

New York State Regents Examination in Geometry

2010 Field Test Analysis, Equating Procedure, and Scaling of Operational Test Forms

Technical Report



Prepared for the New York State Education Department
by Pearson

August 2011

Copyright

Developed and published under contract with the New York State Education Department by Pearson. Copyright © 2010 by the New York State Education Department.

Secure Materials.

All rights reserved. No part of this document may be reproduced or transmitted by any means. Use of these materials is expressly limited to the New York State Education Department.

Table of Contents

Table of Contents.....	i
List of Tables	ii
Section I: Introduction	1
Purpose	1
Section II: Field Test Analysis.....	1
File Merging and Data Clean-up	2
Classical Analysis.....	2
<i>Item Difficulty</i>	3
<i>Point-Biserial Correlation</i>	3
<i>Test Reliability</i>	5
<i>Scoring Reliability</i>	6
<i>Inter-rater Agreement</i>	7
<i>Constructed-Response Item Means and Standard Deviations</i>	13
<i>Intra-class Correlation</i>	13
<i>Weighted Kappa</i>	13
Item Response Theory (IRT) Statistics	14
<i>Item Calibration</i>	15
<i>Item Fit Evaluation</i>	15
Differential Item Functioning (DIF) Statistics.....	17
Section III: Equating Procedure	18
Section IV: Scaling of Operational Test Forms	21
References	23
Appendix A: Classical Item Analysis.....	24
Appendix B: Partial Credit Model Item Analysis.....	36
Appendix C: DIF Statistics	48
Appendix D: Operational Test Maps	55
Appendix E: Scoring Tables.....	65

List of Tables

Table 1. Need/Resource Capacity Category Definitions	1
Table 2. Classical Item Analysis	4
Table 3. Test and Scoring Reliability.....	6
Table 4. Point Differences Between First and Second Reads	8
Table 5. First and Second Read Descriptive Statistics and Agreement	11
Table 6. Partial Credit Model Item Analysis	16
Table 7. Initial Mean Abilities and Equating Constants	20
Table 8. Classical Item Analysis	25
Table 9. Partial Credit Model Item Analysis	37
Table 10. DIF Statistics.....	49
Table 11. Operational Test Map for January 2010.....	56
Table 12. Operational Test Map for June 2010.....	59
Table 13. Operational Test Map for August 2010	62
Table 14. Scoring Tables for January 2010	66
Table 15. Scoring Tables for June 2010	67
Table 16. Scoring Tables for August 2010.....	68

Section I: Introduction

Purpose

The purpose of this report is to document the psychometric work on the New York State Regents Examination in Geometry in 2010. Specifically, contained within this report are procedures for and results of field test analysis, equating, and scaling of operational test forms. Because of a change in vendor mid-year, the field test equating was conducted by Pearson while the scaling was conducted by the previous vendor. Information on test development can be found in the test design and development report for the New York State Regents Examination in Geometry.

Section II: Field Test Analysis

In May 2010, field testing was conducted for the New York State Regents Examination in Geometry to better understand the psychometric quality of the items. The results of this testing are used to help determine which items will be selected for use on operational tests.

Target student samples for participation in this testing were selected such that each would represent the student population expected to take the operational test. The Need/Resource Capacity Categories were used as variables in the sampling plan. See Table 1 for the seven Need/Resource Capacity Categories and their definitions.

Table 1. Need/Resource Capacity Category Definitions

Need/Resource Capacity (N/RC) Category	Definition
High N/RC Districts: New York City	New York City
Large Cities	Buffalo, Rochester, Syracuse, Yonkers
Urban-Suburban	Districts at or above 70 th percentile on the index with at least 100 students per square mile or enrollment greater than 2500
Rural	All districts at or above the 70 th percentile with fewer than 50 students per square mile or enrollment of less than 2500
Average N/RC Districts	All districts between the 20 th and 70 th percentiles on the index
Low N/RC Districts	All districts below the 20 th percentile on the index
Charter Schools	Each charter school is a district

The data collected from field testing were scored by two entities. The multiple-choice items were scored by the New York State Education Department and the constructed-response items were scored by Measurement Incorporated. Therefore, it was necessary to combine data files for data analysis. Both classical and item response theory analyses were conducted using the data to evaluate the quality of the test items.

File Merging and Data Clean-up

Field test forms contained multiple-choice and constructed-response item types. Response data were contained in two separate files. The multiple-choice data file contained 21,294 student records and the constructed-response data file contained 19,848 student records. To combine the two files, the multiple-choice file served as the base file and constructed-response records were merged to the multiple-choice records using unique test booklet numbers. For multiple-choice records that did not have corresponding constructed-response records, constructed-response items were treated as non-attempted and scored as 0. After the exclusion rules were applied, the resulting field test data file contained 19,081 records.

Multiple-choice response data were then compared to the answer key. All item responses not matching the answer key were assigned scores of 0. The responses matching the answer key were assigned scores of 1. With respect to the constructed-response items, scores from 0 to the maximum point value available for each tested item were kept while out of range values were assigned scores of 0. For IRT calibrations, blanks (i.e., missing data) were assigned scores of 0 to be consistent with how operational test items are scored.

The final data file contained both the scored and unscored student responses. Unscored data were used to calculate the percentage of students who selected the various answer choices for the multiple-choice items or the percentage of students who received the range of possible raw score points for the constructed-response items. Thus, the frequency of students leaving items blank can be calculated. The scored data were used for all other analyses.

Classical Analysis

Classical Test Theory is based on the assumption that an observed test score x is composed of both true score t and error score e . This assumption is expressed as follows:

$$x = t + e$$

In other words, error is associated with measuring a student's true score. For example, the choice of test items or the administration conditions may influence student responses making a student's observed score higher or lower than the

student's true score. The error is considered random. After repeated administrations, the mean of the error scores is virtually zero. Thus, a student's observed score is expected to equal his or her true score. This expectation is expressed as follows:

$$E(x) = t$$

Using a Classical Test Theory framework, field test data can be analyzed to provide information about the quality of test items. Item difficulties, point-biserial correlations, reliability estimates, and various statistics related to rater agreement have been calculated and are summarized in the following section.

Item Difficulty

Item difficulty is an indication of student performance on a specific item. Because this examination contains polytomous items, item means are not appropriate for comparing difficulty across items. Instead weighted item means were calculated by dividing an item's mean by the maximum points possible for that item.

For multiple-choice items, the item difficulty is the proportion of students who answer an item correctly. If 90% of the student responses to a multiple-choice item are correct, then this item is considered easier than a multiple-choice item with correct responses by 30% of the students.

Point-Biserial Correlation

The point-biserial correlation is another classical statistic that can be used to evaluate items. For multiple-choice items, it is the correlation between students' performance on a given item (correct or incorrect) and overall performance scores. This statistic is used to evaluate how well an item identifies students who understand the concept being measured and can be generalized for constructed-response items. The possible range for the point-biserial correlation is -1 to 1, with higher values being more desirable.

Table 2 presents a summary of the classical item analysis for each of the field test forms. The first three columns identify the form number, the number of students who took each form, and the number of items on each field test form. The remaining columns are divided into two sections (i.e., item difficulty and point-biserial correlations). Recall that for constructed-response items, item means were divided by the maximum number of points possible in order to place them in the same metric as the multiple-choice items. All item difficulties were below 0.90. With respect to the point-biserial correlations, most of these correlations fell between 0.25 and 0.50.

Table 2. Classical Item Analysis

Form	N-Count	No. of Items	Item Difficulty			Point-Biserial		
			<0.50	0.50 to 0.90	>0.90	<0.25	0.25 to 0.50	>0.50
931	795	10	5	5	0	1	7	2
932	804	10	4	6	0	0	5	5
933	800	10	5	5	0	1	6	3
934	802	10	6	4	0	1	5	4
935	780	10	6	4	0	0	7	3
936	782	10	8	2	0	2	5	3
937	769	10	4	6	0	0	6	4
938	785	10	7	3	0	0	6	4
939	783	10	5	5	0	0	7	3
940	800	10	4	6	0	1	6	3
941	795	10	8	2	0	0	7	3
942	811	10	4	5	0	0	7	2
943	801	10	5	4	0	0	5	4
944	811	10	5	5	0	0	5	5
945	802	10	6	4	0	2	5	3
946	807	10	6	4	0	0	7	3
947	793	10	6	3	0	0	5	4
948	797	10	5	5	0	1	6	3
949	793	10	5	5	0	2	4	4
950	790	10	4	6	0	0	5	5
951	789	10	7	3	0	0	7	3
952	796	10	5	4	0	0	6	3
953	796	10	5	5	0	1	5	4
954	800	10	8	2	0	1	6	3

* For some forms, the item counts in the 'Item Difficulty' and 'Point-Biserial' columns may not sum to the value in the 'No. of Items' column due to 'DNS' (do not score) items.

In addition to the summary information provided in Table 2, all of the classical item statistics are provided in Appendix A. 'Max' is the maximum number of possible points. 'N-Count' refers to the number of student records in the analysis. 'Alpha' contains the internal consistency statistics discussed below. For multiple-choice

items, 'B' represents the proportion of students who left the item blank and 'M1' through 'M4' are the proportions of students who selected each of the four answer choices. For constructed-response items, 'B' represents the proportion of students who left the item blank and 'M0' through 'M6' are the proportions of students who received scores 0 through 6. 'Mean' is the average of the scores received by the students. The final column contains the point-biserial correlation for each item. There are some instances of items missing statistics; this occurs when an item was not scored.

Test Reliability

Classical analysis can also be used to measure the reliability of the test. Reliability is the consistency of the results obtained from a measurement with respect to time or among items or subjects that constitute a test. As such, test reliability can be estimated in a variety of ways. Internal consistency indices are a measure of how consistently examinees respond to items within a test. Two factors influence estimates of internal consistency: test length and homogeneity of items. In general, the more items on the examination the higher the reliability, and the more similar the items are the higher the reliability.

Cronbach's α (alpha) (Cronbach, 1951) has an important use as a measure of the internal consistency of a test. This formula is the extension of an earlier version, the Kuder-Richardson Formula 20 (KR-20), which is the equivalent for dichotomous items.

Table 3 contains the internal consistency statistics for all of the field test forms. These statistics ranged from 0.48 to 0.68 and are based solely on the items in the individual field test forms. It is expected that these statistics associated with the operational tests would be greater because there are more items on the operational test forms.

Table 3. Test and Scoring Reliability

Form Number	Test Reliability	Scoring Reliability
931	0.54	0.84
932	0.65	0.89
933	0.60	0.92
934	0.66	0.90
935	0.62	0.95
936	0.57	0.93
937	0.65	0.94
938	0.63	0.86
939	0.62	0.94
940	0.59	0.94
941	0.63	0.80
942	0.51	0.95
943	0.58	0.94
944	0.68	0.90
945	0.48	0.93
946	0.63	0.89
947	0.61	0.83
948	0.62	0.88
949	0.59	0.86
950	0.66	0.90
951	0.59	0.86
952	0.59	0.92
953	0.58	0.85
954	0.61	0.87

Scoring Reliability

One concern with constructed-response items is the reliability of the scoring process (i.e., consistency of the score assignment). Constructed-response items must be read by scorers who assign scores based on a comparison between the rubric and students' responses. Consistency in the way scores are assigned is a critical part of the reliability of the assessment. To measure this consistency, 10% of the test booklets are scored a second time (i.e., second read scores) and compared to the original set of scores (i.e., first read scores).

As an overall measure of scoring reliability, the Pearson Correlation Coefficient between the first and second scores for each of the constructed-response items was computed. This statistic is often used as an overall indicator of scoring reliability and generally ranges from 0 to near 1. Table 3 contains the results from these analyses in the column headed Scoring Reliability. The correlations ranged from 0.80 to 0.95, indicating high scoring reliability.

Inter-rater Agreement

For each constructed-response item, the difference between the first and second reads was computed. When examining inter-rater agreement statistics, it should be kept in mind that the maximum number of points per item varies as shown in the 'Score Points' column of the following tables.

Table 4 contains the proportion of occurrence of these differences for each item. There were no instances of the first read and second read differing by more than 2.

