558379	0326	11	31	FT	C	A		7597	0.676	0.086	0.140	0.676	0.096	0.002	0.374	-0.228	-0.312	0.374	-0.298	-0.389	0.026	-3.8
Science	8	558381	0327	11	32	FT	D	D		7597	0.407	0.139	0.185	0.266	0.407	0.002	0.282	-0.291	-0.317	-0.217	0.282	0.902	0.025	5.4
Science	8	558383	0328	11	33	FT	B	B		7597	0.345	0.077	0.345	0.388	0.187	0.004	0.040	-0.353	0.040	0.136	-0.230	1.207	0.026	9.9
Science	8	549801	0329	11	69	MX	C	D	2	7597	0.463	0.149	0.123	0.463	0.262	0.003	0.326	-0.285	-0.422	0.326	-0.220	0.638	0.025	1.8
Science	8	549848	0330	11	70	MX	D	A	2	7597	0.485	0.160	0.150	0.203	0.485	0.002	0.404	-0.377	-0.371	-0.332	0.404	0.533	0.025	-5.7
Science	8	558388	0331	12	26	FT	D	A		7607	0.124	0.203	0.477	0.193	0.124	0.003	-0.066	0.019	0.175	-0.097	-0.066	2.688	0.036	9.9
Science	8	558390	0332	12	27	FT	B	B		7607	0.460	0.116	0.460	0.122	0.300	0.002	0.277	-0.268	0.277	-0.342	-0.197	0.653	0.025	6.1
Science	8	558392	0333	12	28	FT	C	A		7607	0.530	0.149	0.163	0.530	0.156	0.002	0.328	-0.321	-0.299	0.328	-0.217	0.323	0.025	2.2
Science	8	558393	0334	12	29	FT	B	B		7607	0.281	0.209	0.281	0.213	0.294	0.002	0.181	-0.187	0.181	-0.362	-0.069	1.554	0.027	9.9
Science	8	558386	0335	12	30	FT	A	A		7607	0.446	0.446	0.073	0.109	0.369	0.002	0.206	0.206	-0.349	-0.353	-0.077	0.718	0.025	9.9
Science	8	558387	0336	12	31	FT	B	C		7607	0.474	0.133	0.474	0.200	0.190	0.002	0.407	-0.404	0.407	-0.389	-0.304	0.584	0.025	-4.7
Science	8	558389	0337	12	32	FT	B	D		7607	0.804	0.069	0.804	0.056	0.068	0.002	0.463	-0.348	0.463	-0.319	-0.302	-1.162	0.031	-9.9
Science	8	558395	0338	12	33	FT	C	B		7607	0.762	0.121	0.049	0.762	0.065	0.004	0.363	-0.162	-0.352	0.363	-0.338	-0.883	0.029	0.1
Science	8	549837	0339	12	69	MX	C	A	2	7607	0.602	0.121	0.190	0.602	0.085	0.002	0.454	-0.402	-0.353	0.454	-0.353	-0.018	0.025	-9.9
Science	8	549795	0340	12	70	MX	D	C	2	7607	0.413	0.127	0.321	0.136	0.413	0.002	0.295	-0.362	-0.188	-0.376	0.295	0.873	0.025	5.0
Science	8	558397	0341	13	26	FT	B	A		7567	0.358	0.139	0.358	0.143	0.358	0.002	0.178	-0.265	0.178	-0.278	-0.084	1.151	0.026	9.9
Science	8	558401	0342	13	27	FT	C	A		7567	0.256	0.138	0.279	0.256	0.324	0.002	0.041	-0.260	-0.016	0.041	0.029	1.698	0.028	9.9
Science	8	558398	0343	13	28	FT	D	B		7567	0.348	0.218	0.190	0.240	0.348	0.002	0.279	-0.199	-0.319	-0.340	0.279	1.199	0.026	4.0
Science	8	558404	0344	13	29	FT	B	B		7567	0.407	0.064	0.407	0.285	0.242	0.003	0.185	-0.390	0.185	-0.191	-0.077	0.915	0.025	9.9

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	8	558394	0345	13	30	FT	C	D		7567	0.879	0.034	0.049	0.879	0.036	0.002	0.420	-0.282	-0.294	0.420	-0.244	-1.781	0.037	-9.9
Science	8	558396	0346	13	31	FT	D	B		7567	0.601	0.083	0.176	0.138	0.601	0.002	0.435	-0.422	-0.389	-0.285	0.435	-0.001	0.025	-8.3
Science	8	558400	0347	13	32	FT	B	A		7567	0.299	0.460	0.299	0.134	0.104	0.002	0.020	0.116	0.020	-0.264	-0.238	1.460	0.027	9.9
Science	8	558399	0348	13	33	FT	D	C		7567	0.371	0.236	0.156	0.235	0.371	0.002	0.187	-0.196	-0.124	-0.214	0.187	1.087	0.026	9.9
Science	8	549813	0349	13	69	MX	B	A	2	7567	0.496	0.152	0.496	0.192	0.156	0.003	0.300	-0.302	0.300	-0.313	-0.166	0.491	0.025	4.2
Science	8	549830	0350	13	70	MX	D	D	2	7567	0.786	0.069	0.071	0.072	0.786	0.001	0.454	-0.319	-0.364	-0.282	0.454	-1.027	0.030	-9.9
Science	8	558406	0351	14	26	FT	A	A		7596	0.195	0.195	0.424	0.103	0.275	0.002	-0.032	-0.032	0.003	-0.308	0.238	2.092	0.031	9.9
Science	8	558409	0352	14	27	FT	C	A		7596	0.571	0.113	0.230	0.571	0.083	0.003	0.332	-0.329	-0.198	0.332	-0.372	0.127	0.025	3.4
Science	8	558411	0353	14	28	FT	C	B		7596	0.345	0.344	0.184	0.345	0.124	0.003	0.249	-0.157	-0.320	0.249	-0.346	1.210	0.026	9.9
Science	8	558413	0354	14	29	FT	A	B		7596	0.312	0.312	0.318	0.141	0.226	0.003	0.190	0.190	-0.105	-0.349	-0.212	1.386	0.027	9.9
Science	8	558403	0355	14	30	FT	A	C		7596	0.761	0.761	0.127	0.068	0.041	0.002	0.481	0.481	-0.385	-0.340	-0.306	-0.888	0.029	-9.9
Science	8	558405	0356	14	31	FT	D	D		7596	0.777	0.035	0.050	0.135	0.777	0.003	0.370	-0.325	-0.311	-0.216	0.370	-0.987	0.029	-3.5
Science	8	558408	0357	14	32	FT	D	A		7596	0.489	0.144	0.225	0.139	0.489	0.003	0.415	-0.324	-0.422	-0.357	0.415	0.516	0.025	-5.5
Science	8	558407	0358	14	33	FT	C	B		7596	0.568	0.065	0.232	0.568	0.129	0.005	0.324	-0.384	-0.240	0.324	-0.236	0.138	0.025	1.2
Science	8	549838	0359	14	69	MX	C	C	3	7596	0.408	0.225	0.204	0.408	0.162	0.001	0.297	-0.291	-0.338	0.297	-0.194	0.899	0.025	6.3
Science	8	549800	0360	14	70	MX	C	B	1	7596	0.858	0.038	0.055	0.858	0.048	0.001	0.370	-0.277	-0.215	0.370	-0.249	-1.600	0.034	-5.4
Science	8	558414	0361	15	26	FT	C	A		7588	0.642	0.123	0.087	0.642	0.145	0.002	0.399	-0.302	-0.402	0.399	-0.260	-0.218	0.026	-5.5
Science	8	558417	0362	15	27	FT	A	A		7588	0.401	0.401	0.245	0.136	0.217	0.002	0.296	0.296	-0.343	-0.328	-0.187	0.929	0.025	4.2
Science	8	558418	0363	15	28	FT	C	D		7588	0.626	0.199	0.107	0.626	0.066	0.002	0.269	-0.139	-0.261	0.269	-0.305	-0.139	0.026	6.6
Science	8	558423	0364	15	29	FT	C	D		7588	0.339	0.264	0.176	0.339	0.218	0.002	0.131	-0.157	-0.173	0.131	-0.074	1.236	0.026	9.9
Science	8	558412	0365	15	30	FT	D	D		7588	0.224	0.201	0.246	0.327	0.224	0.002	0.108	-0.178	-0.180	-0.056	0.108	1.888	0.029	9.9
Science	8	558416	0366	15	31	FT	C	C		7588	0.249	0.219	0.225	0.249	0.305	0.002	0.153	-0.143	-0.306	0.153	-0.085	1.732	0.028	9.9
Science	8	558415	0367	15	32	FT	A	B		7588	0.545	0.545	0.192	0.150	0.109	0.003	0.424	0.424	-0.311	-0.411	-0.361	0.250	0.025	-8.2
Science	8	558419	0368	15	33	FT	B	A		7588	0.550	0.088	0.550	0.115	0.243	0.004	0.378	-0.412	0.378	-0.421	-0.218	0.230	0.025	-3.6
Science	8	549752	0369	15	69	MX	C	C	2	7588	0.555	0.132	0.179	0.555	0.130	0.003	0.398	-0.371	-0.354	0.398	-0.263	0.202	0.025	-5.8
Science	8	549656	0370	15	70	MX	B	D	2	7588	0.553	0.127	0.553	0.106	0.212	0.002	0.366	-0.311	0.366	-0.424	-0.244	0.215	0.025	-1.5
Science	8	558424	0371	16	26	FT	C	A		7601	0.226	0.249	0.325	0.226	0.199	0.001	0.027	-0.101	-0.016	0.027	0.033	1.867	0.029	9.9
Science	8	558450	0372	16	27	FT	C	A		7601	0.248	0.290	0.292	0.248	0.167	0.002	-0.007	0.093	-0.061	-0.007	-0.015	1.731	0.028	9.9
Science	8	558430	0373	16	28	FT	D	C		7601	0.462	0.146	0.164	0.226	0.462	0.002	0.386	-0.376	-0.372	-0.323	0.386	0.637	0.025	-4.8
Science	8	558427	0374	16	29	FT	C	C		7601	0.441	0.112	0.344	0.441	0.100	0.002	0.228	-0.379	-0.077	0.228	-0.360	0.733	0.025	9.9
Science	8	558422	0375	16	30	FT	B	A		7601	0.764	0.051	0.764	0.096	0.087	0.002	0.425	-0.282	0.425	-0.301	-0.321	-0.889	0.029	-8.8
Science	8	558421	0376	16	31	FT	C	C		7601	0.341	0.187	0.226	0.341	0.244	0.002	0.203	-0.280	-0.160	0.203	-0.181	1.221	0.026	9.9
Science	8	558425	0377	16	32	FT	D	D		7601	0.499	0.109	0.117	0.271	0.499	0.003	0.399	-0.415	-0.391	-0.301	0.399	0.463	0.025	-6.2
Science	8	558428	0378	16	33	FT	B	B		7601	0.491	0.296	0.491	0.097	0.113	0.003	0.160	-0.053	0.160	-0.293	-0.183	0.503	0.025	9.9
Science	8	549599	0379	16	69	MX	C	C	2	7601	0.541	0.095	0.153	0.541	0.208	0.003	0.235	-0.305	-0.315	0.235	-0.042	0.271	0.025	9.2
Science	8	549708	0380	16	70	MX	D	D	2	7601	0.419	0.136	0.184	0.257	0.419	0.003	0.289	-0.355	-0.319	-0.190	0.289	0.837	0.025	4.6
Science	8	558434	0381	17	26	FT	B	A		7602	0.274	0.226	0.274	0.272	0.227	0.001	0.172	-0.240	0.172	-0.215	-0.084	1.555	0.027	9.9
Science	8	558432	0382	17	27	FT	A	A		7602	0.516	0.516	0.260	0.128	0.095	0.001	0.317	0.317	-0.216	-0.348	-0.299	0.361	0.025	1.4
Science	8	558437	0383	17	28	FT	B	D		7602	0.274	0.359	0.274	0.146	0.217	0.003	-0.048	0.141	-0.048	-0.201	0.092	1.552	0.027	9.9
Science	8	558440	0384	17	29	FT	C	D		7602	0.380	0.178	0.285	0.380	0.155	0.002	0.074	-0.073	-0.063	0.074	-0.081	0.999	0.025	9.9
Science	8	558431	0385	17	30	FT	D	A		7602	0.381	0.182	0.189	0.245	0.381	0.002	0.198	-0.177	-0.220	-0.190	0.198	0.996	0.025	9.9
Science	8	558429	0386	17	31	FT	B	B		7602	0.619	0.259	0.619	0.070	0.051	0.001	0.424	-0.334	0.424	-0.354	-0.371	-0.126	0.025	-8.9
Science	8	558433	0387	17	32	FT	A	C		7602	0.360	0.360	0.185	0.235	0.219	0.002	0.267	0.267	-0.275	-0.317	-0.196	1.100	0.026	6.1