Table 4. Point Differences Between First and Second Reads

			Difference (First Read minus Second Read)						
Form	Item	Score Points	-3	-2	-1	0	1	2	3
931	8	2	0.00	0.01	0.08	0.84	0.06	0.01	0.00
931	9	2	0.00	0.01	0.10	0.82	0.06	0.00	0.00
931	10	6	0.00	0.03	0.12	0.67	0.12	0.06	0.00
932	8	4	0.00	0.01	0.10	0.84	0.04	0.02	0.00
932	9	2	0.00	0.00	0.03	0.93	0.04	0.00	0.00
932	10	2	0.00	0.00	0.07	0.82	0.10	0.01	0.00
933	8	2	0.00	0.00	0.05	0.87	0.09	0.00	0.00
933	9	6	0.00	0.05	0.13	0.64	0.13	0.05	0.00
933	10	2	0.00	0.00	0.09	0.82	0.08	0.01	0.00
934	8	2	0.00	0.00	0.04	0.88	0.08	0.00	0.00
934	9	4	0.00	0.07	0.09	0.76	0.05	0.04	0.00
934	10	4	0.00	0.01	0.16	0.71	0.12	0.00	0.00
935	8	2	0.00	0.02	0.07	0.87	0.03	0.01	0.00
935	9	2	0.00	0.00	0.05	0.91	0.03	0.02	0.00
935	10	6	0.00	0.02	0.06	0.82	0.07	0.02	0.00
936	8	2	0.00	0.01	0.05	0.90	0.04	0.00	0.00
936	9	4	0.00	0.04	0.07	0.85	0.02	0.03	0.00
936	10	4	0.00	0.02	0.09	0.76	0.12	0.01	0.00
937	8	2	0.00	0.00	0.06	0.86	0.06	0.02	0.00
937	9	2	0.00	0.01	0.09	0.83	0.07	0.00	0.00
937	10	6	0.00	0.01	0.04	0.89	0.04	0.02	0.00
938	8	2	0.00	0.01	0.08	0.89	0.03	0.00	0.00
938	9	4	0.00	0.05	0.13	0.61	0.13	0.09	0.00
938	10	4	0.00	0.08	0.15	0.57	0.15	0.05	0.00
939	8	2	0.00	0.00	0.00	0.99	0.01	0.00	0.00
939	9	2	0.00	0.01	0.12	0.79	0.08	0.00	0.00
939	10	6	0.00	0.02	0.02	0.81	0.09	0.07	0.00
940	8	2	0.00	0.00	0.05	0.88	0.07	0.00	0.00
940	9	4	0.00	0.05	0.07	0.82	0.03	0.04	0.00
940	10	4	0.00	0.03	0.08	0.80	0.09	0.01	0.00
941	8	2	0.00	0.02	0.09	0.80	0.07	0.01	0.00

Table 4. Point Differences Between First and Second Reads (continued)

			Difference (First Read minus Second Read)						
Form	Item	Score Points	-3	-2	-1	0	1	2	3
941	9	4	0.00	0.03	0.08	0.79	0.07	0.04	0.00
941	10	2	0.00	0.02	0.10	0.81	0.06	0.01	0.00
942	8	2	0.00	0.00	0.02	0.98	0.01	0.00	0.00
942	10	4	0.00	0.02	0.02	0.93	0.02	0.01	0.00
943	8	2	0.00	0.00	0.01	0.98	0.01	0.00	0.00
943	9	2	0.00	0.00	0.06	0.84	0.10	0.01	0.00
944	8	2	0.00	0.03	0.06	0.78	0.10	0.03	0.00
944	9	4	0.00	0.04	0.04	0.80	0.09	0.04	0.00
944	10	4	0.00	0.02	0.10	0.77	0.09	0.01	0.00
945	8	2	0.00	0.01	0.09	0.75	0.10	0.06	0.00
945	9	6	0.00	0.02	0.01	0.91	0.04	0.01	0.00
945	10	2	0.00	0.00	0.11	0.79	0.10	0.00	0.00
946	8	2	0.00	0.00	0.05	0.90	0.05	0.00	0.00
946	9	4	0.00	0.05	0.17	0.65	0.11	0.02	0.00
946	10	4	0.00	0.05	0.11	0.72	0.08	0.04	0.00
947	8	2	0.00	0.05	0.07	0.78	0.08	0.02	0.00
947	10	2	0.00	0.01	0.04	0.93	0.02	0.00	0.00
948	8	2	0.00	0.02	0.11	0.76	0.10	0.02	0.00
948	9	4	0.00	0.08	0.08	0.71	0.09	0.04	0.00
948	10	4	0.00	0.05	0.11	0.73	0.09	0.02	0.00
949	8	2	0.00	0.01	0.08	0.80	0.10	0.02	0.00
949	9	2	0.00	0.01	0.06	0.91	0.02	0.00	0.00
949	10	6	0.00	0.09	0.05	0.74	0.09	0.03	0.00
950	8	2	0.00	0.00	0.02	0.95	0.03	0.00	0.00
950	9	4	0.00	0.03	0.12	0.72	0.11	0.02	0.00
950	10	4	0.00	0.01	0.05	0.78	0.12	0.05	0.00
951	8	2	0.00	0.02	0.05	0.79	0.12	0.02	0.00
951	9	6	0.00	0.02	0.08	0.78	0.07	0.05	0.00
951	10	2	0.00	0.03	0.10	0.75	0.09	0.02	0.00
952	9	4	0.00	0.01	0.04	0.87	0.05	0.03	0.00
952	10	2	0.00	0.01	0.04	0.87	0.07	0.01	0.00

Table 4. Point Differences Between First and Second Reads (continued)

			Difference (First Read minus Second Read)						
Form	Item	Score Points	-3	-2	-1	0	1	2	3
953	8	2	0.00	0.00	0.10	0.79	0.11	0.00	0.00
953	9	2	0.00	0.00	0.07	0.90	0.03	0.01	0.00
953	10	6	0.00	0.02	0.11	0.72	0.14	0.02	0.00
954	8	2	0.00	0.04	0.10	0.78	0.02	0.07	0.00
954	9	4	0.00	0.01	0.13	0.77	0.08	0.02	0.00
954	10	4	0.00	0.02	0.11	0.73	0.09	0.06	0.00

Table 5 contains additional summary information regarding the first and second reads. In the fourth column the percent of exact matches between the first and second scores is provided. 'Adj.' is the percentage of differences with a magnitude of one. 'Total' is the sum of the two prior columns and contains values between 86.6% and 100%. These values indicate a high degree of agreement.

Table 5. First and Second Read Descriptive Statistics and Agreement

				Agreement (%)			Raw Score Mean		Raw Score Standard Deviation			
Form	Item	Score Points	Total N-Count	Exact	Adj.	Total	First Read	Second Read	First Read	Second Read	Intra-Class Correlation	Wt Kappa
931	8	2	258	84.5	14.0	98.4	0.2	0.2	0.57	0.54	0.67	0.58
931	9	2	250	82.4	16.8	99.2	1.3	1.4	0.89	0.87	0.87	0.79
931	10	6	258	67.4	23.3	90.7	0.7	0.7	1.15	1.20	0.78	0.59
932	8	4	270	83.7	13.3	97.0	0.4	0.4	0.96	0.95	0.86	0.73
932	9	2	272	93.4	6.6	100.0	0.7	0.7	0.64	0.62	0.92	0.90
932	10	2	260	82.3	16.9	99.2	1.1	1.0	0.90	0.90	0.88	0.81
933	8	2	252	86.5	13.5	100.0	0.9	0.9	0.91	0.93	0.92	0.86
933	9	6	256	64.1	25.8	89.8	1.5	1.6	2.12	2.04	0.92	0.78
933	10	2	270	82.2	17.0	99.3	0.8	0.8	0.89	0.88	0.87	0.80
934	8	2	258	88.4	11.6	100.0	0.6	0.5	0.83	0.81	0.92	0.85
934	9	4	258	76.0	13.2	89.1	1.5	1.6	1.63	1.58	0.89	0.80
934	10	4	258	70.5	28.7	99.2	0.8	0.9	1.07	1.13	0.87	0.72
935	8	2	252	87.3	10.3	97.6	0.3	0.3	0.67	0.73	0.80	0.72
935	9	2	258	90.7	7.8	98.4	0.7	0.7	0.88	0.87	0.91	0.88
935	10	6	260	82.3	13.1	95.4	1.5	1.5	2.10	2.11	0.96	0.89
936	8	2	260	90.0	9.2	99.2	0.9	0.9	0.95	0.96	0.93	0.89
936	9	4	260	84.6	8.5	93.1	0.5	0.5	0.95	1.06	0.83	0.72
936	10	4	258	76.0	20.9	96.9	1.7	1.7	1.72	1.74	0.94	0.85
937	8	2	252	85.7	11.9	97.6	0.9	0.8	0.94	0.93	0.88	0.83
937	9	2	248	83.1	16.1	99.2	0.7	0.8	0.83	0.85	0.86	0.80
937	10	6	256	89.1	7.8	96.9	0.9	0.8	1.87	1.83	0.97	0.90
938	8	2	262	88.5	10.7	99.2	0.6	0.6	0.86	0.88	0.91	0.85
938	9	4	254	60.6	26.0	86.6	2.0	1.9	1.44	1.39	0.80	0.67
938	10	4	246	56.9	30.1	87.0	1.7	1.8	1.52	1.56	0.83	0.67
939	8	2	250	99.2	0.8	100.0	0.8	0.8	0.81	0.80	0.99	0.99
939	9	2	258	79.1	20.2	99.2	0.8	0.9	0.89	0.89	0.86	0.77
939	10	6	258	80.6	10.9	91.5	1.3	1.2	1.98	1.86	0.94	0.85
940	8	2	254	88.2	11.8	100.0	0.9	0.9	0.89	0.88	0.93	0.88
940	9	4	264	81.8	9.8	91.7	2.5	2.5	1.52	1.53	0.91	0.84
940	10	4	258	79.8	16.3	96.1	1.0	1.0	1.38	1.44	0.92	0.82

Table 5. First and Second Read Descriptive Statistics and Agreement (continued)

				Agreement (%)			Raw Score Mean		Raw Score Standard Deviation			
Form	Item	Score Points	Total N-Count	Exact	Adj.	Total	First Read	Second Read	First Read	Second Read	Intra-Class Correlation	Wt Kappa
941	8	2	268	79.9	16.4	96.3	0.9	0.9	0.86	0.88	0.79	0.74
941	9	4	262	78.6	14.5	93.1	0.6	0.6	0.97	0.94	0.77	0.69
941	10	2	270	81.5	15.6	97.0	0.9	0.9	0.92	0.90	0.84	0.78
942	8	2	258	97.7	2.3	100.0	1.5	1.5	0.78	0.76	0.98	0.97
942	10	4	270	92.6	4.4	97.0	0.2	0.2	0.69	0.75	0.85	0.75
943	8	2	270	97.8	2.2	100.0	1.5	1.5	0.83	0.82	0.98	0.97
943	9	2	272	83.8	15.4	99.3	0.8	0.8	0.93	0.92	0.89	0.82
944	8	2	270	77.8	16.3	94.1	0.7	0.6	0.88	0.82	0.73	0.67
944	9	4	268	79.9	12.7	92.5	1.2	1.1	1.64	1.64	0.92	0.83
944	10	4	280	77.1	19.3	96.4	1.1	1.1	1.41	1.44	0.92	0.82
945	8	2	270	74.8	18.5	93.3	1.0	0.8	0.96	0.93	0.75	0.68
945	9	6	268	91.0	5.2	96.3	1.1	1.1	2.13	2.06	0.98	0.93
945	10	2	276	79.0	21.0	100.0	0.8	0.9	0.84	0.85	0.85	0.77
946	8	2	266	90.2	9.8	100.0	0.4	0.4	0.75	0.77	0.92	0.85
946	9	4	260	65.4	27.7	93.1	1.7	1.9	1.45	1.49	0.88	0.75
946	10	4	262	71.8	19.1	90.8	1.0	1.0	1.30	1.24	0.83	0.71
947	8	2	262	77.9	15.3	93.1	0.7	0.7	0.89	0.92	0.74	0.68
947	10	2	254	92.9	6.3	99.2	0.9	0.9	0.84	0.84	0.93	0.91
948	8	2	264	75.8	21.2	97.0	1.1	1.1	0.93	0.94	0.81	0.72
948	9	4	274	70.8	16.8	87.6	1.6	1.6	1.68	1.67	0.88	0.77
948	10	4	262	72.5	20.6	93.1	1.2	1.3	1.54	1.53	0.90	0.78
949	8	2	266	79.7	18.0	97.7	0.6	0.6	0.70	0.69	0.72	0.68
949	9	2	250	91.2	8.0	99.2	0.3	0.4	0.72	0.74	0.90	0.84
949	10	6	256	74.2	14.1	88.3	1.0	1.0	1.54	1.58	0.87	0.74
950	8	2	266	94.7	5.3	100.0	1.2	1.2	0.73	0.74	0.95	0.93
950	9	4	260	71.5	23.1	94.6	1.1	1.1	1.45	1.46	0.89	0.77
950	10	4	258	77.5	17.1	94.6	0.9	0.8	1.39	1.33	0.90	0.78
951	8	2	270	78.5	17.0	95.6	0.8	0.8	0.83	0.83	0.75	0.71
951	9	6	264	78.0	15.2	93.2	0.8	0.8	1.64	1.66	0.92	0.78
951	10	2	260	75.4	19.2	94.6	0.5	0.5	0.76	0.83	0.68	0.60

Table 5. First and Second Read Descriptive Statistics and Agreement (continued)

				Agreement (%)			Raw Score Mean		Raw Score Standard Deviation			
Form	Item	Score Points	Total N-Count	Exact	Adj.	Total	First Read	Second Read	First Read	Second Read	Intra-Class Correlation	Wt Kappa
952	9	4	260	86.9	9.2	96.2	0.8	0.8	1.35	1.30	0.93	0.86
952	10	2	270	87.4	11.1	98.5	0.7	0.7	0.89	0.88	0.89	0.84
953	8	2	264	78.8	21.2	100.0	1.0	1.0	0.84	0.84	0.85	0.77
953	9	2	268	89.6	9.7	99.3	1.2	1.3	0.83	0.81	0.91	0.87
953	10	6	246	71.5	25.2	96.7	0.6	0.5	0.93	0.83	0.76	0.60
954	8	2	264	78.0	11.4	89.4	0.7	0.7	0.93	0.92	0.69	0.65
954	9	4	264	76.5	21.2	97.7	1.0	1.0	1.34	1.35	0.92	0.81
954	10	4	262	72.5	19.8	92.4	1.7	1.6	1.43	1.40	0.88	0.78

* Adj. = difference of one

Constructed-Response Item Means and Standard Deviations

The average score for each constructed-response item was computed based on the first and second reads. In addition, the standard deviation of the scores was computed.