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	8	558435	0388	17	33	FT	D	D		7602	0.294	0.127	0.475	0.101	0.294	0.003	0.165	-0.396	-0.034	-0.432	0.165	1.440	0.027	9.9
Science	8	549790	0389	17	69	MX	C	A	2	7602	0.485	0.114	0.311	0.485	0.088	0.002	0.308	-0.351	-0.194	0.308	-0.377	0.506	0.025	2.9
Science	8	549726	0390	17	70	MX	B	D	1	7602	0.425	0.415	0.425	0.082	0.076	0.002	0.264	-0.157	0.264	-0.424	-0.377	0.785	0.025	6.3
Science	8	558441	0391	18	26	FT	B	A		7611	0.217	0.268	0.217	0.088	0.425	0.002	-0.090	0.053	-0.090	-0.347	0.245	1.946	0.029	9.9
Science	8	558444	0392	18	27	FT	A	A		7611	0.400	0.400	0.217	0.225	0.155	0.002	0.351	0.351	-0.379	-0.315	-0.297	0.939	0.025	0.6
Science	8	558445	0393	18	28	FT	C	C		7611	0.335	0.262	0.224	0.335	0.175	0.003	0.096	-0.074	-0.125	0.096	-0.088	1.263	0.026	9.9
Science	8	558447	0394	18	29	FT	D	C		7611	0.408	0.103	0.258	0.227	0.408	0.003	0.175	-0.224	-0.176	-0.127	0.175	0.904	0.025	9.9
Science	8	558438	0395	18	30	FT	B	C		7611	0.374	0.138	0.374	0.238	0.247	0.003	0.189	-0.226	0.189	-0.188	-0.144	1.068	0.026	9.9
Science	8	558439	0396	18	31	FT	D	D		7611	0.673	0.082	0.114	0.127	0.673	0.003	0.484	-0.383	-0.369	-0.382	0.484	-0.375	0.026	-9.9
Science	8	558442	0397	18	32	FT	C	B		7611	0.590	0.182	0.151	0.590	0.073	0.003	0.475	-0.367	-0.432	0.475	-0.371	0.041	0.025	-9.9
Science	8	558443	0398	18	33	FT	A	A		7611	0.511	0.511	0.121	0.148	0.217	0.004	0.325	0.325	-0.386	-0.324	-0.185	0.417	0.025	2.2
Science	8	549713	0399	18	69	MX	A	D	2	7611	0.749	0.749	0.112	0.079	0.058	0.002	0.531	0.531	-0.395	-0.420	-0.363	-0.802	0.028	-9.9
Science	8	549725	0400	18	70	MX	D	A	2	7611	0.607	0.138	0.103	0.150	0.607	0.002	0.482	-0.380	-0.408	-0.395	0.482	-0.042	0.025	-9.9
Science	11	554450	0401	All	1	OP	D	C	2	129722	0.711	0.027	0.200	0.060	0.711	0.001	0.350	-0.208	-0.294	-0.241	0.350	-0.609	0.007	-4.1
Science	11	554318	0402	All	2	OP	B	A	2	129722	0.851	0.066	0.851	0.061	0.020	0.002	0.304	-0.126	0.304	-0.281	-0.203	-1.523	0.008	-4.4
Science	11	554474	0403	All	3	OP	A	C	2	129722	0.508	0.508	0.044	0.129	0.315	0.004	0.325	0.325	-0.306	-0.271	-0.281	0.366	0.006	9.9
Science	11	554297	0404	All	4	OP	A	A	2	129722	0.568	0.568	0.196	0.189	0.047	0.001	0.244	0.244	-0.122	-0.261	-0.274	0.096	0.006	9.9
Science	11	554457	0405	All	5	OP	D	D	2	129722	0.554	0.103	0.156	0.185	0.554	0.003	0.390	-0.240	-0.289	-0.406	0.390	0.164	0.006	-9.9
Science	11	554426	0406	All	6	OP	D	A	1	129722	0.832	0.027	0.123	0.017	0.832	0.001	0.283	-0.247	-0.181	-0.208	0.283	-1.372	0.008	6.8
Science	11	554425	0407	All	7	OP	A	A	2	129722	0.472	0.472	0.098	0.359	0.069	0.003	0.290	0.290	-0.390	-0.209	-0.265	0.530	0.006	9.9
Science	11	554383	0408	All	8	OP	C	A	2	129722	0.583	0.087	0.184	0.583	0.145	0.001	0.337	-0.273	-0.246	0.337	-0.295	0.028	0.006	6.2
Science	11	554357	0409	All	10	OP	D	A	3	129722	0.571	0.113	0.177	0.138	0.571	0.001	0.404	-0.327	-0.303	-0.392	0.404	0.085	0.006	-9.9
Science	11	554358	0410	All	11	OP	C	B	2	129722	0.404	0.111	0.124	0.404	0.358	0.003	0.150	-0.302	-0.385	0.150	0.022	0.843	0.006	9.9
Science	11	554359	0411	All	12	OP	C	B	2	129722	0.637	0.088	0.112	0.637	0.161	0.002	0.302	-0.185	-0.255	0.302	-0.244	-0.236	0.006	9.9
Science	11	554361	0412	All	13	OP	A	A	2	129722	0.437	0.437	0.137	0.209	0.209	0.007	0.239	0.239	-0.261	-0.217	-0.194	0.695	0.006	9.9
Science	11	554414	0413	All	14	OP	B	A	2	129722	0.550	0.207	0.550	0.190	0.049	0.003	0.313	-0.226	0.313	-0.285	-0.283	0.177	0.006	9.9
Science	11	554386	0414	All	15	OP	C	A	2	129722	0.559	0.135	0.102	0.559	0.199	0.004	0.449	-0.385	-0.415	0.449	-0.327	0.136	0.006	-9.9
Science	11	554411	0415	All	16	OP	D	C	2	129722	0.423	0.250	0.126	0.197	0.423	0.005	0.403	-0.216	-0.523	-0.465	0.403	0.762	0.006	-9.9
Science	11	554385	0416	All	17	OP	B	A	2	129722	0.657	0.130	0.657	0.098	0.112	0.003	0.504	-0.407	0.504	-0.428	-0.336	-0.324	0.006	-9.9
Science	11	554412	0417	All	18	OP	A	A	2	129722	0.438	0.438	0.088	0.380	0.090	0.004	0.348	0.348	-0.395	-0.276	-0.350	0.688	0.006	9.9
Science	11	554481	0418	All	25	OP	D	A	2	129722	0.592	0.215	0.130	0.058	0.592	0.004	0.474	-0.352	-0.466	-0.372	0.474	-0.011	0.006	-9.9
Science	11	549918	0419	All	26	OP	D	C	2	129722	0.625	0.075	0.196	0.102	0.625	0.001	0.456	-0.350	-0.359	-0.393	0.456	-0.162	0.006	-9.9
Science	11	554392	0420	All	27	OP	D	C	2	129722	0.566	0.111	0.066	0.256	0.566	0.001	0.345	-0.250	-0.361	-0.279	0.345	0.106	0.006	8.8
Science	11	554293	0421	All	28	OP	B	D	2	129722	0.828	0.086	0.828	0.031	0.053	0.002	0.372	-0.274	0.372	-0.250	-0.233	-1.326	0.008	-9.9
Science	11	554428	0422	All	29	OP	B	D	2	129722	0.516	0.120	0.516	0.234	0.124	0.006	0.414	-0.281	0.414	-0.403	-0.344	0.333	0.006	-9.9
Science	11	554476	0423	All	30	OP	A	B	2	129722	0.800	0.800	0.046	0.062	0.080	0.011	0.447	0.447	-0.354	-0.361	-0.255	-1.132	0.007	-9.9
Science	11	554370	0424	All	31	OP	C	C	2	129722	0.516	0.254	0.191	0.516	0.037	0.002	0.320	-0.219	-0.328	0.320	-0.312	0.330	0.006	9.9
Science	11	554337	0425	All	32	OP	A	D	2	129722	0.604	0.604	0.146	0.064	0.184	0.002	0.418	0.418	-0.345	-0.426	-0.296	-0.069	0.006	-9.9
Science	11	554372	0426	All	33	OP	C	A	2	129722	0.622	0.198	0.124	0.622	0.054	0.002	0.356	-0.239	-0.321	0.356	-0.321	-0.158	0.006	3.6
Science	11	549936	0427	All	34	OP	C	B	2	129722	0.267	0.343	0.176	0.267	0.212	0.003	0.242	-0.154	-0.372	0.242	-0.310	1.545	0.007	9.9
Science	11	554393	0428	All	35	OP	D	A	2	129722	0.846	0.045	0.055	0.052	0.846	0.002	0.495	-0.291	-0.366	-0.348	0.495	-1.467	0.008	-9.9
Science	11	554396	0429	All	36	OP	C	A	2	129722	0.563	0.083	0.186	0.563	0.164	0.004	0.377	-0.381	-0.330	0.377	-0.257	0.126	0.006	-2.9
Science	11	549964	0430	All	37	OP	C	A	2	129722	0.412	0.103	0.213	0.412	0.268	0.004	0.286	-0.385	-0.416	0.286	-0.087	0.808	0.006	9.9