Table 5 contains the means and standard deviations for the first and second read scores. The largest difference between the item means for the first and second scores was 0.2, while there were minimal differences among standard deviation statistics.

Intra-class Correlation

The intra-class correlation was computed for each item. This correlation is an estimate of the reliability of scoring based on an average of the first and second reads. Correlations greater than 0.60 are considered very strong because they explain more than one-third of the variance in scores. All items had intra-class correlations greater than or equal to 0.67 (See Table 5). Consistent with other information provided in the table, these values indicate a very high level of scoring reliability.

Weighted Kappa

Weighted Kappa (Cohen, 1968) was calculated for each item based on the first and second reads. This statistic produces an estimate of the reliability of the score classifications relative to what would be expected to occur by chance.

Weighted Kappa is an estimate of the reliability of the score classifications. That is, the Kappa statistic is a measure of reproducibility for categorical data. Guidelines for the evaluation of this statistic are:

- $k > 0.75$ denotes excellent reproducibility
- $0.4 < k \leq 0.75$ denotes good reproducibility
- $0 < k \leq 0.4$ denotes marginal reproducibility

The results found in Table 5 show a high degree of consistency between the first and second reads. The Weighted Kappa statistics ranged from 0.58 to 0.99, which in all cases indicates good to excellent reproducibility.

Based on the scoring reliability analyses, there is strong evidence that the scoring of the constructed-response items was performed in a highly reliable manner.

Item Response Theory (IRT) Statistics

As discussed above, the item mean is a statistic used to evaluate item difficulty. However, many different test forms are used during field testing and different samples of students are responding to these items. The average ability of the different samples of students varies and a direct comparison of item means across test forms may lead to inaccurate interpretations. Therefore, Item Response Theory (IRT) was also used to evaluate item difficulty.

Specifically, the Rasch Partial Credit Model (PCM) (Masters, 1982) was used. With use of this model, the difficulty of items and the ability of examinees are placed on the same metric. Thus, the difficulty of an item and the ability of a person can be meaningfully compared across field test forms. Also, the use of this model provides greater flexibility in situations where different samples or test forms are used because the parameters generated are generally not considered to be sample dependent or test dependent. A description of this model, results of item calibration, and item fit evaluation are below.

The PCM provides an overall difficulty estimate for each item. Specifically for constructed-response items when there are several points possible, individual estimates of difficulty for each of the possible score points are also calculated (i.e., step values). Each step value represents the difficulty of a student receiving a particular score point given that they have already received the prior score point. For example, if a 3-point item had step values of -1.0, 1.0, and 0.0, one could say that it is relatively easy to obtain a score of 1. However, it is much more difficult to obtain a 2 given the student has the ability to score a 1 because the difference in difficulty between a 1 and a 2 is much greater than the difference between a 0 and a 1. Also, the difference between a 2 and a 3 is not as great as the difference between a 1 and a 2. Thus, with this example, a small step is needed to go from a 0 to a 1, a large

step is needed to move from a 1 to a 2, and a moderate step is needed to proceed from a 2 to a 3.

Item Calibration

As discussed above, the use of Rasch item difficulty statistics provides an advantage over the use of classical item means because they can be compared across test forms. Different samples of students responded to the various test forms. Although the samples were selected to be similar with respect to student ability, there are differences. By equating the test forms (See the Equating Procedure section below), the Rasch item difficulties account for those differences and these statistics can be compared across test forms.

Rasch item difficulty values generally range from -3.00 to +3.00. An item with a Rasch difficulty greater than +2.0 is considered very difficult and should be examined carefully. If the item is measuring an important concept that students are having difficulty with, then the item can be useful. However, if the item is measuring a trivial concept or is written in a confusing manner, then it may not be appropriate to use on an operational test form. Likewise, any item with a Rasch difficulty less than -2.0 is considered very easy and usually provides little information regarding student achievement. The vast majority of test items should range between -2.0 and +2.0. This range represents approximately two standard deviations around the average difficulty of 0. Thus, one would expect that, based on chance, roughly 5% of the items will fall outside of that range and therefore, these are items that should be closely examined for content.

Item Fit Evaluation

The INFIT statistic is used to determine whether items are functioning in a way that is congruent with the assumptions of the Rasch model. Under these assumptions, how a student will respond to an item depends on the proficiency of the student and the difficulty of the item, both of which are on the same measurement scale. If an item is as difficult as a student is able, the student will have a 50% chance of getting the item correct. If a student is more able than an item is difficult, under the assumptions of the Rasch model, that student has a greater than 50% chance of correctly answering the item. On the other hand, if the item is more difficult than the student is able, he or she has a less than 50% chance of correctly responding to the item. Rasch fit statistics estimate the extent to which an item is functioning in this predicted manner. Items showing a poor fit with the Rasch model typically have values outside the range of 0.7 to 1.3.

Table 6 contains a summary of the Partial Credit Model item analysis for each of the field test forms. The first column lists the form numbers. The next two columns list the number of students who participated and the number of items on each field

test form. The remaining columns are divided into two sections. The first section pertains to the Rasch item difficulties while the second pertains to the INFIT statistics. Nearly all of the items fell within the moderate -2.0 to +2.0 difficulty range and no items had an INFIT statistic outside the typical range.

Table 6. Partial Credit Model Item Analysis

Form	N-Count	No. of Items	Rasch			INFIT		
			<-2.0	-2.0 to 2.0	>2.0	<-0.70	-0.70 to 1.30	>1.30
931	795	10	0	8	2	0	10	0
932	804	10	0	9	1	0	10	0
933	800	10	0	10	0	0	10	0
934	802	10	0	10	0	0	10	0
935	780	10	0	9	1	0	10	0
936	782	10	0	10	0	0	10	0
937	769	10	0	10	0	0	10	0
938	785	10	0	9	1	0	10	0
939	783	10	0	10	0	0	10	0
940	800	10	0	10	0	0	10	0
941	795	10	0	10	0	0	10	0
942	811	10	0	8	1	0	9	0
943	801	10	0	9	0	0	9	0
944	811	10	0	9	1	0	10	0
945	802	10	0	10	0	0	10	0
946	807	10	0	10	0	0	10	0
947	793	10	0	9	0	0	9	0
948	797	10	0	10	0	0	10	0
949	793	10	0	10	0	0	10	0
950	790	10	0	10	0	0	10	0
951	789	10	0	10	0	0	10	0
952	796	10	0	9	0	0	9	0
953	796	10	0	9	1	0	10	0
954	800	10	0	10	0	0	10	0

* For some forms, the item counts in the 'Rasch' and 'INFIT' columns may not sum to the value in the 'No. of Items' column due to 'DNS' (do not score) items.

All of the individual IRT item statistics are provided in Appendix B. The column titled RID contains the Rasch item difficulty statistics. S1–S6 contain the step values for the constructed-response items. Finally, INFIT contains the INFIT statistic for each item.

Differential Item Functioning (DIF) Statistics

Statistical procedures are employed to observe whether, on the basis of data, there exists the possibility of unfair treatment of different populations. DIF statistics are used to identify items for which members of a focal group have a different probability of getting the items correct than members of a reference group after the groups have been matched on ability level on the test.

For the multiple-choice items, the Mantel-Haenszel Delta (MHD) DIF statistics were computed (Dorans & Holland, 1992) to classify test items in three levels of DIF for each comparison: negligible DIF (A), moderate DIF (B), and large DIF (C). An item was flagged if it exhibited a B or C category of DIF using the following rules derived from National Assessment of Educational Progress (NAEP) guidelines (Allen, Carlson, & Zalanak, 1999):

- MHD not significantly different from 0 (based on $\alpha = 0.05$) **or** $|MHD| < 1.0$ are classified as A.
- MHD significantly different from 0 and $\{|MHD| \geq 1.0 \text{ and } < 1.5\}$ **or** MHD not significantly different from 0 and $|MHD| \geq 1.0$ are classified as B.
- $|MHD| \geq 1.5$ and significantly different from 0 are classified as C.

For the constructed-response items, the effect size of the standardized mean difference (SMD) was used to flag DIF. The SMD reflects the size of the differences in performance on constructed-response items between student groups matched on the total score. It is the difference between the unweighted item mean of the focal group and the weighted item mean of the reference group. The weights applied to the reference group are applied so that the weighted number of reference group students is the same as in the focal group (within the same ability group). The SMD is divided by the total group item standard deviation to get a measure of the effect size (ES) for the SMD. The SMD effect size groups each item into one of three categories: negligible DIF (AA), moderate DIF (BB), and large DIF (CC). Only categories BB and CC were flagged in the results.

- Probability is > 0.05 **or** if $|ES| \leq 0.17$, classified as AA.
- Probability is > 0.05 and if $0.17 < |ES| \leq 0.25$, classified as BB.
- Probability is > 0.05 and if $|ES| > 0.25$, classified as CC.

Although DIF statistics are typically conducted by gender and ethnicity, the low n-counts for ethnic subgroups did not allow for these statistics to be meaningful. The n-counts for gender allowed for comparisons to be made, but were still somewhat low so resulting statistics should be interpreted with caution.

The DIF statistics for gender are shown in Appendix C. Flagging of items appears in the 'DIF Category' column and if an item is flagged, the 'Favored Group' column indicates which gender is favored.

Section III: Equating Procedure

The 2010 field test administration for the New York State Regents Examination in Geometry consisted of 22 field test forms numbered 933–954 and two anchor forms labeled 931 and 932. The field test forms contained multiple-choice and constructed-response items. All students participating in the field test were administered one of the 24 test forms. The test forms were spiraled within the classroom so that the groups of students taking each form were equivalent. A complete listing of these field test forms can be seen in Appendix A where item type (e.g., multiple-choice, constructed-response) and the maximum points for each item are displayed.

The anchor forms were equated to the item bank using a common-item equating design. The anchor item difficulty parameters were fixed to their 2009 item bank values. This places the item difficulty estimates and the ability estimates of the students taking the anchor forms onto the item bank scale. After the anchor forms were placed onto the bank scale, the average of the two mean ability estimates for the two forms was computed using ability estimates of non-extreme students. This average ability estimate was used to equate the remaining field test forms as well as updating the item parameters for the anchor forms.

As part of the anchor item equating, an item-stability check was performed. After fixing all of the items to their 2009 bank values, any item with a displacement value with a magnitude greater than 0.30 was no longer fixed and the test form was reanalyzed. If more than one item had a displacement value with a magnitude greater than 0.30, then the item with the largest displacement was freed and the test form was reanalyzed. In a stepwise fashion, this procedure was repeated until all remaining fixed anchor items had displacements with magnitudes less than or equal to 0.30.

Applying the anchor item-stability check to each of the anchor forms resulted in a few items having a displacement value with a magnitude greater than 0.30. For form 931, two items were freed and for form 932 one item was freed. This indicates a strong level of stability in the items used on the anchor forms.

The equated mean ability estimate for form 931 was 0.44 and for form 932 the mean was 0.52. This produced a target mean ability estimate of 0.48 which was used for the remainder of the equating process.

After the anchor forms were equated and the target mean was computed, the field test forms were equated using the equivalent groups design. The first step was to calibrate each form separately where all the item parameters were free to estimate (without constraint). From those initial calibrations, the mean ability estimates for each field test form were obtained. The second step was to determine the equating constant for each form by subtracting the mean ability for a given field test form from the target mean ability calculated from the anchor forms (i.e., forms 931 and 932). The respective equating constant was then added to each of the item parameters on a given form. If the resulting mean of the ability estimates for those students did not equal that of the target mean, then the procedure was repeated until the mean abilities for each of the field test forms equaled the target mean ability. Table 7 shows the mean abilities and constants used for the equating.

Table 7. Initial Mean Abilities and Equating Constants

Form Number	Mean Ability	Constant
931	-0.18	0.64
932	0.05	0.42
933	0.31	0.18
934	-0.18	0.63
935	-0.17	0.62
936	-0.54	0.97
937	0.04	0.42
938	-0.41	0.85
939	0.40	0.08
940	0.16	0.31
941	-0.36	0.80
942	0.18	0.28
943	0.13	0.31
944	-0.14	0.59
945	-0.16	0.61
946	-0.04	0.50
947	-0.17	0.59
948	-0.08	0.53
949	-0.37	0.81
950	-0.08	0.54
951	-0.39	0.83
952	-0.31	0.71
953	0.18	0.30
954	-0.48	0.91

The equated item parameters for the field test items can now be compared across test forms since the equating process places all items on the same scale. In addition, when items are combined to form unique operational test forms, raw score to scale score tables can be generated based on these parameters. The following section contains a description of the development of the operational test forms and scoring tables.

Section IV: Scaling of Operational Test Forms

Operational test items are selected based on content coverage, content accuracy, and statistical quality. The sets of items on each operational test conform to the coverage suggested by content experts. These expert judgments are based on the learning standards established by the New York State Education Department. With respect to statistical quality, classical and Rasch statistics are examined to determine how well items function. Also, items are selected such that they range in difficulty in order to measure students across ability levels. Appendix D contains the 2010 operational test maps with content information regarding each item included on the January 2010, June 2010, and August 2010 operational test forms.

In order to limit wide fluctuations of raw scores that correspond to scale scores of 65 and 85 across administrations, the average Rasch item difficulty for the operational test is considered. For this examination, an average Rasch difficulty of approximately 0.279 is used as a target for each administration. In most cases, meeting this target will provide raw scores of similar magnitude to other forms. However, differences with these scores also occur due to the distribution of the Rasch item difficulty parameters.

Scoring tables display the relationship between raw scores on the operational test and assigned scale scores. Appendix E contains the scoring tables used for the January, June, and August 2010 operational test forms. Four steps are taken in order to produce these tables and resulting conversion charts.

The first step is to develop a raw score (i.e., number of points on the test form) to theta (i.e., student ability) to scale score relationship for the baseline operational test form. This relationship is determined when standards are set and then used for every administration moving forward until the standards are revisited. The baseline form was determined by the New York State Education Department to be June 2009. The raw score to theta relationship from that examination was used and then scale scores are calculated based on the raw score cuts according to the following formula:

$$p(x) = m_3x^3 + m_2x^2 + m_1x + m_0$$

The raw score of zero was assigned a scale score of zero and the maximum raw score was assigned a scale score of 100. The raw scores corresponding to the scale scores of 65 and 85 were also fixed. The polynomial relationship shown above was then used to assign all scale scores to the remaining raw scores. The resulting values for $m_1 - m_3$ are the transformation constants used to produce the final raw score to scale score table.

The second step is to develop a raw score to theta relationship for the new operational test form using the field test equated PCM item parameters. This is

accomplished by doing a calibration where all items are anchored to their field test parameters. One modification that is made is that for 6-point items, a constant based on existing bank values is used in place of the field test parameters. The number of points on the test form (i.e., raw score) expected across student ability levels is based on the difficulty of the items on the form. Thus, given a particular student ability level (i.e., theta), if the points are more difficult to earn on the new test than the points on the June 2009 test, the number of points expected of this student on the new test will be less than the number of points expected of this student on the baseline form.

The third step is to use linear interpolation to determine the raw score to theta to scale score relationship for the new test. The theta values associated with scale scores of 65 and 85 on the baseline form are used along with the raw score to theta relationship developed in the previous step. In other words, the baseline 65 and 85 theta values are used as reference points and linear interpolation assigns the other scale scores.

Finally, a conversion chart is created based on the scoring table generated in the third step. Scale scores are rounded to the nearest whole number in all cases except for 0, 65, 85, and 100. A raw score of zero is assigned a scale score of zero. The maximum raw score is assigned a scale score of 100. With respect to 65 and 85 scale scores, the raw scores with scale scores of 65 or 85 after rounding are assigned those values.

References

- Allen, N.L., Carlson, J.E., and Zalanak, C.A. 1999. *The NAEP 1996 Technical Report*. Washington, DC: National Center for Education Statistics.
- Cohen, J. 1968. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70, 213–20.
- Cronbach, L.J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Dorans, N.J. and Holland, P.W. 1992. DIF Detection and Description: Mantel–Haenszel and Standardization. In *Differential Item Functioning: Theory and Practice*, edited by P.W. Holland and H. Wainer, 35–66. Hillsdale, NJ: Erlbaum.
- Masters, G. N. 1982. A Rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.

Appendix A: Classical Item Analysis

Table 8. Classical Item Analysis

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	931	MC	01	1	795	0.54	0.00		0.18	0.48	0.23	0.11			0.48	0.46
2010_Geom_FT	931	MC	02	1	795	0.54	0.01		0.78	0.04	0.02	0.16			0.78	0.38
2010_Geom_FT	931	MC	03	1	795	0.54	0.01		0.58	0.10	0.19	0.12			0.58	0.28
2010_Geom_FT	931	MC	04	1	795	0.54	0.01		0.06	0.19	0.18	0.55			0.55	0.45
2010_Geom_FT	931	MC	05	1	795	0.54	0.04		0.38	0.24	0.23	0.12			0.24	0.21
2010_Geom_FT	931	MC	06	1	795	0.54	0.01		0.26	0.05	0.59	0.10			0.59	0.45
2010_Geom_FT	931	MC	07	1	795	0.54	0.01		0.43	0.06	0.47	0.03			0.47	0.39
2010_Geom_FT	931	CR	08	2	795	0.54	0.17	0.67	0.09	0.06					0.22	0.42
2010_Geom_FT	931	CR	09	2	795	0.54	0.04	0.26	0.10	0.60					1.30	0.61
2010_Geom_FT	931	CR	10	6	795	0.54	0.11	0.53	0.18	0.10	0.03	0.02	0.02	0.02	0.72	0.68
2010_Geom_FT	932	MC	01	1	804	0.65	0.00		0.18	0.04	0.06	0.71			0.71	0.50
2010_Geom_FT	932	MC	02	1	804	0.65	0.00		0.05	0.29	0.64	0.02			0.64	0.53
2010_Geom_FT	932	MC	03	1	804	0.65	0.00		0.14	0.52	0.21	0.13			0.52	0.33
2010_Geom_FT	932	MC	04	1	804	0.65	0.01		0.69	0.07	0.11	0.11			0.69	0.51
2010_Geom_FT	932	MC	05	1	804	0.65	0.01		0.15	0.14	0.17	0.53			0.53	0.49
2010_Geom_FT	932	MC	06	1	804	0.65	0.00		0.13	0.07	0.44	0.35			0.44	0.33
2010_Geom_FT	932	MC	07	1	804	0.65	0.01		0.31	0.58	0.07	0.03			0.58	0.46
2010_Geom_FT	932	CR	08	4	804	0.65	0.06	0.74	0.08	0.02	0.09	0.01			0.43	0.59
2010_Geom_FT	932	CR	09	2	804	0.65	0.05	0.33	0.46	0.15					0.77	0.56
2010_Geom_FT	932	CR	10	2	804	0.65	0.11	0.28	0.23	0.38					0.99	0.63
2010_Geom_FT	933	MC	01	1	800	0.60	0.02		0.14	0.73	0.07	0.05			0.73	0.41

Table 8. Classical Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	933	MC	02	1	800	0.60	0.00		0.01	0.06	0.88	0.04			0.88	0.31
2010_Geom_FT	933	MC	03	1	800	0.60	0.01		0.43	0.17	0.03	0.37			0.43	0.39
2010_Geom_FT	933	MC	04	1	800	0.60	0.01		0.24	0.31	0.26	0.19			0.26	0.13
2010_Geom_FT	933	MC	05	1	800	0.60	0.01		0.11	0.17	0.12	0.60			0.60	0.34
2010_Geom_FT	933	MC	06	1	800	0.60	0.01		0.67	0.11	0.05	0.16			0.67	0.50
2010_Geom_FT	933	MC	07	1	800	0.60	0.02		0.17	0.76	0.02	0.03			0.76	0.33
2010_Geom_FT	933	CR	08	2	800	0.60	0.11	0.37	0.16	0.36					0.87	0.58
2010_Geom_FT	933	CR	09	6	800	0.60	0.01	0.44	0.27	0.05	0.07	0.04	0.04	0.09	1.42	0.84
2010_Geom_FT	933	CR	10	2	800	0.60	0.11	0.37	0.22	0.30					0.82	0.68
2010_Geom_FT	934	MC	01	1	802	0.66	0.00		0.14	0.23	0.20	0.43			0.43	0.56
2010_Geom_FT	934	MC	02	1	802	0.66	0.00		0.11	0.53	0.23	0.12			0.53	0.32
2010_Geom_FT	934	MC	03	1	802	0.66	0.01		0.17	0.10	0.16	0.56			0.56	0.37
2010_Geom_FT	934	MC	04	1	802	0.66	0.01		0.67	0.19	0.08	0.06			0.67	0.47
2010_Geom_FT	934	MC	05	1	802	0.66	0.02		0.15	0.39	0.22	0.23			0.39	0.35
2010_Geom_FT	934	MC	06	1	802	0.66	0.01		0.17	0.22	0.53	0.07			0.53	0.21
2010_Geom_FT	934	MC	07	1	802	0.66	0.01		0.31	0.26	0.03	0.39			0.39	0.39
2010_Geom_FT	934	CR	08	2	802	0.66	0.07	0.53	0.15	0.26					0.66	0.67
2010_Geom_FT	934	CR	09	4	802	0.66	0.08	0.37	0.05	0.19	0.09	0.21			1.57	0.75
2010_Geom_FT	934	CR	10	4	802	0.66	0.14	0.35	0.27	0.13	0.03	0.09			0.95	0.73
2010_Geom_FT	935	MC	01	1	780	0.62	0.01		0.07	0.12	0.50	0.30			0.50	0.51
2010_Geom_FT	935	MC	02	1	780	0.62	0.01		0.14	0.33	0.13	0.39			0.39	0.32
2010_Geom_FT	935	MC	03	1	780	0.62	0.01		0.57	0.18	0.13	0.11			0.57	0.43