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	11	549974	0431	All	38	OP	C	B	2	129722	0.553	0.196	0.098	0.553	0.149	0.004	0.259	-0.264	-0.319	0.259	-0.076	0.164	0.006	9.9
Science	11	554467	0432	All	45	OP	B	D	2	129722	0.367	0.133	0.367	0.375	0.119	0.006	0.132	-0.214	0.132	-0.099	-0.083	1.016	0.006	9.9
Science	11	554465	0433	All	46	OP	C	A	2	129722	0.532	0.110	0.283	0.532	0.068	0.007	0.345	-0.416	-0.218	0.345	-0.311	0.262	0.006	9.9
Science	11	554468	0434	All	47	OP	A	D	2	129722	0.281	0.281	0.260	0.281	0.167	0.010	0.220	0.220	-0.321	-0.205	-0.157	1.465	0.007	9.9
Science	11	549919	0435	All	49	OP	C	A	2	129722	0.503	0.195	0.192	0.503	0.106	0.004	0.315	-0.266	-0.333	0.315	-0.155	0.390	0.006	9.9
Science	11	554301	0436	All	50	OP	D	A	2	129722	0.805	0.131	0.028	0.035	0.805	0.001	0.244	-0.097	-0.256	-0.272	0.244	-1.169	0.007	9.9
Science	11	549973	0437	All	51	OP	A	A	2	129722	0.881	0.881	0.062	0.029	0.027	0.001	0.336	0.336	-0.202	-0.233	-0.240	-1.781	0.009	-9.9
Science	11	554417	0438	All	52	OP	C	A	2	129722	0.498	0.085	0.164	0.498	0.250	0.003	0.227	-0.187	-0.259	0.227	-0.153	0.410	0.006	9.9
Science	11	554421	0439	All	53	OP	C	B	2	129722	0.500	0.213	0.141	0.500	0.143	0.003	0.295	-0.159	-0.381	0.295	-0.275	0.406	0.006	9.9
Science	11	549910	0440	All	54	OP	B	C	2	129722	0.839	0.056	0.839	0.051	0.053	0.001	0.371	-0.219	0.371	-0.271	-0.257	-1.414	0.008	-9.9
Science	11	549894	0441	All	55	OP	C	A	2	129722	0.796	0.064	0.070	0.796	0.069	0.001	0.391	-0.302	-0.289	0.391	-0.228	-1.104	0.007	-9.9
Science	11	554314	0442	All	56	OP	A	B	2	129722	0.434	0.434	0.174	0.244	0.145	0.002	0.456	0.456	-0.437	-0.416	-0.416	0.711	0.006	-9.9
Science	11	554374	0443	All	57	OP	D	D	2	129722	0.559	0.102	0.144	0.190	0.559	0.005	0.287	-0.356	-0.230	-0.187	0.287	0.138	0.006	9.9
Science	11	549990	0444	All	58	OP	C	D	2	129722	0.603	0.156	0.141	0.603	0.092	0.007	0.424	-0.262	-0.373	0.424	-0.398	-0.068	0.006	-9.9
Science	11	554409	0445	All	59	OP	A	C	2	129722	0.466	0.466	0.215	0.266	0.051	0.002	0.357	0.357	-0.386	-0.253	-0.377	0.568	0.006	7.5
Science	11	549949	0446	All	60	OP	A	A	2	129722	0.445	0.445	0.232	0.187	0.132	0.003	0.326	0.326	-0.327	-0.312	-0.230	0.656	0.006	9.9
Science	11	554302	0447	All	61	OP	C	B	2	129722	0.727	0.055	0.073	0.727	0.143	0.002	0.453	-0.383	-0.340	0.453	-0.311	-0.682	0.007	-9.9
Science	11	554298	0448	All	62	OP	A	A	2	129722	0.457	0.457	0.169	0.141	0.229	0.004	0.426	0.426	-0.438	-0.449	-0.297	0.605	0.006	-9.9
Science	11	549922	0449	1	21	MX	A	A	2	10898	0.396	0.396	0.461	0.101	0.032	0.009	-0.012	-0.012	0.096	-0.159	-0.187	0.868	0.021	9.9
Science	11	554379	0450	1	22	MX	C	D	2	10898	0.395	0.190	0.263	0.395	0.124	0.028	0.141	-0.095	-0.083	0.141	-0.209	0.875	0.021	9.9
Science	11	554319	0451	1	23	MX	D	C	2	10898	0.304	0.262	0.337	0.089	0.304	0.008	0.246	-0.184	-0.259	-0.434	0.246	1.332	0.022	5.6
Science	11	554313	0452	1	24	MX	C	A	2	10898	0.411	0.158	0.235	0.411	0.188	0.009	0.172	-0.226	-0.106	0.172	-0.159	0.801	0.021	9.9
Science	11	558061	0453	1	39	FT	A	C		10898	0.507	0.507	0.189	0.225	0.073	0.005	0.450	0.450	-0.398	-0.392	-0.389	0.353	0.021	-9.9
Science	11	558059	0454	1	40	FT	D	A		10898	0.541	0.123	0.247	0.082	0.541	0.006	0.424	-0.423	-0.322	-0.362	0.424	0.198	0.021	-9.9
Science	11	557965	0455	1	41	FT	C	B		10898	0.455	0.250	0.095	0.455	0.184	0.016	0.042	0.005	-0.305	0.042	0.095	0.595	0.021	9.9
Science	11	558053	0456	1	66	FT	B	A		10898	0.542	0.160	0.542	0.133	0.158	0.007	0.368	-0.362	0.368	-0.363	-0.177	0.195	0.021	-4.9
Science	11	558055	0457	1	67	FT	D	B		10898	0.113	0.167	0.158	0.554	0.113	0.008	0.044	-0.138	-0.245	0.016	0.044	2.682	0.031	9.9
Science	11	558054	0458	1	68	FT	B	B		10898	0.173	0.332	0.173	0.059	0.430	0.007	0.143	-0.105	0.143	-0.413	-0.170	2.149	0.026	9.9
Science	11	558058	0459	1	69	FT	D	A		10898	0.553	0.075	0.225	0.139	0.553	0.009	0.380	-0.397	-0.246	-0.363	0.380	0.144	0.021	-4.4
Science	11	558060	0460	1	71	FT	B	A		10898	0.373	0.205	0.373	0.142	0.263	0.016	0.149	-0.136	0.149	-0.354	0.013	0.979	0.021	9.9
Science	11	558067	0461	1	72	FT	C	C		10898	0.511	0.049	0.256	0.511	0.170	0.014	0.301	-0.166	-0.281	0.301	-0.246	0.337	0.021	3.1
Science	11	549998	0462	2	21	MX	D	A	2	10802	0.411	0.273	0.134	0.167	0.411	0.015	0.315	-0.180	-0.479	-0.323	0.315	0.826	0.021	1.5
Science	11	554449	0463	2	22	MX	B	A	1	10802	0.506	0.171	0.506	0.116	0.199	0.008	0.341	-0.270	0.341	-0.378	-0.256	0.383	0.021	1.0
Science	11	549908	0464	2	23	MX	C	B	1	10802	0.561	0.121	0.048	0.561	0.262	0.008	0.475	-0.404	-0.390	0.475	-0.400	0.124	0.021	-9.9
Science	11	554291	0465	2	24	MX	A	C	2	10802	0.574	0.574	0.140	0.115	0.163	0.008	0.491	0.491	-0.463	-0.386	-0.373	0.064	0.021	-9.9
Science	11	558064	0466	2	39	FT	C	A		10802	0.593	0.237	0.077	0.593	0.088	0.005	0.340	-0.290	-0.387	0.340	-0.138	-0.030	0.021	-0.4
Science	11	558065	0467	2	40	FT	A	A		10802	0.437	0.437	0.125	0.173	0.257	0.007	0.383	0.383	-0.445	-0.379	-0.283	0.705	0.021	-3.9
Science	11	558066	0468	2	41	FT	A	C		10802	0.469	0.469	0.172	0.170	0.185	0.004	0.421	0.421	-0.422	-0.390	-0.318	0.554	0.021	-7.6
Science	11	558068	0469	2	66	FT	B	D		10802	0.503	0.146	0.503	0.244	0.103	0.004	0.191	-0.140	0.191	-0.181	-0.146	0.397	0.021	9.9
Science	11	558071	0470	2	67	FT	A	D		10802	0.686	0.686	0.129	0.087	0.094	0.004	0.441	0.441	-0.333	-0.375	-0.282	-0.496	0.022	-9.9
Science	11	558073	0471	2	68	FT	A	A		10802	0.368	0.368	0.224	0.235	0.167	0.006	0.354	0.354	-0.376	-0.348	-0.297	1.035	0.021	1.1
Science	11	558077	0472	2	69	FT	A	A		10802	0.431	0.431	0.091	0.420	0.053	0.005	0.259	0.259	-0.367	-0.185	-0.271	0.731	0.021	9.9
Science	11	558074	0473	2	71	FT	B	B		10802	0.480	0.221	0.480	0.127	0.133	0.039	0.100	0.046	0.100	-0.256	-0.084	0.503	0.021	9.9

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	11	558076	0474	2	72	FT	D	D		10802	0.602	0.044	0.137	0.179	0.602	0.038	0.453	-0.378	-0.368	-0.362	0.453	-0.071	0.021	-9.9
Science	11	554403	0475	3	21	MX	A	A	2	10796	0.234	0.234	0.081	0.616	0.063	0.006	0.207	0.207	-0.293	-0.191	-0.298	1.777	0.024	9.9
Science	11	549961	0476	3	22	MX	B	B	2	10796	0.501	0.246	0.501	0.139	0.104	0.009	0.342	-0.208	0.342	-0.413	-0.295	0.409	0.021	0.5
Science	11	554366	0477	3	23	MX	A	A	2	10796	0.752	0.752	0.066	0.114	0.060	0.008	0.456	0.456	-0.324	-0.356	-0.293	-0.866	0.024	-9.9
Science	11	549965	0478	3	24	MX	B	C	2	10796	0.197	0.134	0.197	0.335	0.327	0.007	-0.001	-0.065	-0.001	0.028	0.013	2.019	0.025	9.9
Science	11	558078	0479	3	39	FT	B	A		10796	0.334	0.403	0.334	0.147	0.113	0.003	0.129	-0.038	0.129	-0.348	-0.122	1.210	0.022	9.9
Science	11	558075	0480	3	40	FT	D	D		10796	0.634	0.095	0.115	0.153	0.634	0.003	0.458	-0.370	-0.365	-0.355	0.458	-0.224	0.021	-9.9
Science	11	558079	0481	3	41	FT	A	B		10796	0.689	0.689	0.215	0.059	0.032	0.004	0.381	0.381	-0.298	-0.322	-0.247	-0.509	0.022	-4.7
Science	11	558082	0482	3	66	FT	B	A		10796	0.810	0.046	0.810	0.089	0.050	0.005	0.477	-0.301	0.477	-0.379	-0.286	-1.243	0.026	-9.9
Science	11	558085	0483	3	67	FT	B	B		10796	0.564	0.313	0.564	0.079	0.039	0.006	0.416	-0.339	0.416	-0.401	-0.306	0.116	0.021	-9.1
Science	11	558083	0484	3	68	FT	B	B		10796	0.437	0.143	0.437	0.284	0.126	0.010	0.373	-0.409	0.373	-0.303	-0.331	0.706	0.021	-1.9
Science	11	558081	0485	3	69	FT	C	A		10796	0.387	0.182	0.220	0.387	0.186	0.024	0.371	-0.296	-0.298	0.371	-0.445	0.945	0.021	-0.6
Science	11	558084	0486	3	71	FT	C	C		10796	0.491	0.120	0.262	0.491	0.105	0.023	0.145	-0.032	-0.193	0.145	-0.083	0.455	0.021	9.9
Science	11	558090	0487	3	72	FT	C	A		10796	0.505	0.208	0.219	0.505	0.047	0.020	0.353	-0.344	-0.260	0.353	-0.305	0.388	0.021	-1.2
Science	11	549989	0488	4	21	MX	B	A	2	10811	0.684	0.077	0.684	0.187	0.047	0.005	0.343	-0.330	0.343	-0.229	-0.244	-0.474	0.022	-2.0
Science	11	549984	0489	4	22	MX	C	A	2	10811	0.483	0.091	0.246	0.483	0.173	0.007	0.317	-0.265	-0.337	0.317	-0.199	0.488	0.021	2.8
Science	11	549932	0490	4	23	MX	C	C	2	10811	0.445	0.107	0.263	0.445	0.178	0.006	0.341	-0.381	-0.291	0.341	-0.261	0.665	0.021	1.3
Science	11	554478	0491	4	24	MX	A	D	2	10811	0.253	0.253	0.329	0.186	0.226	0.007	0.125	0.125	-0.044	-0.240	-0.187	1.642	0.023	9.9
Science	11	558092	0492	4	39	FT	B	D		10811	0.547	0.057	0.547	0.302	0.087	0.006	0.337	-0.257	0.337	-0.291	-0.298	0.195	0.021	0.0
Science	11	558089	0493	4	40	FT	D	C		10811	0.433	0.187	0.189	0.184	0.433	0.007	0.005	0.052	-0.091	0.035	0.005	0.722	0.021	9.9
Science	11	558088	0494	4	41	FT	C	A		10811	0.656	0.086	0.218	0.656	0.036	0.005	0.435	-0.348	-0.352	0.435	-0.303	-0.330	0.022	-9.9
Science	11	558093	0495	4	66	FT	B	B		10811	0.752	0.038	0.752	0.115	0.066	0.030	0.430	-0.277	0.430	-0.375	-0.224	-0.849	0.024	-9.9
Science	11	558091	0496	4	67	FT	D	B		10811	0.292	0.406	0.103	0.168	0.292	0.031	0.027	0.106	-0.280	-0.132	0.027	1.422	0.022	9.9
Science	11	558094	0497	4	68	FT	D	A		10811	0.335	0.148	0.355	0.132	0.335	0.031	0.072	-0.202	0.045	-0.165	0.072	1.192	0.022	9.9
Science	11	558097	0498	4	69	FT	B	A		10811	0.202	0.080	0.202	0.465	0.221	0.031	-0.025	-0.331	-0.025	0.124	0.003	1.959	0.025	9.9
Science	11	558096	0499	4	71	FT	A	C		10811	0.555	0.555	0.092	0.139	0.144	0.069	0.375	0.375	-0.305	-0.278	-0.310	0.156	0.021	-6.1
Science	11	558101	0500	4	72	FT	C	B		10811	0.606	0.128	0.164	0.606	0.032	0.069	0.397	-0.377	-0.251	0.397	-0.313	-0.082	0.021	-9.3
Science	11	549946	0501	5	21	MX	D	A	2	10812	0.246	0.158	0.286	0.299	0.246	0.010	0.248	-0.369	-0.198	-0.305	0.248	1.680	0.024	5.6
Science	11	549988	0502	5	22	MX	B	A	2	10812	0.380	0.156	0.380	0.246	0.209	0.009	0.054	-0.174	0.054	-0.065	0.070	0.970	0.021	9.9
Science	11	554446	0503	5	23	MX	C	D	2	10812	0.552	0.205	0.157	0.552	0.077	0.008	0.434	-0.372	-0.383	0.434	-0.322	0.165	0.021	-9.9
Science	11	554375	0504	5	24	MX	D	B	2	10812	0.441	0.067	0.090	0.393	0.441	0.009	0.198	-0.335	-0.367	-0.080	0.198	0.677	0.021	9.9
Science	11	558103	0505	5	39	FT	C	A		10812	0.770	0.107	0.070	0.770	0.050	0.003	0.431	-0.282	-0.352	0.431	-0.289	-0.972	0.024	-9.9
Science	11	558100	0506	5	40	FT	B	C		10812	0.101	0.138	0.101	0.462	0.295	0.004	0.005	-0.089	0.005	-0.031	0.074	2.840	0.033	9.9
Science	11	558102	0507	5	41	FT	A	B		10812	0.692	0.692	0.096	0.117	0.091	0.004	0.495	0.495	-0.380	-0.430	-0.294	-0.521	0.022	-9.9
Science	11	558105	0508	5	66	FT	A	A		10812	0.219	0.219	0.231	0.311	0.234	0.006	-0.016	-0.016	-0.204	-0.013	0.285	1.849	0.024	9.9
Science	11	558107	0509	5	67	FT	C	A		10812	0.545	0.107	0.206	0.545	0.135	0.006	0.351	-0.413	-0.187	0.351	-0.321	0.200	0.021	-3.0
Science	11	558106	0510	5	68	FT	A	C		10812	0.160	0.160	0.267	0.263	0.302	0.007	-0.097	-0.097	0.096	0.070	0.202	2.270	0.027	9.9
Science	11	558108	0511	5	69	FT	B	C		10812	0.359	0.109	0.359	0.319	0.206	0.007	0.156	-0.301	0.156	-0.147	-0.071	1.073	0.021	9.9
Science	11	558111	0512	5	71	FT	C	A		10812	0.317	0.144	0.430	0.317	0.088	0.021	0.106	-0.283	-0.002	0.106	-0.236	1.284	0.022	9.9
Science	11	558114	0513	5	72	FT	D	D		10812	0.333	0.111	0.239	0.295	0.333	0.021	0.168	-0.334	-0.113	-0.139	0.168	1.200	0.022	9.9
Science	11	554348	0514	6	21	MX	B	A	1	10753	0.241	0.393	0.241	0.114	0.245	0.006	0.204	-0.200	0.204	-0.379	-0.168	1.726	0.024	9.9
Science	11	554368	0515	6	22	MX	D	A	2	10753	0.476	0.129	0.177	0.211	0.476	0.007	0.354	-0.321	-0.278	-0.348	0.354	0.520	0.021	0.0
Science	11	549962	0516	6	23	MX	B	D	2	10753	0.209	0.323	0.209	0.178	0.281	0.009	0.088	-0.099	0.088	-0.207	-0.032	1.929	0.025	9.9

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	11	554482	0517	6	24	MX	C	B	2	10753	0.333	0.140	0.119	0.333	0.400	0.007	0.236	-0.353	-0.312	0.236	-0.154	1.210	0.022	9.9
Science	11	558117	0518	6	39	FT	A	D		10753	0.294	0.294	0.329	0.186	0.186	0.004	0.110	0.110	-0.124	-0.101	-0.128	1.417	0.023	9.9
Science	11	558112	0519	6	40	FT	C	C		10753	0.480	0.099	0.349	0.480	0.068	0.004	0.430	-0.405	-0.377	0.430	-0.389	0.501	0.021	-8.7
Science	11	558113	0520	6	41	FT	B	A		10753	0.332	0.034	0.332	0.046	0.585	0.002	0.448	-0.462	0.448	-0.480	-0.427	1.218	0.022	-7.2
Science	11	558115	0521	6	66	FT	D	A		10753	0.457	0.094	0.245	0.175	0.457	0.028	0.464	-0.463	-0.402	-0.434	0.464	0.610	0.021	-9.9
Science	11	558119	0522	6	67	FT	B	C		10753	0.239	0.538	0.239	0.139	0.054	0.030	0.043	0.041	0.043	-0.168	-0.370	1.736	0.024	9.9
Science	11	558118	0523	6	68	FT	D	C		10753	0.525	0.158	0.154	0.134	0.525	0.029	0.363	-0.296	-0.302	-0.322	0.363	0.294	0.021	-2.8
Science	11	558120	0524	6	69	FT	A	A		10753	0.565	0.565	0.150	0.191	0.064	0.030	0.421	0.421	-0.371	-0.313	-0.368	0.108	0.021	-9.9
Science	11	558123	0525	6	71	FT	C	B		10753	0.708	0.077	0.099	0.708	0.064	0.052	0.376	-0.244	-0.341	0.376	-0.217	-0.611	0.023	-5.3
Science	11	558126	0526	6	72	FT	D	A		10753	0.250	0.349	0.102	0.248	0.250	0.052	0.178	-0.150	-0.448	-0.127	0.178	1.674	0.024	9.9
Science	11	549923	0527	7	21	MX	C	A	2	10754	0.694	0.055	0.087	0.694	0.159	0.005	0.394	-0.290	-0.370	0.394	-0.239	-0.528	0.022	-8.1
Science	11	549958	0528	7	22	MX	B	C	2	10754	0.753	0.079	0.753	0.086	0.076	0.005	0.443	-0.334	0.443	-0.310	-0.306	-0.868	0.024	-9.9
Science	11	554448	0529	7	23	MX	B	A	2	10754	0.545	0.168	0.545	0.203	0.076	0.007	0.263	-0.164	0.263	-0.198	-0.351	0.204	0.021	8.1
Science	11	554391	0530	7	24	MX	A	D	2	10754	0.392	0.392	0.297	0.133	0.171	0.007	0.271	0.271	-0.243	-0.322	-0.229	0.919	0.021	7.7
Science	11	558128	0531	7	39	FT	C	B		10754	0.732	0.062	0.093	0.732	0.112	0.002	0.449	-0.355	-0.356	0.449	-0.281	-0.740	0.023	-9.9
Science	11	558124	0532	7	40	FT	B	D		10754	0.360	0.110	0.360	0.144	0.382	0.004	0.114	-0.308	0.114	-0.280	0.033	1.076	0.021	9.9
Science	11	558125	0533	7	41	FT	C	A		10754	0.307	0.100	0.502	0.307	0.087	0.004	0.070	-0.278	0.007	0.070	-0.185	1.345	0.022	9.9
Science	11	558127	0534	7	66	FT	B	A		10754	0.651	0.048	0.651	0.207	0.063	0.030	0.252	-0.288	0.252	-0.165	-0.153	-0.305	0.022	7.8
Science	11	558131	0535	7	67	FT	B	C		10754	0.671	0.131	0.671	0.063	0.103	0.031	0.435	-0.375	0.435	-0.351	-0.260	-0.410	0.022	-9.9
Science	11	558135	0536	7	68	FT	A	A		10754	0.246	0.246	0.412	0.237	0.070	0.034	0.098	0.098	0.014	-0.253	-0.231	1.688	0.024	9.9
Science	11	558130	0537	7	69	FT	D	D		10754	0.293	0.146	0.368	0.163	0.293	0.031	0.186	-0.272	-0.092	-0.335	0.186	1.424	0.023	9.9
Science	11	558134	0538	7	71	FT	D	A		10754	0.733	0.062	0.073	0.087	0.733	0.044	0.528	-0.332	-0.413	-0.422	0.528	-0.745	0.023	-9.9
Science	11	558142	0539	7	72	FT	A	C		10754	0.289	0.289	0.240	0.302	0.123	0.046	0.129	0.129	-0.076	-0.155	-0.140	1.443	0.023	9.9
Science	11	554328	0540	8	21	MX	A	C	2	10846	0.590	0.590	0.122	0.246	0.033	0.009	0.315	0.315	-0.288	-0.255	-0.163	-0.003	0.021	0.7
Science	11	549939	0541	8	22	MX	D	B	2	10846	0.727	0.040	0.164	0.062	0.727	0.007	0.499	-0.315	-0.398	-0.379	0.499	-0.703	0.023	-9.9
Science	11	554338	0542	8	23	MX	D	A	2	10846	0.402	0.131	0.308	0.150	0.402	0.008	0.214	-0.259	-0.201	-0.151	0.214	0.869	0.021	9.9
Science	11	554340	0543	8	24	MX	D	A	2	10846	0.538	0.220	0.166	0.068	0.538	0.007	0.305	-0.225	-0.205	-0.409	0.305	0.238	0.021	3.0
Science	11	558136	0544	8	39	FT	D	B		10846	0.269	0.232	0.195	0.297	0.269	0.006	0.281	-0.242	-0.376	-0.306	0.281	1.551	0.023	4.5
Science	11	558137	0545	8	40	FT	B	A		10846	0.266	0.199	0.266	0.226	0.303	0.006	0.093	-0.188	0.093	-0.206	0.033	1.573	0.023	9.9
Science	11	558138	0546	8	41	FT	C	A		10846	0.512	0.125	0.314	0.512	0.047	0.002	0.241	-0.276	-0.159	0.241	-0.263	0.360	0.021	9.9
Science	11	558143	0547	8	66	FT	C	D		10846	0.336	0.209	0.304	0.336	0.147	0.004	-0.091	0.020	0.099	-0.091	0.208	1.192	0.022	9.9
Science	11	558139	0548	8	67	FT	B	D		10846	0.510	0.240	0.510	0.183	0.062	0.005	0.301	-0.199	0.301	-0.294	-0.347	0.369	0.021	4.0
Science	11	558141	0549	8	68	FT	C	A		10846	0.327	0.227	0.085	0.327	0.356	0.004	0.211	-0.186	-0.382	0.211	-0.172	1.239	0.022	9.9
Science	11	558148	0550	8	69	FT	B	A		10846	0.212	0.209	0.212	0.267	0.306	0.006	0.093	0.003	0.093	-0.231	-0.078	1.904	0.025	9.9
Science	11	558449	0551	8	71	FT	C	C		10846	0.523	0.135	0.115	0.523	0.171	0.056	0.301	-0.101	-0.422	0.301	-0.254	0.307	0.021	4.5
Science	11	558150	0552	8	72	FT	A	D		10846	0.707	0.707	0.060	0.065	0.111	0.056	0.383	0.383	-0.313	-0.335	-0.233	-0.596	0.022	-7.4
Science	11	554406	0553	9	21	MX	C	C	2	10829	0.452	0.183	0.263	0.452	0.089	0.013	0.245	-0.174	-0.286	0.245	-0.113	0.624	0.021	9.9
Science	11	549892	0554	9	22	MX	A	A	2	10829	0.709	0.709	0.196	0.052	0.033	0.009	0.342	0.342	-0.251	-0.292	-0.211	-0.615	0.023	-3.6
Science	11	554351	0555	9	23	MX	D	D	2	10829	0.666	0.087	0.101	0.135	0.666	0.010	0.511	-0.365	-0.357	-0.437	0.511	-0.385	0.022	-9.9
Science	11	554341	0556	9	24	MX	A	A	2	10829	0.728	0.728	0.134	0.061	0.066	0.011	0.333	0.333	-0.173	-0.284	-0.291	-0.719	0.023	-2.7
Science	11	558153	0557	9	39	FT	D	D		10829	0.515	0.104	0.211	0.166	0.515	0.004	0.336	-0.314	-0.231	-0.353	0.336	0.334	0.021	-0.3
Science	11	558149	0558	9	40	FT	B	C		10829	0.265	0.383	0.265	0.171	0.175	0.005	0.004	0.137	0.004	-0.182	-0.146	1.564	0.023	9.9
Science	11	558147	0559	9	41	FT	C	B		10829	0.278	0.161	0.172	0.278	0.383	0.006	-0.041	0.002	-0.103	-0.041	0.138	1.491	0.023	9.9