Table 8. Classical Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	935	MC	04	1	780	0.62	0.02		0.13	0.14	0.35	0.36			0.36	0.42
2010_Geom_FT	935	MC	05	1	780	0.62	0.01		0.13	0.78	0.05	0.03			0.78	0.41
2010_Geom_FT	935	MC	06	1	780	0.62	0.01		0.66	0.04	0.18	0.11			0.66	0.34
2010_Geom_FT	935	MC	07	1	780	0.62	0.02		0.13	0.38	0.39	0.09			0.39	0.43
2010_Geom_FT	935	CR	08	2	780	0.62	0.05	0.79	0.06	0.10					0.27	0.45
2010_Geom_FT	935	CR	09	2	780	0.62	0.16	0.33	0.18	0.32					0.83	0.64
2010_Geom_FT	935	CR	10	6	780	0.62	0.15	0.38	0.04	0.18	0.04	0.06	0.02	0.12	1.64	0.87
2010_Geom_FT	936	MC	01	1	782	0.57	0.01		0.65	0.18	0.08	0.07			0.65	0.48
2010_Geom_FT	936	MC	02	1	782	0.57	0.00		0.27	0.26	0.13	0.34			0.34	0.50
2010_Geom_FT	936	MC	03	1	782	0.57	0.00		0.06	0.59	0.23	0.11			0.59	0.42
2010_Geom_FT	936	MC	04	1	782	0.57	0.00		0.23	0.22	0.51	0.04			0.22	0.16
2010_Geom_FT	936	MC	05	1	782	0.57	0.00		0.17	0.33	0.33	0.17			0.33	0.11
2010_Geom_FT	936	MC	06	1	782	0.57	0.02		0.27	0.16	0.28	0.27			0.27	0.25
2010_Geom_FT	936	MC	07	1	782	0.57	0.00		0.09	0.31	0.34	0.26			0.34	0.48
2010_Geom_FT	936	CR	08	2	782	0.57	0.05	0.44	0.11	0.40					0.90	0.58
2010_Geom_FT	936	CR	09	4	782	0.57	0.13	0.64	0.09	0.04	0.07	0.02			0.48	0.61
2010_Geom_FT	936	CR	10	4	782	0.57	0.07	0.34	0.16	0.07	0.07	0.29			1.68	0.77
2010_Geom_FT	937	MC	01	1	769	0.65	0.01		0.59	0.13	0.16	0.10			0.59	0.56
2010_Geom_FT	937	MC	02	1	769	0.65	0.00		0.19	0.07	0.64	0.09			0.64	0.48
2010_Geom_FT	937	MC	03	1	769	0.65	0.01		0.06	0.20	0.13	0.60			0.60	0.47
2010_Geom_FT	937	MC	04	1	769	0.65	0.01		0.07	0.54	0.23	0.14			0.54	0.46
2010_Geom_FT	937	MC	05	1	769	0.65	0.01		0.22	0.61	0.05	0.12			0.61	0.41

Table 8. Classical Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	937	MC	06	1	769	0.65	0.01		0.43	0.06	0.25	0.26			0.43	0.41
2010_Geom_FT	937	MC	07	1	769	0.65	0.01		0.31	0.01	0.65	0.02			0.65	0.32
2010_Geom_FT	937	CR	08	2	769	0.65	0.08	0.50	0.09	0.33					0.76	0.62
2010_Geom_FT	937	CR	09	2	769	0.65	0.08	0.43	0.23	0.26					0.75	0.62
2010_Geom_FT	937	CR	10	6	769	0.65	0.11	0.66	0.06	0.01	0.02	0.02	0.03	0.08	0.89	0.76
2010_Geom_FT	938	MC	01	1	785	0.63	0.00		0.22	0.12	0.17	0.50			0.50	0.53
2010_Geom_FT	938	MC	02	1	785	0.63	0.01		0.15	0.15	0.47	0.21			0.21	0.32
2010_Geom_FT	938	MC	03	1	785	0.63	0.00		0.69	0.20	0.04	0.06			0.69	0.35
2010_Geom_FT	938	MC	04	1	785	0.63	0.01		0.36	0.16	0.34	0.13			0.36	0.41
2010_Geom_FT	938	MC	05	1	785	0.63	0.00		0.42	0.42	0.10	0.05			0.42	0.37
2010_Geom_FT	938	MC	06	1	785	0.63	0.01		0.06	0.62	0.11	0.21			0.62	0.40
2010_Geom_FT	938	MC	07	1	785	0.63	0.04		0.14	0.19	0.50	0.14			0.14	0.26
2010_Geom_FT	938	CR	08	2	785	0.63	0.12	0.57	0.10	0.21					0.52	0.62
2010_Geom_FT	938	CR	09	4	785	0.63	0.04	0.17	0.18	0.26	0.20	0.15			1.89	0.64
2010_Geom_FT	938	CR	10	4	785	0.63	0.11	0.24	0.18	0.13	0.14	0.20			1.65	0.79
2010_Geom_FT	939	MC	01	1	783	0.62	0.01		0.13	0.48	0.26	0.14			0.48	0.31
2010_Geom_FT	939	MC	02	1	783	0.62	0.01		0.52	0.21	0.11	0.15			0.52	0.48
2010_Geom_FT	939	MC	03	1	783	0.62	0.00		0.04	0.84	0.05	0.07			0.84	0.33
2010_Geom_FT	939	MC	04	1	783	0.62	0.00		0.05	0.14	0.75	0.06			0.75	0.35
2010_Geom_FT	939	MC	05	1	783	0.62	0.00		0.07	0.05	0.83	0.05			0.83	0.41
2010_Geom_FT	939	MC	06	1	783	0.62	0.01		0.07	0.76	0.05	0.10			0.76	0.49
2010_Geom_FT	939	MC	07	1	783	0.62	0.02		0.42	0.16	0.20	0.20			0.42	0.42

Table 8. Classical Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	939	CR	08	2	783	0.62	0.11	0.31	0.35	0.23					0.82	0.67
2010_Geom_FT	939	CR	09	2	783	0.62	0.02	0.45	0.20	0.32					0.84	0.60
2010_Geom_FT	939	CR	10	6	783	0.62	0.09	0.57	0.04	0.11	0.06	0.04	0.02	0.08	1.16	0.80
2010_Geom_FT	940	MC	01	1	800	0.59	0.00		0.20	0.53	0.08	0.19			0.53	0.33
2010_Geom_FT	940	MC	02	1	800	0.59	0.00		0.07	0.11	0.62	0.20			0.62	0.33
2010_Geom_FT	940	MC	03	1	800	0.59	0.00		0.70	0.09	0.08	0.12			0.70	0.42
2010_Geom_FT	940	MC	04	1	800	0.59	0.01		0.32	0.34	0.20	0.14			0.34	0.22
2010_Geom_FT	940	MC	05	1	800	0.59	0.01		0.38	0.10	0.27	0.24			0.38	0.41
2010_Geom_FT	940	MC	06	1	800	0.59	0.01		0.10	0.06	0.06	0.78			0.78	0.46
2010_Geom_FT	940	MC	07	1	800	0.59	0.01		0.07	0.06	0.66	0.20			0.66	0.34
2010_Geom_FT	940	CR	08	2	800	0.59	0.03	0.45	0.23	0.30					0.83	0.55
2010_Geom_FT	940	CR	09	4	800	0.59	0.05	0.17	0.05	0.23	0.08	0.43			2.46	0.73
2010_Geom_FT	940	CR	10	4	800	0.59	0.13	0.41	0.18	0.05	0.07	0.17			1.14	0.72
2010_Geom_FT	941	MC	01	1	795	0.63	0.02		0.42	0.09	0.42	0.05			0.42	0.32
2010_Geom_FT	941	MC	02	1	795	0.63	0.01		0.24	0.37	0.25	0.13			0.37	0.27
2010_Geom_FT	941	MC	03	1	795	0.63	0.01		0.11	0.17	0.48	0.23			0.23	0.48
2010_Geom_FT	941	MC	04	1	795	0.63	0.02		0.15	0.61	0.17	0.06			0.61	0.41
2010_Geom_FT	941	MC	05	1	795	0.63	0.02		0.40	0.23	0.34	0.01			0.34	0.42
2010_Geom_FT	941	MC	06	1	795	0.63	0.01		0.05	0.12	0.06	0.76			0.76	0.41
2010_Geom_FT	941	MC	07	1	795	0.63	0.03		0.24	0.11	0.47	0.14			0.47	0.34
2010_Geom_FT	941	CR	08	2	795	0.63	0.13	0.42	0.22	0.23					0.69	0.64
2010_Geom_FT	941	CR	09	4	795	0.63	0.04	0.57	0.22	0.07	0.06	0.03			0.68	0.75

Table 8. Classical Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	941	CR	10	2	795	0.63	0.08	0.44	0.14	0.34					0.82	0.62
2010_Geom_FT	942	MC	01	1	811	0.51	0.00		0.14	0.40	0.26	0.19			0.40	0.47
2010_Geom_FT	942	MC	02	1	811	0.51	0.00		0.63	0.29	0.03	0.04			0.63	0.41
2010_Geom_FT	942	MC	03	1	811	0.51	0.00		0.32	0.08	0.30	0.30			0.32	0.25
2010_Geom_FT	942	MC	04	1	811	0.51	0.00		0.17	0.44	0.27	0.11			0.44	0.42
2010_Geom_FT	942	MC	05	1	811	0.51	0.01		0.05	0.06	0.79	0.09			0.79	0.34
2010_Geom_FT	942	MC	06	1	811	0.51	0.00		0.03	0.14	0.04	0.78			0.78	0.44
2010_Geom_FT	942	MC	07	1	811	0.51	0.01		0.17	0.14	0.64	0.04			0.64	0.49
2010_Geom_FT	942	CR	08	2	811	0.51	0.04	0.14	0.15	0.67					1.48	0.64
2010_Geom_FT	942	CR	09													
2010_Geom_FT	942	CR	10	4	811	0.51	0.09	0.79	0.02	0.07	0.01	0.02			0.25	0.53
2010_Geom_FT	943	MC	01	1	801	0.58	0.00		0.09	0.12	0.67	0.12			0.67	0.51
2010_Geom_FT	943	MC	02	1	801	0.58	0.00		0.28	0.57	0.10	0.04			0.57	0.55
2010_Geom_FT	943	MC	03	1	801	0.58	0.01		0.07	0.41	0.26	0.26			0.26	0.45
2010_Geom_FT	943	MC	04	1	801	0.58	0.01		0.06	0.09	0.38	0.46			0.38	0.39
2010_Geom_FT	943	MC	05	1	801	0.58	0.01		0.82	0.07	0.08	0.02			0.82	0.40
2010_Geom_FT	943	MC	06	1	801	0.58	0.01		0.07	0.19	0.28	0.45			0.45	0.49
2010_Geom_FT	943	MC	07	1	801	0.58	0.00		0.06	0.40	0.07	0.46			0.46	0.37
2010_Geom_FT	943	CR	08	2	801	0.58	0.08	0.12	0.05	0.74					1.53	0.51
2010_Geom_FT	943	CR	09	2	801	0.58	0.02	0.50	0.17	0.31					0.79	0.67
2010_Geom_FT	943	CR	10													
2010_Geom_FT	944	MC	01	1	811	0.68	0.01		0.50	0.24	0.16	0.09			0.50	0.52

Table 8. Classical Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	944	MC	02	1	811	0.68	0.00		0.14	0.24	0.57	0.06			0.57	0.36
2010_Geom_FT	944	MC	03	1	811	0.68	0.01		0.06	0.60	0.21	0.12			0.60	0.45
2010_Geom_FT	944	MC	04	1	811	0.68	0.02		0.30	0.18	0.28	0.21			0.21	0.43
2010_Geom_FT	944	MC	05	1	811	0.68	0.00		0.12	0.07	0.60	0.20			0.60	0.47
2010_Geom_FT	944	MC	06	1	811	0.68	0.01		0.38	0.16	0.24	0.21			0.38	0.52
2010_Geom_FT	944	MC	07	1	811	0.68	0.02		0.76	0.14	0.03	0.05			0.76	0.42
2010_Geom_FT	944	CR	08	2	811	0.68	0.07	0.54	0.11	0.28					0.68	0.61
2010_Geom_FT	944	CR	09	4	811	0.68	0.09	0.51	0.07	0.07	0.05	0.21			1.19	0.72
2010_Geom_FT	944	CR	10	4	811	0.68	0.11	0.37	0.20	0.11	0.09	0.12			1.17	0.73
2010_Geom_FT	945	MC	01	1	802	0.48	0.00		0.04	0.06	0.58	0.32			0.58	0.32
2010_Geom_FT	945	MC	02	1	802	0.48	0.00		0.49	0.13	0.10	0.28			0.49	0.37
2010_Geom_FT	945	MC	03	1	802	0.48	0.00		0.30	0.22	0.10	0.38			0.38	0.24
2010_Geom_FT	945	MC	04	1	802	0.48	0.00		0.38	0.13	0.38	0.11			0.38	0.10
2010_Geom_FT	945	MC	05	1	802	0.48	0.01		0.24	0.55	0.10	0.11			0.55	0.29
2010_Geom_FT	945	MC	06	1	802	0.48	0.01		0.51	0.20	0.18	0.09			0.51	0.38
2010_Geom_FT	945	MC	07	1	802	0.48	0.01		0.14	0.09	0.54	0.21			0.54	0.36
2010_Geom_FT	945	CR	08	2	802	0.48	0.09	0.43	0.10	0.38					0.86	0.51
2010_Geom_FT	945	CR	09	6	802	0.48	0.07	0.69	0.06	0.02	0.02	0.02	0.03	0.09	0.93	0.78
2010_Geom_FT	945	CR	10	2	802	0.48	0.07	0.37	0.24	0.31					0.87	0.60
2010_Geom_FT	946	MC	01	1	807	0.63	0.01		0.20	0.48	0.19	0.12			0.48	0.38
2010_Geom_FT	946	MC	02	1	807	0.63	0.00		0.06	0.18	0.14	0.61			0.61	0.43
2010_Geom_FT	946	MC	03	1	807	0.63	0.00		0.09	0.10	0.49	0.32			0.49	0.34

Table 8. Classical Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	946	MC	04	1	807	0.63	0.00		0.10	0.70	0.14	0.05			0.70	0.47
2010_Geom_FT	946	MC	05	1	807	0.63	0.00		0.57	0.19	0.15	0.08			0.57	0.33
2010_Geom_FT	946	MC	06	1	807	0.63	0.01		0.07	0.22	0.12	0.57			0.57	0.43
2010_Geom_FT	946	MC	07	1	807	0.63	0.02		0.43	0.40	0.09	0.05			0.43	0.26
2010_Geom_FT	946	CR	08	2	807	0.63	0.08	0.60	0.12	0.20					0.52	0.60
2010_Geom_FT	946	CR	09	4	807	0.63	0.02	0.30	0.12	0.24	0.10	0.22			1.78	0.76
2010_Geom_FT	946	CR	10	4	807	0.63	0.07	0.51	0.10	0.23	0.03	0.07			0.93	0.67
2010_Geom_FT	947	MC	01	1	793	0.61	0.01		0.08	0.43	0.24	0.24			0.43	0.30
2010_Geom_FT	947	MC	02	1	793	0.61	0.01		0.57	0.22	0.13	0.07			0.57	0.48
2010_Geom_FT	947	MC	03	1	793	0.61	0.00		0.26	0.09	0.13	0.52			0.52	0.54
2010_Geom_FT	947	MC	04	1	793	0.61	0.01		0.05	0.78	0.13	0.03			0.78	0.45
2010_Geom_FT	947	MC	05	1	793	0.61	0.01		0.18	0.11	0.24	0.47			0.47	0.51
2010_Geom_FT	947	MC	06	1	793	0.61	0.02		0.07	0.15	0.39	0.36			0.39	0.40
2010_Geom_FT	947	MC	07	1	793	0.61	0.01		0.39	0.21	0.23	0.16			0.23	0.33
2010_Geom_FT	947	CR	08	2	793	0.61	0.07	0.59	0.10	0.24					0.58	0.63
2010_Geom_FT	947	CR	09													
2010_Geom_FT	947	CR	10	2	793	0.61	0.15	0.33	0.23	0.28					0.80	0.72
2010_Geom_FT	948	MC	01	1	797	0.62	0.01		0.18	0.09	0.55	0.17			0.55	0.49
2010_Geom_FT	948	MC	02	1	797	0.62	0.00		0.34	0.05	0.60	0.01			0.60	0.45
2010_Geom_FT	948	MC	03	1	797	0.62	0.01		0.08	0.75	0.09	0.07			0.75	0.38
2010_Geom_FT	948	MC	04	1	797	0.62	0.00		0.27	0.10	0.48	0.15			0.48	0.35
2010_Geom_FT	948	MC	05	1	797	0.62	0.01		0.21	0.52	0.13	0.12			0.21	0.11

Table 8. Classical Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	948	MC	06	1	797	0.62	0.01		0.57	0.10	0.14	0.19			0.57	0.44
2010_Geom_FT	948	MC	07	1	797	0.62	0.02		0.12	0.20	0.29	0.37			0.37	0.39
2010_Geom_FT	948	CR	08	2	797	0.62	0.08	0.28	0.16	0.48					1.11	0.56
2010_Geom_FT	948	CR	09	4	797	0.62	0.06	0.44	0.10	0.15	0.05	0.20			1.34	0.70
2010_Geom_FT	948	CR	10	4	797	0.62	0.08	0.52	0.12	0.06	0.06	0.17			1.10	0.75
2010_Geom_FT	949	MC	01	1	793	0.59	0.01		0.53	0.12	0.24	0.10			0.53	0.43
2010_Geom_FT	949	MC	02	1	793	0.59	0.01		0.19	0.52	0.16	0.12			0.52	0.37
2010_Geom_FT	949	MC	03	1	793	0.59	0.01		0.19	0.34	0.30	0.17			0.30	0.19
2010_Geom_FT	949	MC	04	1	793	0.59	0.01		0.41	0.48	0.06	0.04			0.48	0.19
2010_Geom_FT	949	MC	05	1	793	0.59	0.01		0.07	0.07	0.34	0.52			0.52	0.59
2010_Geom_FT	949	MC	06	1	793	0.59	0.00		0.25	0.65	0.08	0.02			0.65	0.48
2010_Geom_FT	949	MC	07	1	793	0.59	0.03		0.27	0.13	0.51	0.06			0.51	0.41
2010_Geom_FT	949	CR	08	2	793	0.59	0.07	0.48	0.32	0.14					0.60	0.59
2010_Geom_FT	949	CR	09	2	793	0.59	0.10	0.71	0.08	0.11					0.30	0.56
2010_Geom_FT	949	CR	10	6	793	0.59	0.10	0.53	0.14	0.10	0.08	0.01	0.01	0.02	0.80	0.73
2010_Geom_FT	950	MC	01	1	790	0.66	0.01		0.16	0.22	0.58	0.02			0.58	0.53
2010_Geom_FT	950	MC	02	1	790	0.66	0.00		0.04	0.36	0.06	0.54			0.54	0.43
2010_Geom_FT	950	MC	03	1	790	0.66	0.01		0.23	0.22	0.49	0.05			0.22	0.31
2010_Geom_FT	950	MC	04	1	790	0.66	0.01		0.13	0.62	0.09	0.15			0.62	0.39
2010_Geom_FT	950	MC	05	1	790	0.66	0.00		0.16	0.55	0.13	0.16			0.55	0.53
2010_Geom_FT	950	MC	06	1	790	0.66	0.02		0.29	0.25	0.13	0.32			0.29	0.27
2010_Geom_FT	950	MC	07	1	790	0.66	0.01		0.13	0.06	0.02	0.78			0.78	0.35

Table 8. Classical Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	950	CR	08	2	790	0.66	0.09	0.09	0.39	0.43					1.25	0.56
2010_Geom_FT	950	CR	09	4	790	0.66	0.07	0.50	0.09	0.09	0.10	0.14			1.14	0.79
2010_Geom_FT	950	CR	10	4	790	0.66	0.06	0.58	0.11	0.11	0.02	0.12			0.86	0.70
2010_Geom_FT	951	MC	01	1	789	0.59	0.01		0.22	0.30	0.22	0.25			0.30	0.34
2010_Geom_FT	951	MC	02	1	789	0.59	0.01		0.13	0.15	0.58	0.13			0.58	0.27
2010_Geom_FT	951	MC	03	1	789	0.59	0.01		0.04	0.20	0.71	0.05			0.71	0.32
2010_Geom_FT	951	MC	04	1	789	0.59	0.01		0.10	0.36	0.09	0.44			0.44	0.32
2010_Geom_FT	951	MC	05	1	789	0.59	0.01		0.38	0.12	0.06	0.42			0.38	0.43
2010_Geom_FT	951	MC	06	1	789	0.59	0.01		0.30	0.29	0.17	0.23			0.23	0.46
2010_Geom_FT	951	MC	07	1	789	0.59	0.01		0.61	0.12	0.19	0.07			0.61	0.50
2010_Geom_FT	951	CR	08	2	789	0.59	0.08	0.38	0.27	0.27					0.81	0.66
2010_Geom_FT	951	CR	09	6	789	0.59	0.07	0.68	0.07	0.09	0.01	0.02	0.02	0.04	0.74	0.74
2010_Geom_FT	951	CR	10	2	789	0.59	0.13	0.51	0.13	0.22					0.58	0.51
2010_Geom_FT	952	MC	01	1	796	0.59	0.01		0.31	0.51	0.11	0.06			0.31	0.39
2010_Geom_FT	952	MC	02	1	796	0.59	0.01		0.08	0.08	0.15	0.69			0.69	0.30
2010_Geom_FT	952	MC	03	1	796	0.59	0.01		0.12	0.22	0.38	0.27			0.22	0.42
2010_Geom_FT	952	MC	04	1	796	0.59	0.01		0.54	0.10	0.25	0.10			0.54	0.59
2010_Geom_FT	952	MC	05	1	796	0.59	0.01		0.10	0.64	0.12	0.13			0.64	0.48
2010_Geom_FT	952	MC	06	1	796	0.59	0.01		0.12	0.16	0.53	0.19			0.53	0.40
2010_Geom_FT	952	MC	07	1	796	0.59	0.02		0.23	0.19	0.31	0.25			0.31	0.43
2010_Geom_FT	952	CR	08													
2010_Geom_FT	952	CR	09	4	796	0.59	0.15	0.57	0.05	0.15	0.02	0.06			0.65	0.73

Table 8. Classical Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	Alpha	B	M0	M1	M2	M3	M4	M5	M6	Mean	Point-Biserial
2010_Geom_FT	952	CR	10	2	796	0.59	0.16	0.39	0.14	0.31					0.77	0.58
2010_Geom_FT	953	MC	01	1	796	0.58	0.01		0.18	0.25	0.12	0.44			0.44	0.36
2010_Geom_FT	953	MC	02	1	796	0.58	0.01		0.03	0.80	0.10	0.07			0.80	0.39
2010_Geom_FT	953	MC	03	1	796	0.58	0.01		0.25	0.14	0.47	0.12			0.47	0.37
2010_Geom_FT	953	MC	04	1	796	0.58	0.01		0.38	0.40	0.09	0.11			0.38	0.55
2010_Geom_FT	953	MC	05	1	796	0.58	0.01		0.06	0.85	0.07	0.02			0.85	0.38
2010_Geom_FT	953	MC	06	1	796	0.58	0.01		0.17	0.16	0.05	0.61			0.61	0.21
2010_Geom_FT	953	MC	07	1	796	0.58	0.03		0.43	0.22	0.22	0.10			0.43	0.28
2010_Geom_FT	953	CR	08	2	796	0.58	0.01	0.32	0.31	0.35					1.02	0.62
2010_Geom_FT	953	CR	09	2	796	0.58	0.06	0.18	0.29	0.48					1.24	0.60
2010_Geom_FT	953	CR	10	6	796	0.58	0.17	0.45	0.24	0.08	0.01	0.03	0.01	0.01	0.66	0.68
2010_Geom_FT	954	MC	01	1	800	0.61	0.00		0.08	0.09	0.44	0.40			0.40	0.39
2010_Geom_FT	954	MC	02	1	800	0.61	0.00		0.59	0.15	0.15	0.10			0.59	0.24
2010_Geom_FT	954	MC	03	1	800	0.61	0.00		0.09	0.49	0.26	0.16			0.26	0.35
2010_Geom_FT	954	MC	04	1	800	0.61	0.01		0.16	0.25	0.15	0.43			0.43	0.50
2010_Geom_FT	954	MC	05	1	800	0.61	0.00		0.25	0.20	0.51	0.04			0.51	0.34
2010_Geom_FT	954	MC	06	1	800	0.61	0.00		0.15	0.23	0.41	0.21			0.41	0.44
2010_Geom_FT	954	MC	07	1	800	0.61	0.01		0.14	0.05	0.54	0.26			0.26	0.30
2010_Geom_FT	954	CR	08	2	800	0.61	0.10	0.56	0.06	0.29					0.64	0.56
2010_Geom_FT	954	CR	09	4	800	0.61	0.11	0.41	0.25	0.07	0.05	0.12			1.02	0.72
2010_Geom_FT	954	CR	10	4	800	0.61	0.06	0.34	0.13	0.24	0.10	0.15			1.47	0.72

Appendix B: Partial Credit Model Item Analysis

Table 9. Partial Credit Model Item Analysis

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	931	MC	01	1	795	0.5522							0.96
2010_Geom_FT	931	MC	02	1	795	-1.0238							0.99
2010_Geom_FT	931	MC	03	1	795	0.0747							1.13
2010_Geom_FT	931	MC	04	1	795	0.1941							0.96
2010_Geom_FT	931	MC	05	1	795	1.7867							1.15
2010_Geom_FT	931	MC	06	1	795	0.0325							0.96
2010_Geom_FT	931	MC	07	1	795	0.5873							1.04
2010_Geom_FT	931	CR	08	2	795	2.2398	0.5419	-0.5419					1.03
2010_Geom_FT	931	CR	09	2	795	-0.0856	1.1197	-1.1197					0.87
2010_Geom_FT	931	CR	10	6	795	2.0220	-0.4661	-0.4641	0.3908	0.1708	0.1267	0.2418	0.82
2010_Geom_FT	932	MC	01	1	804	-0.6106							0.92
2010_Geom_FT	932	MC	02	1	804	-0.2306							0.91
2010_Geom_FT	932	MC	03	1	804	0.3827							1.18
2010_Geom_FT	932	MC	04	1	804	-0.5384							0.92
2010_Geom_FT	932	MC	05	1	804	0.3037							0.96
2010_Geom_FT	932	MC	06	1	804	0.7641							1.16
2010_Geom_FT	932	MC	07	1	804	0.0449							1.02
2010_Geom_FT	932	CR	08	4	804	2.4683	0.3192	0.2671	-2.3854	1.7991			0.87
2010_Geom_FT	932	CR	09	2	804	1.1078	-1.1086	1.1086					1.00
2010_Geom_FT	932	CR	10	2	804	0.4960	0.1035	-0.1035					0.97
2010_Geom_FT	933	MC	01	1	800	-0.7037							0.95