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	11	558152	0560	9	66	FT	C	A		10829	0.427	0.104	0.382	0.427	0.083	0.004	0.219	-0.220	-0.181	0.219	-0.223	0.740	0.021	9.9
Science	11	558156	0561	9	67	FT	B	D		10829	0.327	0.303	0.327	0.212	0.152	0.005	0.134	-0.132	0.134	-0.216	-0.038	1.230	0.022	9.9
Science	11	558155	0562	9	68	FT	A	C		10829	0.332	0.332	0.222	0.275	0.165	0.007	0.255	0.255	-0.291	-0.269	-0.180	1.203	0.022	9.6
Science	11	558154	0563	9	69	FT	D	A		10829	0.577	0.098	0.211	0.108	0.577	0.006	0.424	-0.401	-0.275	-0.422	0.424	0.047	0.021	-9.3
Science	11	558158	0564	9	71	FT	B	A		10829	0.443	0.131	0.443	0.333	0.083	0.009	0.149	-0.229	0.149	-0.081	-0.137	0.665	0.021	9.9
Science	11	558159	0565	9	72	FT	A	A		10829	0.465	0.465	0.155	0.277	0.094	0.009	0.243	0.243	-0.223	-0.163	-0.322	0.564	0.021	9.9
Science	11	549977	0566	10	21	MX	B	A	2	10837	0.444	0.104	0.444	0.332	0.110	0.009	0.307	-0.319	0.307	-0.236	-0.307	0.666	0.021	5.4
Science	11	554371	0567	10	22	MX	C	A	2	10837	0.444	0.203	0.265	0.444	0.081	0.007	0.269	-0.290	-0.186	0.269	-0.288	0.666	0.021	8.8
Science	11	549968	0568	10	23	MX	A	B	2	10837	0.362	0.362	0.160	0.120	0.350	0.008	0.330	0.330	-0.294	-0.488	-0.259	1.062	0.021	3.4
Science	11	549985	0569	10	24	MX	A	C	1	10837	0.283	0.283	0.160	0.245	0.300	0.012	0.041	0.041	-0.234	-0.080	0.104	1.478	0.023	9.9
Science	11	558163	0570	10	39	FT	D	A		10837	0.701	0.057	0.097	0.139	0.701	0.006	0.479	-0.319	-0.352	-0.398	0.479	-0.577	0.022	-9.9
Science	11	558162	0571	10	40	FT	C	D		10837	0.615	0.125	0.206	0.615	0.048	0.006	0.346	-0.361	-0.190	0.346	-0.358	-0.138	0.021	-0.1
Science	11	558161	0572	10	41	FT	B	A		10837	0.274	0.162	0.274	0.359	0.200	0.004	0.031	-0.043	0.031	-0.015	-0.052	1.527	0.023	9.9
Science	11	558165	0573	10	66	FT	A	B		10837	0.539	0.539	0.062	0.251	0.134	0.014	0.503	0.503	-0.393	-0.443	-0.434	0.224	0.021	-9.9
Science	11	558166	0574	10	67	FT	B	A		10837	0.503	0.120	0.503	0.273	0.098	0.006	0.327	-0.164	0.327	-0.323	-0.301	0.395	0.021	2.0
Science	11	558167	0575	10	68	FT	D	B		10837	0.405	0.173	0.253	0.162	0.405	0.007	0.537	-0.498	-0.490	-0.568	0.537	0.854	0.021	-9.9
Science	11	558168	0576	10	69	FT	B	A		10837	0.513	0.097	0.513	0.244	0.140	0.007	0.294	-0.301	0.294	-0.216	-0.263	0.347	0.021	5.9
Science	11	558171	0577	10	71	FT	D	B		10837	0.247	0.180	0.191	0.349	0.247	0.033	0.273	-0.368	-0.380	-0.217	0.273	1.686	0.024	4.5
Science	11	558173	0578	10	72	FT	B	C		10837	0.555	0.219	0.555	0.120	0.072	0.034	0.383	-0.236	0.383	-0.405	-0.343	0.148	0.021	-5.1
Science	11	549912	0579	11	21	MX	D	A	1	10836	0.493	0.138	0.127	0.235	0.493	0.008	0.449	-0.322	-0.509	-0.374	0.449	0.447	0.021	-9.9
Science	11	554401	0580	11	22	MX	B	C	2	10836	0.476	0.109	0.476	0.284	0.120	0.012	0.188	0.005	0.188	-0.241	-0.172	0.526	0.021	9.9
Science	11	549970	0581	11	23	MX	C	B	2	10836	0.431	0.127	0.351	0.431	0.083	0.009	0.165	-0.238	-0.087	0.165	-0.214	0.736	0.021	9.9
Science	11	549914	0582	11	24	MX	C	A	2	10836	0.757	0.085	0.060	0.757	0.090	0.008	0.466	-0.324	-0.364	0.466	-0.318	-0.890	0.024	-9.9
Science	11	558177	0583	11	39	FT	A	A		10836	0.636	0.636	0.155	0.169	0.036	0.003	0.421	0.421	-0.311	-0.360	-0.321	-0.234	0.021	-9.9
Science	11	558172	0584	11	40	FT	B	D		10836	0.490	0.123	0.490	0.280	0.102	0.005	0.327	-0.256	0.327	-0.285	-0.321	0.458	0.021	2.6
Science	11	558174	0585	11	41	FT	C	B		10836	0.804	0.105	0.053	0.804	0.035	0.003	0.459	-0.316	-0.359	0.459	-0.292	-1.201	0.026	-9.9
Science	11	558176	0586	11	66	FT	A	A		10836	0.640	0.640	0.095	0.172	0.088	0.005	0.360	0.360	-0.288	-0.281	-0.268	-0.250	0.021	-1.7
Science	11	558178	0587	11	67	FT	C	A		10836	0.529	0.128	0.230	0.529	0.107	0.006	0.353	-0.342	-0.258	0.353	-0.312	0.280	0.021	-1.9
Science	11	558180	0588	11	68	FT	D	B		10836	0.340	0.221	0.247	0.186	0.340	0.005	0.340	-0.306	-0.327	-0.391	0.340	1.180	0.022	3.0
Science	11	558185	0589	11	69	FT	B	B		10836	0.368	0.334	0.368	0.119	0.172	0.006	-0.081	0.288	-0.081	-0.235	-0.069	1.038	0.021	9.9
Science	11	558182	0590	11	71	FT	D	A		10836	0.466	0.148	0.202	0.152	0.466	0.032	0.444	-0.365	-0.402	-0.428	0.444	0.572	0.021	-9.9
Science	11	558184	0591	11	72	FT	A	C		10836	0.256	0.256	0.203	0.242	0.266	0.033	0.178	0.178	-0.207	-0.199	-0.166	1.634	0.023	9.9
Science	11	550000	0592	12	21	MX	B	C	2	10748	0.384	0.242	0.384	0.122	0.144	0.107	0.299	-0.216	0.299	-0.365	-0.154	0.966	0.021	5.7
Science	11	554369	0593	12	22	MX	C	A	2	10748	0.548	0.082	0.291	0.548	0.071	0.009	0.344	-0.344	-0.243	0.344	-0.333	0.193	0.021	-0.5
Science	11	549924	0594	12	23	MX	B	A	2	10748	0.533	0.228	0.533	0.111	0.117	0.010	0.434	-0.350	0.434	-0.421	-0.328	0.264	0.021	-9.9
Science	11	554397	0595	12	24	MX	A	D	2	10748	0.295	0.295	0.256	0.173	0.265	0.010	0.314	0.314	-0.294	-0.370	-0.324	1.421	0.023	5.0
Science	11	558186	0596	12	39	FT	B	B		10748	0.609	0.058	0.609	0.120	0.210	0.002	0.471	-0.346	0.471	-0.429	-0.365	-0.102	0.021	-9.9
Science	11	558183	0597	12	40	FT	D	D		10748	0.519	0.118	0.173	0.187	0.519	0.003	0.342	-0.359	-0.306	-0.232	0.342	0.327	0.021	1.7
Science	11	558187	0598	12	41	FT	A	B		10748	0.421	0.421	0.182	0.255	0.138	0.003	0.301	0.301	-0.402	-0.187	-0.283	0.788	0.021	6.1
Science	11	558189	0599	12	66	FT	D	A		10748	0.311	0.114	0.216	0.352	0.311	0.006	0.379	-0.501	-0.414	-0.334	0.379	1.335	0.022	-2.2
Science	11	558192	0600	12	67	FT	B	A		10748	0.386	0.243	0.386	0.158	0.206	0.007	0.322	-0.263	0.322	-0.311	-0.348	0.954	0.021	4.2
Science	11	558190	0601	12	68	FT	D	D		10748	0.487	0.095	0.116	0.295	0.487	0.006	0.345	-0.384	-0.361	-0.238	0.345	0.474	0.021	1.0
Science	11	558191	0602	12	69	FT	D	D		10748	0.554	0.071	0.074	0.294	0.554	0.007	0.408	-0.410	-0.363	-0.306	0.408	0.164	0.021	-7.9

Information											Proportions						Correlations					Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Key	R.C.	D.O.K.	N	P-Val.	A	B	C	D	Omit	Pt. Bis.	A	B	C	D	Logit	Logit SE	Fit
Science	11	558195	0603	12	71	FT	C	A		10748	0.164	0.233	0.377	0.164	0.200	0.027	-0.099	0.101	0.168	-0.099	0.105	2.272	0.027	9.9
Science	11	558196	0604	12	72	FT	C	C		10748	0.722	0.066	0.096	0.722	0.089	0.026	0.461	-0.347	-0.338	0.461	-0.322	-0.686	0.023	-9.9

Appendix G:

2008 Grades 4, 8, and 11 Open-Ended Statistics

Table Legend

Column	Description
Cont.	Tested Content
Grade	Tested Grade
ID	Item identification number
Pub. ID	Public item identification number
Form	Form on which the item appeared 'All' = common form
Seq.	Sequence number of the item
Status	Status of the item at the time of administration 'OP' = common, 'MX' = matrix 'FT' = field test
Pts.	Maximum points possible
R.C.	Reporting category
D.O.K.	Depth of knowledge
N	Number of students
Mean	Mean score
Proportion [0-4, Omit]	Proportion achieving score point
Pt. Bis.	Item total correlation
Correlation [0-4]	Point biserial of the score point
Logit	Rasch logit difficulty
Logit SE	Rasch logit difficulty standard error
Fit	Rasch outfit statistic

Information												Proportions						Correlations						Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Pts.	R.C.	D.O.K.	N	Mean	0	1	2	3	4	Omit	Pt. Bis.	0	1	2	3	4	Logit	Logit SE	Fit
Science	4	554181	0605	All	32	OP	2	A	3	126117	1.365	0.130	0.375	0.495			0.004	0.349	-0.253	-0.153	0.319			0.214	0.005	9.9
Science	4	553977	0606	All	33	OP	2	C	2	126117	1.353	0.200	0.246	0.553			0.007	0.317	-0.283	-0.057	0.278			0.392	0.004	9.9
Science	4	554006	0607	All	34	OP	2	A	3	126117	1.138	0.212	0.437	0.350			0.008	0.440	-0.375	-0.036	0.358			0.792	0.005	9.9
Science	4	554021	0608	All	67	OP	2	B	3	126117	1.641	0.084	0.190	0.726			0.005	0.503	-0.414	-0.223	0.454			-0.395	0.005	-9.9
Science	4	553993	0609	All	68	OP	2	D	2	126117	0.948	0.241	0.570	0.189			0.008	0.358	-0.340	0.115	0.226			1.291	0.005	9.9
Science	4	549578	0610	1	35	MX	2	C	2	7141	0.976	0.285	0.453	0.261			0.013	0.289	-0.248	0.022	0.230			1.149	0.019	9.9
Science	4	558199	0611	1	69	FT	2	B		3356	0.456	0.582	0.380	0.038			0.011	0.371	-0.352	0.280	0.197			2.869	0.033	-2.0
Science	4	554247	0612	2	35	MX	2	B	3	7006	0.628	0.472	0.427	0.100			0.013	0.460	-0.446	0.287	0.268			2.183	0.021	-3.3
Science	4	558208	0613	2	69	FT	2	A		3324	1.172	0.121	0.587	0.292			0.010	0.389	-0.341	-0.038	0.285			0.620	0.032	2.2
Science	4	554166	0614	3	35	MX	2	A	3	6978	0.450	0.620	0.311	0.069			0.022	0.291	-0.278	0.192	0.181			2.602	0.022	8.7
Science	4	558210	0615	3	69	FT	2	C		3310	0.731	0.371	0.528	0.102			0.014	0.418	-0.409	0.262	0.221			2.017	0.031	-0.4
Science	4	554260	0616	4	35	MX	2	A	3	7022	0.426	0.625	0.324	0.051			0.018	0.244	-0.232	0.173	0.142			2.740	0.023	9.9
Science	4	558216	0617	4	69	FT	2	B		3349	1.003	0.220	0.556	0.224			0.012	0.371	-0.325	0.046	0.269			1.169	0.029	2.8
Science	4	554261	0618	5	35	MX	2	C	2	7005	0.809	0.379	0.434	0.188			0.016	0.386	-0.324	0.068	0.317			1.603	0.019	3.3
Science	4	558221	0619	5	69	FT	2	C		3331	0.592	0.529	0.350	0.121			0.019	0.393	-0.361	0.184	0.285			2.158	0.028	-1.3
Science	4	549577	0620	6	35	MX	2	B	2	7021	0.558	0.491	0.459	0.049			0.016	0.450	-0.448	0.370	0.185			2.608	0.023	-4.4
Science	4	558226	0621	6	69	FT	2	A		3335	0.988	0.213	0.585	0.201			0.019	0.488	-0.455	0.120	0.318			1.225	0.031	-3.1
Science	4	554093	0622	7	35	MX	2	B	3	6991	0.468	0.620	0.293	0.088			0.015	0.277	-0.273	0.185	0.171			2.453	0.021	9.9
Science	4	558231	0623	7	69	FT	2	D		3331	0.569	0.573	0.286	0.141			0.019	0.362	-0.360	0.205	0.244			2.171	0.027	3.4
Science	4	554231	0624	8	35	MX	2	C	3	7003	0.190	0.825	0.161	0.014			0.016	0.196	-0.194	0.175	0.081			3.770	0.030	6.8
Science	4	558236	0625	8	69	FT	2	A		3320	1.146	0.275	0.304	0.421			0.014	0.549	-0.503	-0.004	0.459			0.903	0.025	-3.2
Science	4	554080	0626	9	35	MX	2	C	2	7006	0.311	0.722	0.244	0.034			0.020	0.371	-0.369	0.312	0.175			3.188	0.025	-3.1
Science	4	558243	0627	9	69	FT	2	A		3332	1.189	0.098	0.615	0.287			0.009	0.452	-0.351	-0.119	0.359			0.512	0.033	-1.9
Science	4	554035	0628	10	35	MX	2	A	3	7011	0.940	0.250	0.560	0.190			0.013	0.456	-0.424	0.132	0.300			1.347	0.021	-1.3
Science	4	558245	0629	10	69	FT	2	C		3341	1.243	0.176	0.404	0.420			0.011	0.219	-0.201	-0.014	0.170			0.626	0.028	9.9
Science	4	554217	0630	11	35	MX	2	A	2	7023	0.981	0.295	0.430	0.276			0.012	0.548	-0.514	0.111	0.401			1.201	0.019	-7.5
Science	4	558249	0631	11	69	FT	2	A		3327	0.999	0.290	0.421	0.289			0.013	0.464	-0.378	-0.020	0.401			1.216	0.027	0.8
Science	4	554150	0632	12	35	MX	2	A	2	7000	1.457	0.198	0.147	0.655			0.016	0.458	-0.428	-0.076	0.415			0.248	0.018	6.2
Science	4	558255	0633	12	69	FT	2	A		3318	0.921	0.318	0.442	0.239			0.014	0.444	-0.409	0.104	0.326			1.383	0.027	0.4
Science	4	554123	0634	13	35	MX	2	A	3	6969	1.001	0.279	0.441	0.280			0.015	0.492	-0.448	0.068	0.372			1.155	0.019	-2.4
Science	4	558261	0635	13	69	FT	2	B		3311	1.070	0.152	0.625	0.222			0.010	0.315	-0.320	0.079	0.184			0.970	0.032	5.3
Science	4	554138	0636	14	35	MX	2	B	3	6985	1.368	0.219	0.194	0.587			0.013	0.442	-0.361	-0.161	0.432			0.422	0.018	7.6
Science	4	558266	0637	14	69	FT	2	A		3300	1.308	0.105	0.482	0.413			0.007	0.413	-0.377	-0.075	0.310			0.324	0.030	1.4
Science	4	554136	0638	15	35	MX	2	A	3	7034	1.043	0.309	0.338	0.352			0.011	0.485	-0.455	0.057	0.384			1.065	0.018	1.5
Science	4	558270	0639	15	69	FT	2	C		3333	0.669	0.480	0.371	0.149			0.014	0.342	-0.334	0.181	0.223			2.007	0.028	7.5
Science	4	554050	0640	16	35	MX	2	B	3	6964	1.138	0.216	0.430	0.354			0.013	0.497	-0.428	-0.034	0.404			0.805	0.019	-1.3
Science	4	558276	0641	16	69	FT	2	B		3299	0.700	0.439	0.423	0.139			0.015	0.377	-0.376	0.222	0.222			2.000	0.029	4.2
Science	4	554064	0642	17	35	MX	2	A	3	6967	1.019	0.259	0.464	0.277			0.012	0.409	-0.352	0.018	0.324			1.063	0.019	5.7
Science	4	558281	0643	17	69	FT	2	B		3276	0.893	0.373	0.361	0.266			0.012	0.462	-0.446	0.136	0.340			1.403	0.026	1.5
Science	4	554165	0644	18	35	MX	2	B	3	6991	0.797	0.393	0.418	0.189			0.014	0.461	-0.430	0.164	0.329			1.650	0.019	-0.2
Science	4	558285	0645	18	69	FT	2	B		3323	1.206	0.209	0.376	0.415			0.010	0.411	-0.386	0.002	0.317			0.721	0.027	6.1
Science	8	549849	0646	All	34	OP	2	D	3	136648	1.322	0.064	0.551	0.386			0.006	0.345	-0.263	-0.149	0.284			-0.628	0.005	9.9
Science	8	549591	0647	All	35	OP	2	A	3	136648	0.851	0.287	0.575	0.138			0.053	0.207	-0.099	-0.083	0.250			0.910	0.005	9.9
Science	8	549873	0648	All	36	OP	2	C	2	136648	0.812	0.401	0.386	0.213			0.022	0.438	-0.409	0.139	0.325			0.890	0.004	7.4
Science	8	549662	0649	All	71	OP	2	A	3	136648	1.342	0.154	0.350	0.496			0.012	0.375	-0.330	-0.074	0.309			-0.328	0.004	9.9
Science	8	549606	0650	All	72	OP	2	B	3	136648	0.979	0.306	0.410	0.284			0.020	0.530	-0.487	0.084	0.406			0.514	0.004	-9.9

Information												Proportions						Correlations						Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Pts.	R.C.	D.O.K.	N	Mean	0	1	2	3	4	Omit	Pt. Bis.	0	1	2	3	4	Logit	Logit SE	Fit
Science	8	549874	0651	1	37	MX	2	B	3	7664	1.122	0.257	0.364	0.379			0.032	0.544	-0.491	0.001	0.441			0.174	0.017	-7.1
Science	8	558291	0652	1	73	FT	2	B		3665	1.279	0.218	0.285	0.497			0.016	0.506	-0.383	-0.194	0.491			-0.068	0.024	-1.5
Science	8	549864	0653	2	37	MX	2	C	3	7582	0.598	0.490	0.421	0.088			0.029	0.510	-0.492	0.330	0.293			1.560	0.020	-9.1
Science	8	558303	0654	2	73	FT	2	B		3656	1.277	0.217	0.289	0.494			0.020	0.567	-0.480	-0.122	0.507			-0.071	0.025	-5.4
Science	8	549605	0655	3	37	MX	2	A	3	7544	1.096	0.228	0.448	0.324			0.036	0.447	-0.412	0.032	0.335			0.230	0.018	1.2
Science	8	558310	0656	3	73	FT	2	B		3647	0.758	0.447	0.348	0.205			0.026	0.542	-0.496	0.158	0.425			1.053	0.025	-5.9
Science	8	549704	0657	4	37	MX	2	B	3	7605	0.634	0.511	0.345	0.145			0.057	0.472	-0.442	0.211	0.342			1.318	0.018	-4.6
Science	8	558320	0658	4	73	FT	2	B		3683	0.977	0.197	0.629	0.174			0.024	0.490	-0.413	0.062	0.354			0.611	0.030	-4.8
Science	8	549634	0659	5	37	MX	2	A	3	7581	0.294	0.725	0.255	0.019			0.031	0.339	-0.333	0.296	0.142			2.749	0.025	-1.3
Science	8	558328	0660	5	73	FT	2	A		3674	0.700	0.441	0.418	0.141			0.079	0.472	-0.413	0.161	0.362			1.258	0.027	-3.0
Science	8	549863	0661	6	37	MX	2	B	3	7561	1.420	0.103	0.373	0.523			0.028	0.434	-0.366	-0.141	0.360			-0.617	0.019	0.3
Science	8	558336	0662	6	73	FT	2	A		3644	1.129	0.255	0.361	0.384			0.024	0.544	-0.483	-0.016	0.449			0.227	0.025	-4.7
Science	8	549840	0663	7	37	MX	2	D	3	7579	0.673	0.493	0.341	0.166			0.037	0.552	-0.518	0.227	0.406			1.183	0.018	-9.9
Science	8	558346	0664	7	73	FT	2	A		3651	0.520	0.561	0.357	0.081			0.030	0.492	-0.467	0.309	0.306			1.759	0.029	-6.6
Science	8	549816	0665	8	37	MX	2	B	3	7575	1.238	0.168	0.427	0.406			0.029	0.517	-0.464	-0.050	0.403			-0.117	0.018	-6.5
Science	8	558356	0666	8	73	FT	2	A		3669	1.053	0.191	0.565	0.244			0.013	0.483	-0.448	0.070	0.329			0.366	0.028	-4.0
Science	8	549793	0667	9	37	MX	2	D	3	7605	1.007	0.265	0.463	0.272			0.035	0.429	-0.392	0.064	0.317			0.456	0.018	4.2
Science	8	558366	0668	9	73	FT	2	A		3673	1.189	0.154	0.502	0.344			0.026	0.454	-0.338	-0.128	0.392			-0.002	0.028	-0.7
Science	8	549648	0669	10	37	MX	2	D	2	7583	0.618	0.533	0.316	0.151			0.041	0.493	-0.477	0.245	0.346			1.323	0.018	-5.1
Science	8	558374	0670	10	73	FT	2	A		3668	1.212	0.204	0.380	0.416			0.022	0.442	-0.383	-0.055	0.367			0.040	0.025	1.8
Science	8	549633	0671	11	37	MX	2	C	3	7597	0.976	0.315	0.394	0.291			0.036	0.425	-0.418	0.117	0.301			0.517	0.017	4.1
Science	8	558382	0672	11	73	FT	2	A		3659	0.591	0.557	0.294	0.149			0.086	0.379	-0.354	0.162	0.287			1.418	0.026	3.1
Science	8	549592	0673	12	37	MX	2	D	3	7607	1.069	0.281	0.368	0.350			0.046	0.461	-0.405	-0.001	0.383			0.313	0.017	1.3
Science	8	558391	0674	12	73	FT	2	A		3668	0.494	0.604	0.297	0.098			0.055	0.378	-0.344	0.184	0.283			1.743	0.028	0.3
Science	8	549839	0675	13	37	MX	2	B	3	7567	1.132	0.226	0.416	0.358			0.038	0.528	-0.516	0.070	0.378			0.167	0.018	-6.7
Science	8	558402	0676	13	73	FT	2	A		3658	0.987	0.118	0.777	0.105			0.020	0.483	-0.423	0.108	0.299			0.583	0.037	-5.5
Science	8	549817	0677	14	37	MX	2	A	3	7596	0.603	0.546	0.304	0.150			0.024	0.509	-0.480	0.227	0.377			1.352	0.018	-7.1
Science	8	558410	0678	14	73	FT	2	A		3670	0.930	0.244	0.582	0.174			0.036	0.344	-0.331	0.128	0.208			0.736	0.029	3.8
Science	8	549689	0679	15	37	MX	2	A	2	7588	1.108	0.274	0.344	0.382			0.034	0.516	-0.486	0.040	0.407			0.232	0.017	-4.0
Science	8	558420	0680	15	73	FT	2	D		3650	0.900	0.335	0.429	0.235			0.015	0.392	-0.353	0.080	0.299			0.767	0.025	3.5
Science	8	549769	0681	16	37	MX	2	A	2	7601	1.044	0.238	0.481	0.281			0.020	0.376	-0.323	0.009	0.295			0.350	0.018	5.9
Science	8	558426	0682	16	73	FT	2	A		3659	0.786	0.377	0.460	0.163			0.038	0.419	-0.381	0.150	0.297			1.075	0.027	0.2
Science	8	549744	0683	17	37	MX	2	B	3	7602	0.968	0.373	0.286	0.341			0.039	0.516	-0.482	0.066	0.428			0.498	0.016	-4.0
Science	8	558436	0684	17	73	FT	2	A		3670	1.201	0.281	0.237	0.482			0.030	0.466	-0.441	0.000	0.397			0.117	0.023	2.2
Science	8	549731	0685	18	37	MX	2	C	3	7611	1.008	0.348	0.297	0.356			0.027	0.466	-0.422	0.025	0.396			0.454	0.016	2.4
Science	8	558446	0686	18	73	FT	2	B		3664	0.291	0.745	0.218	0.037			0.041	0.304	-0.285	0.212	0.194			2.470	0.034	2.2
Science	11	549997	0687	All	9	OP	2	A	3	129722	0.794	0.398	0.411	0.192			0.060	0.457	-0.440	0.189	0.311			0.889	0.004	-6.4
Science	11	554388	0688	All	19	OP	4	A/D	3	129722	1.481	0.270	0.260	0.255	0.149	0.066	0.065	0.614	-0.523	-0.084	0.209	0.303	0.281	0.988	0.003	-9.9
Science	11	554415	0689	All	20	OP	4	A/C	3	129722	0.892	0.403	0.364	0.179	0.048	0.007	0.082	0.566	-0.523	0.144	0.330	0.241	0.111	1.957	0.004	-9.9
Science	11	554346	0690	All	42	OP	2	A	3	129722	1.156	0.199	0.446	0.355			0.035	0.498	-0.421	-0.052	0.405			0.038	0.004	-9.9
Science	11	549915	0691	All	43	OP	2	C	3	129722	0.888	0.494	0.123	0.382			0.152	0.631	-0.607	0.063	0.582			0.603	0.004	-9.9
Science	11	554402	0692	All	44	OP	2	B	3	129722	0.838	0.340	0.482	0.178			0.081	0.494	-0.420	0.104	0.385			0.824	0.005	-9.9
Science	11	554469	0693	All	48	OP	4	A/D	3	129722	1.027	0.447	0.243	0.184	0.088	0.038	0.164	0.608	-0.592	0.125	0.306	0.283	0.219	1.453	0.003	-9.9
Science	11	554294	0694	All	63	OP	2	D	3	129722	0.481	0.609	0.301	0.090			0.073	0.521	-0.506	0.333	0.330			1.655	0.005	-9.9
Science	11	554321	0695	All	64	OP	2	A	3	129722	0.961	0.271	0.498	0.231			0.132	0.535	-0.494	0.120	0.378			0.514	0.004	-9.9
Science	11	558050	0696	1	65	FT	2	A		5346	0.766	0.329	0.575	0.096			0.069	0.487	-0.502	0.355	0.206			1.177	0.025	-6.1