Table 9. Partial Credit Model Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	933	MC	02	1	800	-1.8668							1.00
2010_Geom_FT	933	MC	03	1	800	0.8239							1.07
2010_Geom_FT	933	MC	04	1	800	1.7237							1.29
2010_Geom_FT	933	MC	05	1	800	0.0241							1.10
2010_Geom_FT	933	MC	06	1	800	-0.3662							0.89
2010_Geom_FT	933	MC	07	1	800	-0.8489							1.05
2010_Geom_FT	933	CR	08	2	800	0.7265	0.5639	-0.5639					1.08
2010_Geom_FT	933	CR	09	6	800	1.4025	-0.7678	0.9554	-0.6459	0.5299	0.2588	-0.3304	0.64
2010_Geom_FT	933	CR	10	2	800	0.8449	0.1631	-0.1631					0.83
2010_Geom_FT	934	MC	01	1	802	0.7515							0.87
2010_Geom_FT	934	MC	02	1	802	0.2820							1.11
2010_Geom_FT	934	MC	03	1	802	0.1402							1.07
2010_Geom_FT	934	MC	04	1	802	-0.4076							0.93
2010_Geom_FT	934	MC	05	1	802	0.9739							1.08
2010_Geom_FT	934	MC	06	1	802	0.2879							1.22
2010_Geom_FT	934	MC	07	1	802	0.9556							1.04
2010_Geom_FT	934	CR	08	2	802	1.0719	0.6201	-0.6201					0.86
2010_Geom_FT	934	CR	09	4	802	0.8143	1.4099	-1.8172	0.7639	-0.3566			0.96
2010_Geom_FT	934	CR	10	4	802	1.3993	-0.6483	0.1262	1.3534	-0.8313			0.83
2010_Geom_FT	935	MC	01	1	780	0.4549							0.94
2010_Geom_FT	935	MC	02	1	780	0.9542							1.16
2010_Geom_FT	935	MC	03	1	780	0.0695							1.01
2010_Geom_FT	935	MC	04	1	780	1.1447							1.04

Table 9. Partial Credit Model Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	935	MC	05	1	780	-1.1315							0.95
2010_Geom_FT	935	MC	06	1	780	-0.3853							1.10
2010_Geom_FT	935	MC	07	1	780	0.9801							1.03
2010_Geom_FT	935	CR	08	2	780	2.0957	1.1293	-1.1293					1.16
2010_Geom_FT	935	CR	09	2	780	0.7831	0.3871	-0.3871					0.97
2010_Geom_FT	935	CR	10	6	780	1.2387	1.4607	-2.3104	1.1253	-0.4271	1.1517	-1.0003	0.62
2010_Geom_FT	936	MC	01	1	782	-0.3271							0.91
2010_Geom_FT	936	MC	02	1	782	1.1766							0.90
2010_Geom_FT	936	MC	03	1	782	-0.0464							0.99
2010_Geom_FT	936	MC	04	1	782	1.8341							1.16
2010_Geom_FT	936	MC	05	1	782	1.2153							1.25
2010_Geom_FT	936	MC	06	1	782	1.5213							1.11
2010_Geom_FT	936	MC	07	1	782	1.1509							0.92
2010_Geom_FT	936	CR	08	2	782	0.5894	1.1066	-1.1066					1.00
2010_Geom_FT	936	CR	09	4	782	1.9934	0.4877	-0.3031	-1.1097	0.9251			0.87
2010_Geom_FT	936	CR	10	4	782	0.6376	0.2604	0.6497	0.0726	-0.9827			0.79
2010_Geom_FT	937	MC	01	1	769	0.0133							0.85
2010_Geom_FT	937	MC	02	1	769	-0.2486							0.94
2010_Geom_FT	937	MC	03	1	769	-0.0205							0.98
2010_Geom_FT	937	MC	04	1	769	0.2715							1.02
2010_Geom_FT	937	MC	05	1	769	-0.0681							1.06
2010_Geom_FT	937	MC	06	1	769	0.8292							1.08
2010_Geom_FT	937	MC	07	1	769	-0.3056							1.17

Table 9. Partial Credit Model Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	937	CR	08	2	769	0.9700	1.1329	-1.1329					1.03
2010_Geom_FT	937	CR	09	2	769	1.0353	0.0354	-0.0354					0.98
2010_Geom_FT	937	CR	10	6	769	1.7615	1.2039	0.8228	-0.6290	-0.6862	-0.1973	-0.5142	0.84
2010_Geom_FT	938	MC	01	1	785	0.4200							0.91
2010_Geom_FT	938	MC	02	1	785	1.9065							1.03
2010_Geom_FT	938	MC	03	1	785	-0.5464							1.07
2010_Geom_FT	938	MC	04	1	785	1.0645							1.00
2010_Geom_FT	938	MC	05	1	785	0.7795							1.06
2010_Geom_FT	938	MC	06	1	785	-0.1655							1.03
2010_Geom_FT	938	MC	07	1	785	2.4440							1.08
2010_Geom_FT	938	CR	08	2	785	1.3045	0.9739	-0.9739					0.88
2010_Geom_FT	938	CR	09	4	785	0.5465	-0.7264	-0.6980	0.4525	0.9719			1.16
2010_Geom_FT	938	CR	10	4	785	0.7055	-0.1559	0.0004	0.0406	0.1149			0.78
2010_Geom_FT	939	MC	01	1	783	0.6396							1.24
2010_Geom_FT	939	MC	02	1	783	0.4241							1.00
2010_Geom_FT	939	MC	03	1	783	-1.5088							1.03
2010_Geom_FT	939	MC	04	1	783	-0.8304							1.05
2010_Geom_FT	939	MC	05	1	783	-1.3995							0.91
2010_Geom_FT	939	MC	06	1	783	-0.9322							0.86
2010_Geom_FT	939	MC	07	1	783	0.9138							1.06
2010_Geom_FT	939	CR	08	2	783	0.9698	-0.5762	0.5762					0.85
2010_Geom_FT	939	CR	09	2	783	0.8654	0.2176	-0.2176					1.06
2010_Geom_FT	939	CR	10	6	783	1.6882	1.3240	-1.7260	0.1343	0.3491	0.8869	-0.9683	0.73

Table 9. Partial Credit Model Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	940	MC	01	1	800	0.3593							1.08
2010_Geom_FT	940	MC	02	1	800	-0.0440							1.07
2010_Geom_FT	940	MC	03	1	800	-0.4761							0.96
2010_Geom_FT	940	MC	04	1	800	1.2514							1.15
2010_Geom_FT	940	MC	05	1	800	1.0539							1.00
2010_Geom_FT	940	MC	06	1	800	-0.9105							0.90
2010_Geom_FT	940	MC	07	1	800	-0.2537							1.06
2010_Geom_FT	940	CR	08	2	800	0.8197	0.2098	-0.2098					1.01
2010_Geom_FT	940	CR	09	4	800	0.1461	1.0728	-1.5897	1.4920	-0.9752			0.92
2010_Geom_FT	940	CR	10	4	800	1.1199	0.2103	0.8870	-0.4493	-0.6480			0.83
2010_Geom_FT	941	MC	01	1	795	0.7780							1.13
2010_Geom_FT	941	MC	02	1	795	1.0312							1.18
2010_Geom_FT	941	MC	03	1	795	1.8178							0.93
2010_Geom_FT	941	MC	04	1	795	-0.1029							1.03
2010_Geom_FT	941	MC	05	1	795	1.1590							1.01
2010_Geom_FT	941	MC	06	1	795	-0.9760							0.97
2010_Geom_FT	941	MC	07	1	795	0.5377							1.09
2010_Geom_FT	941	CR	08	2	795	1.0428	0.1286	-0.1286					0.91
2010_Geom_FT	941	CR	09	4	795	1.9219	-0.6418	0.0562	-0.2882	0.8738			0.74
2010_Geom_FT	941	CR	10	2	795	0.7495	0.7412	-0.7412					0.96
2010_Geom_FT	942	MC	01	1	811	0.8779							0.96
2010_Geom_FT	942	MC	02	1	811	-0.1518							1.04
2010_Geom_FT	942	MC	03	1	811	1.2671							1.21

Table 9. Partial Credit Model Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	942	MC	04	1	811	0.7183							1.02
2010_Geom_FT	942	MC	05	1	811	-1.0311							1.05
2010_Geom_FT	942	MC	06	1	811	-0.9679							0.94
2010_Geom_FT	942	MC	07	1	811	-0.2232							0.93
2010_Geom_FT	942	CR	08	2	811	-0.5080	0.4745	-0.4745					0.78
2010_Geom_FT	942	CR	09										
2010_Geom_FT	942	CR	10	4	811	2.5072	1.5883	-2.3098	1.7512	-1.0296			0.80
2010_Geom_FT	943	MC	01	1	801	-0.3545							0.93
2010_Geom_FT	943	MC	02	1	801	0.1211							0.90
2010_Geom_FT	943	MC	03	1	801	1.6575							0.99
2010_Geom_FT	943	MC	04	1	801	1.0240							1.08
2010_Geom_FT	943	MC	05	1	801	-1.2289							1.02
2010_Geom_FT	943	MC	06	1	801	0.6816							0.97
2010_Geom_FT	943	MC	07	1	801	0.6127							1.12
2010_Geom_FT	943	CR	08	2	801	-0.5208	1.6875	-1.6875					1.13
2010_Geom_FT	943	CR	09	2	801	0.8306	0.4727	-0.4727					0.88
2010_Geom_FT	943	CR	10										
2010_Geom_FT	944	MC	01	1	811	0.4466							0.93
2010_Geom_FT	944	MC	02	1	811	0.1102							1.13
2010_Geom_FT	944	MC	03	1	811	-0.0408							1.02
2010_Geom_FT	944	MC	04	1	811	2.0058							0.97
2010_Geom_FT	944	MC	05	1	811	-0.0792							0.97
2010_Geom_FT	944	MC	06	1	811	1.0207							0.93

Table 9. Partial Credit Model Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	944	MC	07	1	811	-1.0006							0.97
2010_Geom_FT	944	CR	08	2	811	1.0756	0.9896	-0.9896					1.02
2010_Geom_FT	944	CR	09	4	811	1.1261	1.1423	-0.3988	0.3658	-1.1093			1.08
2010_Geom_FT	944	CR	10	4	811	1.2536	-0.2694	0.0617	-0.0013	0.2090			0.98
2010_Geom_FT	945	MC	01	1	802	0.0634							1.02
2010_Geom_FT	945	MC	02	1	802	0.4775							0.97
2010_Geom_FT	945	MC	03	1	802	0.9880							1.09
2010_Geom_FT	945	MC	04	1	802	0.9822							1.20
2010_Geom_FT	945	MC	05	1	802	0.2163							1.05
2010_Geom_FT	945	MC	06	1	802	0.4000							0.98
2010_Geom_FT	945	MC	07	1	802	0.2499							1.00
2010_Geom_FT	945	CR	08	2	802	0.6850	1.2060	-1.2060					1.00
2010_Geom_FT	945	CR	09	6	802	1.3680	1.4532	0.4131	-0.3005	-0.2898	-0.3464	-0.9297	0.56
2010_Geom_FT	945	CR	10	2	802	0.6878	0.1657	-0.1657					0.86
2010_Geom_FT	946	MC	01	1	807	0.5326							1.04
2010_Geom_FT	946	MC	02	1	807	-0.0837							0.98
2010_Geom_FT	946	MC	03	1	807	0.5097							1.08
2010_Geom_FT	946	MC	04	1	807	-0.5170							0.91
2010_Geom_FT	946	MC	05	1	807	0.1183							1.08
2010_Geom_FT	946	MC	06	1	807	0.1241							0.98
2010_Geom_FT	946	MC	07	1	807	0.7740							1.16
2010_Geom_FT	946	CR	08	2	807	1.3589	0.7751	-0.7751					0.92
2010_Geom_FT	946	CR	09	4	807	0.6512	0.2011	-0.9328	0.9817	-0.2500			0.87