Information												Proportions						Correlations						Rasch		
Cont.	Grade	ID	Pub. ID	Form	Seq.	Status	Pts.	R.C.	D.O.K.	N	Mean	0	1	2	3	4	Omit	Pt. Bis.	0	1	2	3	4	Logit	Logit SE	Fit
Science	11	558063	0697	1	70	FT	4	A/B		5243	1.185	0.348	0.279	0.233	0.118	0.022	0.116	0.665	-0.567	-0.026	0.314	0.367	0.211	1.471	0.016	-9.9
Science	11	558069	0698	2	65	FT	2	A		5277	0.612	0.455	0.479	0.067			0.112	0.466	-0.456	0.342	0.227			1.626	0.025	-5.2
Science	11	558070	0699	2	70	FT	4	A/D		5194	0.605	0.585	0.266	0.112	0.030	0.006	0.099	0.384	-0.332	0.118	0.228	0.173	0.129	2.248	0.019	5.4
Science	11	558080	0700	3	65	FT	2	C		5108	0.110	0.908	0.075	0.018			0.261	0.361	-0.366	0.305	0.196			2.984	0.040	-7.7
Science	11	558086	0701	3	70	FT	4	B/A		5253	1.252	0.387	0.229	0.187	0.139	0.059	0.112	0.621	-0.564	0.022	0.222	0.333	0.273	1.249	0.014	-5.9
Science	11	558095	0702	4	65	FT	2	D		5157	0.176	0.859	0.106	0.035			0.093	0.313	-0.311	0.232	0.200			2.486	0.032	-2.5
Science	11	558099	0703	4	70	FT	4	A/B		5371	0.594	0.597	0.254	0.109	0.036	0.003	0.131	0.483	-0.474	0.237	0.275	0.212	0.066	2.369	0.019	-5.0
Science	11	558104	0704	5	65	FT	2	D		5309	0.800	0.362	0.476	0.162			0.083	0.604	-0.568	0.251	0.401			0.932	0.022	-9.9
Science	11	558109	0705	5	70	FT	4	C/A		5365	0.611	0.585	0.267	0.107	0.035	0.006	0.127	0.591	-0.568	0.266	0.343	0.254	0.118	2.193	0.019	-9.9
Science	11	558116	0706	6	65	FT	2	C		5747	0.743	0.418	0.420	0.162			0.090	0.490	-0.433	0.154	0.374			1.066	0.021	-4.7
Science	11	558121	0707	6	70	FT	4	D/A		5188	1.182	0.289	0.349	0.254	0.106	0.002	0.097	0.650	-0.537	-0.053	0.344	0.378	0.059	2.084	0.017	-9.9
Science	11	558129	0708	7	65	FT	2	C		5200	0.977	0.323	0.376	0.301			0.076	0.610	-0.573	0.111	0.467			0.490	0.020	-9.9
Science	11	558133	0709	7	70	FT	4	A/D		5340	0.657	0.560	0.274	0.122	0.037	0.007	0.114	0.559	-0.541	0.247	0.317	0.237	0.119	2.159	0.018	-9.4
Science	11	558140	0710	8	65	FT	2	B		5272	0.416	0.646	0.293	0.061			0.187	0.442	-0.424	0.305	0.267			1.956	0.025	-5.1
Science	11	558145	0711	8	70	FT	4	A/D		5351	0.704	0.533	0.280	0.143	0.035	0.008	0.150	0.538	-0.484	0.159	0.318	0.251	0.135	2.092	0.018	-5.8
Science	11	558151	0712	9	65	FT	2	B		5268	0.041	0.962	0.035	0.003			0.217	0.236	-0.239	0.225	0.081			3.922	0.066	-4.8
Science	11	558160	0713	9	70	FT	4	A/C		5373	1.059	0.398	0.285	0.199	0.095	0.023	0.081	0.630	-0.584	0.086	0.313	0.317	0.193	1.568	0.016	-9.9
Science	11	558164	0714	10	65	FT	2	A		5294	0.628	0.529	0.315	0.156			0.109	0.473	-0.481	0.281	0.302			1.259	0.021	-2.7
Science	11	558169	0715	10	70	FT	4	B/A		5276	0.555	0.683	0.157	0.104	0.036	0.021	0.165	0.534	-0.514	0.190	0.320	0.230	0.209	1.987	0.018	-6.6
Science	11	558175	0716	11	65	FT	2	A		5176	0.568	0.510	0.412	0.078			0.176	0.592	-0.584	0.425	0.308			1.642	0.024	-9.9
Science	11	558179	0717	11	70	FT	4	B/A		5339	0.870	0.507	0.237	0.153	0.084	0.019	0.147	0.575	-0.525	0.119	0.264	0.329	0.188	1.780	0.016	-5.3
Science	11	558188	0718	12	65	FT	2	A		5302	0.498	0.584	0.335	0.082			0.178	0.476	-0.456	0.305	0.295			1.746	0.024	-5.5
Science	11	558193	0719	12	70	FT	4	A/D		5154	0.604	0.618	0.219	0.114	0.039	0.010	0.158	0.524	-0.523	0.241	0.318	0.218	0.115	2.179	0.019	-3.5

Appendix H:

2008 Science Raw to Scale Score Tables

Table Legend

Column	Description
Raw Score	Raw score
Meas.	Rasch ability measure
Meas. SE	Rasch ability measure standard error
Scaled Score	Scaled score
Scaled Score SE	Scaled score standard error
Freq.	Frequency
Freq. %	Frequency percent
Cum. Freq.	Cumulative frequency
Cum. Freq. %	Cumulative frequency percent
Pctile	Percentile

Grade 4

Raw Score	Meas.	Meas. SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Pctile
0	-5.6820	1.8344	1050	324	0	0.0	0	0.0	0
1	-4.4560	1.0156	1050	180	0	0.0	0	0.0	0
2	-3.7317	0.7287	1050	129	0	0.0	0	0.0	0
3	-3.2953	0.6034	1050	107	0	0.0	0	0.0	0
4	-2.9771	0.5296	1050	94	0	0.0	0	0.0	0
5	-2.7235	0.4799	1050	85	1	0.0	1	0.0	1
6	-2.5111	0.4436	1050	78	5	0.0	6	0.0	1
7	-2.3269	0.4157	1050	73	6	0.0	12	0.0	1
8	-2.1635	0.3935	1050	70	10	0.0	22	0.0	1
9	-2.0158	0.3754	1050	66	19	0.0	41	0.0	1
10	-1.8806	0.3603	1050	64	38	0.0	79	0.1	1
11	-1.7555	0.3475	1050	61	40	0.0	119	0.1	1
12	-1.6386	0.3365	1050	59	99	0.1	218	0.2	1
13	-1.5287	0.3270	1050	58	127	0.1	345	0.3	1
14	-1.4245	0.3186	1050	56	171	0.1	516	0.4	1
15	-1.3254	0.3113	1050	55	270	0.2	786	0.6	1
16	-1.2305	0.3049	1050	54	345	0.3	1131	0.9	1
17	-1.1393	0.2992	1050	53	421	0.3	1552	1.2	1
18	-1.0513	0.2941	1050	52	439	0.3	1991	1.6	1
19	-0.9662	0.2896	1055	51	530	0.4	2521	2.0	2
20	-0.8835	0.2856	1069	50	572	0.5	3093	2.4	2
21	-0.8030	0.2820	1084	50	700	0.6	3793	3.0	3
22	-0.7243	0.2789	1098	49	801	0.6	4594	3.6	3
23	-0.6474	0.2761	1111	49	872	0.7	5466	4.3	4
24	-0.5718	0.2736	1125	48	961	0.8	6427	5.1	5
25	-0.4975	0.2715	1138	48	981	0.8	7408	5.9	5
26	-0.4243	0.2697	1151	48	1118	0.9	8526	6.7	6
27	-0.3520	0.2681	1163	47	1184	0.9	9710	7.7	7
28	-0.2805	0.2668	1176	47	1339	1.1	11049	8.7	8
29	-0.2096	0.2658	1189	47	1419	1.1	12468	9.9	9
30	-0.1392	0.2650	1201	47	1455	1.2	13923	11.0	10
31	-0.0692	0.2644	1213	47	1645	1.3	15568	12.3	12
32	0.0006	0.2640	1226	47	1814	1.4	17382	13.7	13
33	0.0703	0.2639	1238	47	1856	1.5	19238	15.2	14
34	0.1400	0.2641	1250	47	1958	1.5	21196	16.8	16
35	0.2098	0.2644	1263	47	2211	1.7	23407	18.5	18
36	0.2798	0.2650	1275	47	2391	1.9	25798	20.4	19
37	0.3502	0.2658	1288	47	2401	1.9	28199	22.3	21
38	0.4211	0.2668	1300	47	2589	2.0	30788	24.4	23
39	0.4926	0.2681	1313	47	2796	2.2	33584	26.6	25
40	0.5649	0.2697	1325	48	2872	2.3	36456	28.8	28
41	0.6381	0.2715	1338	48	3056	2.4	39512	31.3	30
42	0.7124	0.2736	1352	48	3399	2.7	42911	33.9	33
43	0.7880	0.2761	1365	49	3397	2.7	46308	36.6	35
44	0.8650	0.2789	1379	49	3605	2.9	49913	39.5	38
45	0.9436	0.2820	1392	50	3877	3.1	53790	42.5	41
46	1.0241	0.2856	1407	50	3989	3.2	57779	45.7	44
47	1.1068	0.2896	1421	51	4103	3.2	61882	48.9	47
48	1.1919	0.2941	1436	52	4427	3.5	66309	52.4	51
49	1.2798	0.2991	1452	53	4548	3.6	70857	56.0	54
50	1.3710	0.3048	1468	54	4645	3.7	75502	59.7	58
51	1.4659	0.3113	1485	55	4789	3.8	80291	63.5	62
52	1.5650	0.3186	1502	56	4889	3.9	85180	67.4	65
53	1.6692	0.3269	1521	58	5006	4.0	90186	71.3	69
54	1.7791	0.3364	1540	59	5050	4.0	95236	75.3	73

Grade 4

Raw Score	Meas.	Meas. SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Pctile
55	1.8959	0.3475	1561	61	4846	3.8	100082	79.2	77
56	2.0211	0.3603	1583	64	4822	3.8	104904	83.0	81
57	2.1563	0.3755	1607	66	4482	3.5	109386	86.5	85
58	2.3040	0.3937	1633	70	4088	3.2	113474	89.8	88
59	2.4677	0.4160	1662	74	3768	3.0	117242	92.7	91
60	2.6522	0.4440	1694	78	3162	2.5	120404	95.2	94
61	2.8652	0.4805	1732	85	2473	2.0	122877	97.2	96
62	3.1195	0.5305	1777	94	1711	1.4	124588	98.5	98
63	3.4389	0.6045	1833	107	1081	0.9	125669	99.4	99
64	3.8769	0.7300	1911	129	551	0.4	126220	99.8	99
65	4.6035	1.0169	2039	180	168	0.1	126388	100.0	99
66	5.8317	1.8353	2256	324	38	0.0	126426	100.0	99

Grade 8

Raw Score	Meas.	Meas. SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Pctile
0	-5.6721	1.8345	925	351	0	0.0	0	0.0	0
1	-4.4460	1.0156	925	195	2	0.0	2	0.0	1
2	-3.7218	0.7285	925	140	0	0.0	2	0.0	1
3	-3.2859	0.6028	925	115	2	0.0	4	0.0	1
4	-2.9685	0.5288	925	101	4	0.0	8	0.0	1
5	-2.7159	0.4787	925	92	3	0.0	11	0.0	1
6	-2.5046	0.4422	925	85	9	0.0	20	0.0	1
7	-2.3217	0.4141	925	79	20	0.0	40	0.0	1
8	-2.1596	0.3918	925	75	38	0.0	78	0.1	1
9	-2.0134	0.3735	925	72	79	0.1	157	0.1	1
10	-1.8797	0.3582	925	69	156	0.1	313	0.2	1
11	-1.7561	0.3453	925	66	262	0.2	575	0.4	1
12	-1.6407	0.3343	925	64	436	0.3	1011	0.7	1
13	-1.5322	0.3247	925	62	597	0.4	1608	1.2	1
14	-1.4295	0.3164	925	61	792	0.6	2400	1.7	1
15	-1.3318	0.3090	942	59	1043	0.8	3443	2.5	2
16	-1.2383	0.3026	959	58	1271	0.9	4714	3.4	3
17	-1.1485	0.2969	977	57	1577	1.1	6291	4.6	4
18	-1.0619	0.2918	993	56	1778	1.3	8069	5.9	5
19	-0.9780	0.2873	1009	55	1849	1.3	9918	7.2	7
20	-0.8967	0.2833	1025	54	2012	1.5	11930	8.7	8
21	-0.8174	0.2798	1040	54	2172	1.6	14102	10.2	9
22	-0.7400	0.2767	1055	53	2232	1.6	16334	11.9	11
23	-0.6642	0.2739	1069	52	2405	1.7	18739	13.6	13
24	-0.5899	0.2715	1084	52	2463	1.8	21202	15.4	14
25	-0.5168	0.2694	1098	52	2463	1.8	23665	17.2	16
26	-0.4447	0.2676	1111	51	2629	1.9	26294	19.1	18
27	-0.3735	0.2661	1125	51	2753	2.0	29047	21.1	20
28	-0.3031	0.2648	1139	51	2743	2.0	31790	23.1	22
29	-0.2333	0.2638	1152	51	2930	2.1	34720	25.2	24
30	-0.1639	0.2630	1165	50	2967	2.2	37687	27.4	26
31	-0.0949	0.2624	1178	50	3084	2.2	40771	29.6	28
32	-0.0261	0.2621	1192	50	3240	2.4	44011	31.9	31
33	0.0425	0.2620	1205	50	3340	2.4	47351	34.4	33
34	0.1112	0.2621	1218	50	3317	2.4	50668	36.8	36
35	0.1799	0.2625	1231	50	3452	2.5	54120	39.3	38
36	0.2490	0.2630	1244	50	3580	2.6	57700	41.9	41
37	0.3184	0.2638	1258	51	3730	2.7	61430	44.6	43
38	0.3882	0.2649	1271	51	3807	2.8	65237	47.3	46
39	0.4587	0.2661	1284	51	3928	2.9	69165	50.2	49
40	0.5299	0.2677	1298	51	3991	2.9	73156	53.1	52
41	0.6021	0.2695	1312	52	4218	3.1	77374	56.2	55
42	0.6753	0.2716	1326	52	4179	3.0	81553	59.2	58
43	0.7497	0.2740	1340	52	4363	3.2	85916	62.4	61
44	0.8255	0.2768	1355	53	4333	3.1	90249	65.5	64
45	0.9030	0.2799	1370	54	4417	3.2	94666	68.7	67
46	0.9823	0.2834	1385	54	4285	3.1	98951	71.8	70
47	1.0637	0.2874	1400	55	4259	3.1	103210	74.9	73
48	1.1475	0.2918	1416	56	4152	3.0	107362	77.9	76
49	1.2341	0.2968	1433	57	4115	3.0	111477	80.9	79
50	1.3239	0.3025	1450	58	3903	2.8	115380	83.7	82
51	1.4173	0.3089	1468	59	3656	2.7	119036	86.4	85
52	1.5149	0.3161	1487	61	3369	2.4	122405	88.8	88
53	1.6174	0.3244	1506	62	3136	2.3	125541	91.1	90
54	1.7257	0.3339	1527	64	2731	2.0	128272	93.1	92

Grade 8

Raw Score	Meas.	Meas. SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Pctile
55	1.8408	0.3448	1549	66	2404	1.7	130676	94.8	94
56	1.9640	0.3576	1573	68	2032	1.5	132708	96.3	96
57	2.0973	0.3728	1598	71	1615	1.2	134323	97.5	97
58	2.2429	0.3909	1626	75	1229	0.9	135552	98.4	98
59	2.4043	0.4132	1657	79	912	0.7	136464	99.0	99
60	2.5864	0.4411	1692	84	567	0.4	137031	99.4	99
61	2.7967	0.4776	1732	91	387	0.3	137418	99.7	99
62	3.0479	0.5275	1780	101	211	0.2	137629	99.9	99
63	3.3640	0.6015	1841	115	103	0.1	137732	100.0	99
64	3.7981	0.7272	1924	139	42	0.0	137774	100.0	99
65	4.5203	1.0145	2062	194	16	0.0	137790	100.0	99
66	5.7449	1.8338	2297	351	0	0.0	137790	100.0	100

Grade 11

Raw Score	Meas.	Meas. SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Pctile
0	-5.5355	1.8348	1050	187	0	0.0	0	0.0	0
1	-4.3084	1.0163	1050	103	0	0.0	0	0.0	0
2	-3.5824	0.7300	1050	74	0	0.0	0	0.0	0
3	-3.1440	0.6051	1050	62	3	0.0	3	0.0	1
4	-2.8236	0.5316	1050	54	3	0.0	6	0.0	1
5	-2.5679	0.4822	1050	49	19	0.0	25	0.0	1
6	-2.3531	0.4462	1050	45	41	0.0	66	0.1	1
7	-2.1666	0.4185	1050	43	84	0.1	150	0.1	1
8	-2.0009	0.3965	1050	40	173	0.1	323	0.2	1
9	-1.8509	0.3784	1050	39	300	0.2	623	0.5	1
10	-1.7136	0.3632	1050	37	443	0.3	1066	0.8	1
11	-1.5864	0.3504	1050	36	641	0.5	1707	1.3	1
12	-1.4676	0.3392	1050	35	918	0.7	2625	2.0	2
13	-1.3558	0.3295	1057	34	1123	0.9	3748	2.9	2
14	-1.2501	0.3210	1067	33	1329	1.0	5077	3.9	3
15	-1.1495	0.3134	1078	32	1435	1.1	6512	5.0	4
16	-1.0534	0.3066	1087	31	1653	1.3	8165	6.2	6
17	-0.9614	0.3005	1097	31	1796	1.4	9961	7.6	7
18	-0.8727	0.2950	1106	30	1974	1.5	11935	9.1	8
19	-0.7872	0.2900	1115	30	2049	1.6	13984	10.7	10
20	-0.7044	0.2855	1123	29	2196	1.7	16180	12.3	11
21	-0.6240	0.2814	1131	29	2408	1.8	18588	14.2	13
22	-0.5459	0.2776	1139	28	2565	2.0	21153	16.1	15
23	-0.4698	0.2742	1147	28	2649	2.0	23802	18.1	17
24	-0.3955	0.2711	1154	28	2656	2.0	26458	20.2	19
25	-0.3228	0.2683	1162	27	2868	2.2	29326	22.4	21
26	-0.2515	0.2657	1169	27	3004	2.3	32330	24.6	24
27	-0.1815	0.2634	1176	27	3017	2.3	35347	27.0	26
28	-0.1127	0.2613	1183	27	3322	2.5	38669	29.5	28
29	-0.0449	0.2595	1190	26	3201	2.4	41870	31.9	31
30	0.0220	0.2578	1197	26	3317	2.5	45187	34.5	33
31	0.0881	0.2564	1204	26	3412	2.6	48599	37.1	36
32	0.1535	0.2551	1210	26	3498	2.7	52097	39.7	38
33	0.2183	0.2541	1217	26	3600	2.7	55697	42.5	41
34	0.2827	0.2532	1223	26	3496	2.7	59193	45.1	44
35	0.3466	0.2526	1230	26	3549	2.7	62742	47.8	46
36	0.4103	0.2521	1236	26	3657	2.8	66399	50.6	49
37	0.4738	0.2519	1243	26	3813	2.9	70212	53.5	52
38	0.5372	0.2518	1249	26	3472	2.6	73684	56.2	55
39	0.6006	0.2519	1256	26	3529	2.7	77213	58.9	58
40	0.6642	0.2523	1262	26	3671	2.8	80884	61.7	60
41	0.7280	0.2528	1269	26	3560	2.7	84444	64.4	63
42	0.7921	0.2536	1275	26	3575	2.7	88019	67.1	66
43	0.8567	0.2546	1282	26	3399	2.6	91418	69.7	68
44	0.9218	0.2558	1289	26	3380	2.6	94798	72.3	71
45	0.9876	0.2573	1295	26	3230	2.5	98028	74.7	74
46	1.0543	0.2590	1302	26	3223	2.5	101251	77.2	76
47	1.1219	0.2611	1309	27	3101	2.4	104352	79.6	78
48	1.1906	0.2634	1316	27	2938	2.2	107290	81.8	81
49	1.2607	0.2660	1323	27	2705	2.1	109995	83.9	83
50	1.3322	0.2689	1330	27	2613	2.0	112608	85.9	85
51	1.4053	0.2722	1338	28	2549	1.9	115157	87.8	87
52	1.4804	0.2760	1345	28	2274	1.7	117431	89.5	89
53	1.5577	0.2801	1353	29	2167	1.7	119598	91.2	90
54	1.6375	0.2848	1361	29	1907	1.5	121505	92.6	92

Grade 11

Raw Score	Meas.	Meas. SE	Scaled Score	Scaled Score SE	Freq.	Freq. %	Cum. Freq.	Cum. Freq. %	Pctile
55	1.7201	0.2901	1370	30	1803	1.4	123308	94.0	93
56	1.8059	0.2960	1379	30	1632	1.2	124940	95.3	95
57	1.8955	0.3026	1388	31	1463	1.1	126403	96.4	96
58	1.9893	0.3102	1397	32	1165	0.9	127568	97.3	97
59	2.0882	0.3187	1407	32	1001	0.8	128569	98.0	98
60	2.1928	0.3286	1418	33	803	0.6	129372	98.6	98
61	2.3045	0.3399	1429	35	591	0.5	129963	99.1	99
62	2.4244	0.3531	1442	36	466	0.4	130429	99.4	99
63	2.5545	0.3686	1455	38	313	0.2	130742	99.7	99
64	2.6971	0.3872	1469	39	206	0.2	130948	99.8	99
65	2.8557	0.4100	1485	42	114	0.1	131062	99.9	99
66	3.0352	0.4384	1504	45	50	0.0	131112	100.0	99
67	3.2433	0.4754	1525	48	26	0.0	131138	100.0	99
68	3.4928	0.5260	1550	54	12	0.0	131150	100.0	99
69	3.8075	0.6006	1582	61	6	0.0	131156	100.0	99
70	4.2408	0.7268	1626	74	1	0.0	131157	100.0	99
71	4.9627	1.0145	1700	103	0	0.0	131157	100.0	100
72	6.1874	1.8339	1825	187	0	0.0	131157	100.0	100