Table 9. Partial Credit Model Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	946	CR	10	4	807	1.4940	0.5180	-1.6374	1.5938	-0.4744			1.00
2010_Geom_FT	947	MC	01	1	793	0.7428							1.21
2010_Geom_FT	947	MC	02	1	793	0.0634							0.99
2010_Geom_FT	947	MC	03	1	793	0.3047							0.92
2010_Geom_FT	947	MC	04	1	793	-1.0942							0.94
2010_Geom_FT	947	MC	05	1	793	0.5550							0.96
2010_Geom_FT	947	MC	06	1	793	0.9042							1.09
2010_Geom_FT	947	MC	07	1	793	1.8032							1.12
2010_Geom_FT	947	CR	08	2	793	1.1879	0.9691	-0.9691					0.99
2010_Geom_FT	947	CR	09										
2010_Geom_FT	947	CR	10	2	793	0.8034	0.0267	-0.0267					0.79
2010_Geom_FT	948	MC	01	1	797	0.2090							0.91
2010_Geom_FT	948	MC	02	1	797	-0.0160							0.95
2010_Geom_FT	948	MC	03	1	797	-0.8380							0.99
2010_Geom_FT	948	MC	04	1	797	0.5275							1.07
2010_Geom_FT	948	MC	05	1	797	1.9172							1.23
2010_Geom_FT	948	MC	06	1	797	0.1328							0.97
2010_Geom_FT	948	MC	07	1	797	1.0773							1.02
2010_Geom_FT	948	CR	08	2	797	0.2591	0.6625	-0.6625					1.02
2010_Geom_FT	948	CR	09	4	797	0.9518	0.7462	-0.7793	1.0808	-1.0476			1.06
2010_Geom_FT	948	CR	10	4	797	1.1213	0.7345	0.2168	-0.0846	-0.8667			0.83
2010_Geom_FT	949	MC	01	1	793	0.2511							1.00
2010_Geom_FT	949	MC	02	1	793	0.3350							1.07

Table 9. Partial Credit Model Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	949	MC	03	1	793	1.4062							1.22
2010_Geom_FT	949	MC	04	1	793	0.4901							1.26
2010_Geom_FT	949	MC	05	1	793	0.3350							0.81
2010_Geom_FT	949	MC	06	1	793	-0.3044							0.89
2010_Geom_FT	949	MC	07	1	793	0.3709							1.04
2010_Geom_FT	949	CR	08	2	793	1.3552	-0.5143	0.5143					0.93
2010_Geom_FT	949	CR	09	2	793	1.9014	0.9801	-0.9801					0.91
2010_Geom_FT	949	CR	10	6	793	1.9635	-0.1857	-0.8150	-0.4645	1.6044	-0.2836	0.1444	0.86
2010_Geom_FT	950	MC	01	1	790	0.0654							0.90
2010_Geom_FT	950	MC	02	1	790	0.2661							1.00
2010_Geom_FT	950	MC	03	1	790	1.9482							1.10
2010_Geom_FT	950	MC	04	1	790	-0.1542							1.09
2010_Geom_FT	950	MC	05	1	790	0.2163							0.91
2010_Geom_FT	950	MC	06	1	790	1.5437							1.17
2010_Geom_FT	950	MC	07	1	790	-1.1010							1.01
2010_Geom_FT	950	CR	08	2	790	-0.1736	-0.8284	0.8284					0.97
2010_Geom_FT	950	CR	09	4	790	1.2760	0.7418	-0.5096	-0.1581	-0.0741			0.79
2010_Geom_FT	950	CR	10	4	790	1.5031	0.5585	-0.5459	1.6306	-1.6432			1.00
2010_Geom_FT	951	MC	01	1	789	1.4358							1.07
2010_Geom_FT	951	MC	02	1	789	0.0683							1.17
2010_Geom_FT	951	MC	03	1	789	-0.5884							1.04
2010_Geom_FT	951	MC	04	1	789	0.7088							1.11
2010_Geom_FT	951	MC	05	1	789	1.0066							0.97

Table 9. Partial Credit Model Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	951	MC	06	1	789	1.8117							0.94
2010_Geom_FT	951	MC	07	1	789	-0.0680							0.88
2010_Geom_FT	951	CR	08	2	789	0.8423	-0.0751	0.0751					0.82
2010_Geom_FT	951	CR	09	6	789	1.8221	0.9942	-1.2041	1.3825	-0.8915	-0.0729	-0.2082	0.77
2010_Geom_FT	951	CR	10	2	789	1.2418	0.6990	-0.6990					1.11
2010_Geom_FT	952	MC	01	1	796	1.2710							1.06
2010_Geom_FT	952	MC	02	1	796	-0.5317							1.14
2010_Geom_FT	952	MC	03	1	796	1.7758							1.00
2010_Geom_FT	952	MC	04	1	796	0.1815							0.82
2010_Geom_FT	952	MC	05	1	796	-0.3104							0.93
2010_Geom_FT	952	MC	06	1	796	0.2393							1.06
2010_Geom_FT	952	MC	07	1	796	1.2710							1.01
2010_Geom_FT	952	CR	08										
2010_Geom_FT	952	CR	09	4	796	1.7128	1.0700	-1.8734	1.8210	-1.0176			0.81
2010_Geom_FT	952	CR	10	2	796	0.8004	0.7133	-0.7133					1.08
2010_Geom_FT	953	MC	01	1	796	0.7504							1.07
2010_Geom_FT	953	MC	02	1	796	-1.1426							0.96
2010_Geom_FT	953	MC	03	1	796	0.5733							1.07
2010_Geom_FT	953	MC	04	1	796	1.0039							0.86
2010_Geom_FT	953	MC	05	1	796	-1.5608							0.97
2010_Geom_FT	953	MC	06	1	796	-0.0752							1.22
2010_Geom_FT	953	MC	07	1	796	0.7981							1.17
2010_Geom_FT	953	CR	08	2	796	0.4127	-0.2717	0.2717					0.91

Table 9. Partial Credit Model Item Analysis (continued)

Test	Form	Type	Item	Max	N-Count	RID	S1	S2	S3	S4	S5	S6	INFIT
2010_Geom_FT	953	CR	09	2	796	-0.0528	-0.2073	0.2073					0.89
2010_Geom_FT	953	CR	10	6	796	2.2689	-1.0048	-0.2122	1.1957	-1.3629	1.3187	0.0655	0.73
2010_Geom_FT	954	MC	01	1	800	0.8877							1.01
2010_Geom_FT	954	MC	02	1	800	0.0097							1.14
2010_Geom_FT	954	MC	03	1	800	1.5862							1.02
2010_Geom_FT	954	MC	04	1	800	0.7185							0.91
2010_Geom_FT	954	MC	05	1	800	0.3806							1.06
2010_Geom_FT	954	MC	06	1	800	0.8231							0.96
2010_Geom_FT	954	MC	07	1	800	1.6006							1.07
2010_Geom_FT	954	CR	08	2	800	1.0130	1.7848	-1.7848					1.02
2010_Geom_FT	954	CR	09	4	800	1.1838	-0.2603	0.8057	0.0815	-0.6269			0.91
2010_Geom_FT	954	CR	10	4	800	0.8940	0.2416	-1.0674	0.8548	-0.0290			0.95

Appendix C: DIF Statistics

Table 10. DIF Statistics

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
931	1	MC	-0.71	3.24	-0.13		
931	2	MC	-0.08	0.03	-0.01		
931	3	MC	-0.13	0.12	-0.02		
931	4	MC	-0.11	0.08	-0.01		
931	5	MC	-0.91	4.85	-0.16		
931	6	MC	0.44	1.34	0.08		
931	7	MC	-0.50	1.82	-0.10		
931	8	OE		0.57	-0.02		
931	9	OE		1.45	0.06		
931	10	OE		5.04	0.10		
932	1	MC	0.47	1.14	0.06		
932	2	MC	0.31	0.53	0.04		
932	3	MC	-0.34	0.96	-0.05		
932	4	MC	-0.02	0.00	-0.01		
932	5	MC	-0.18	0.21	-0.03		
932	6	MC	0.09	0.06	0.01		
932	7	MC	-0.28	0.54	-0.03		
932	8	OE		0.68	-0.04		
932	9	OE		0.04	0.02		
932	10	OE		0.84	0.03		
933	1	MC	-0.03	0.00	-0.01		
933	2	MC	0.69	1.38	0.06		

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
933	3	MC	-0.68	3.25	-0.13		
933	4	MC	-0.64	2.62	-0.10		
933	5	MC	0.19	0.24	0.03		
933	6	MC	-0.68	2.47	-0.09		
933	7	MC	-0.33	0.60	-0.06		
933	8	OE		0.01	0.01		
933	9	OE		0.74	0.03		
933	10	OE		4.04	0.09		
934	1	MC	0.87	4.32	0.12		
934	2	MC	-0.87	5.93	-0.16		
934	3	MC	0.35	0.96	0.04		
934	4	MC	-0.05	0.02	-0.02		
934	5	MC	-0.56	2.31	-0.10		
934	6	MC	-0.55	2.38	-0.11		
934	7	MC	0.07	0.03	0.01		
934	8	OE		0.03	-0.01		
934	9	OE		1.10	-0.04		
934	10	OE		9.57	0.14		
935	1	MC	0.57	1.91	0.08		
935	2	MC	-0.28	0.55	-0.05		
935	3	MC	0.01	0.00	-0.03		
935	4	MC	0.46	1.30	0.07		

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
935	5	MC	0.00	0.00	0.01		
935	6	MC	-0.98	5.83	-0.16		
935	7	MC	0.40	1.01	0.05		
935	8	OE		0.00	-0.02		
935	9	OE		1.60	-0.07		
935	10	OE		0.66	0.04		
936	1	MC	0.23	0.30	0.04		
936	2	MC	-0.70	2.68	-0.11		
936	3	MC	0.01	0.00	-0.01		
936	4	MC	0.64	2.27	0.14		
936	5	MC	-0.34	0.84	-0.05		
936	6	MC	-0.08	0.04	0.00		
936	7	MC	-0.85	4.17	-0.12		
936	8	OE		0.08	-0.02		
936	9	OE		0.11	-0.03		
936	10	OE		2.65	0.07		
937	1	MC	-1.56	12.10	-0.21	C	M
937	2	MC	0.43	1.05	0.08		
937	3	MC	-0.64	2.44	-0.11		
937	4	MC	0.31	0.63	0.06		
937	5	MC	-1.02	6.74	-0.16	B	M
937	6	MC	-0.36	0.91	-0.07		
937	7	MC	0.11	0.08	0.02		

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
937	8	OE		0.03	0.01		
937	9	OE		3.11	0.10		
937	10	OE		2.33	0.05		
938	1	MC	0.52	1.69	0.09		
938	2	MC	-0.02	0.00	-0.01		
938	3	MC	-0.56	1.87	-0.09		
938	4	MC	-0.01	0.00	0.00		
938	5	MC	-0.66	3.06	-0.13		
938	6	MC	-0.34	0.80	-0.05		
938	7	MC	-0.26	0.25	-0.02		
938	8	OE		0.64	-0.05		
938	9	OE		3.77	0.11		
938	10	OE		0.00	-0.01		
939	1	MC	-0.29	0.64	-0.06		
939	2	MC	-0.13	0.12	-0.01		
939	3	MC	-0.82	2.51	-0.11		
939	4	MC	-0.19	0.19	-0.03		
939	5	MC	0.56	1.14	0.07		
939	6	MC	0.19	0.14	0.03		
939	7	MC	0.47	1.56	0.07		
939	8	OE		4.11	0.11		
939	9	OE		0.39	0.03		
939	10	OE		2.35	-0.05		

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
940	1	MC	-0.11	0.10	-0.04		
940	2	MC	-0.93	6.34	-0.17		
940	3	MC	0.19	0.22	0.04		
940	4	MC	-0.09	0.06	-0.01		
940	5	MC	0.41	1.20	0.09		
940	6	MC	-0.01	0.00	0.00		
940	7	MC	1.01	6.90	0.18	B	F
940	8	OE		0.34	0.03		
940	9	OE		3.07	0.08		
940	10	OE		7.47	-0.13		
941	1	MC	0.04	0.01	0.03		
941	2	MC	0.04	0.01	0.01		
941	3	MC	-0.57	1.50	-0.07		
941	4	MC	-0.76	3.96	-0.13		
941	5	MC	-0.40	0.99	-0.05		
941	6	MC	0.04	0.01	0.01		
941	7	MC	-0.80	4.79	-0.14		
941	8	OE		5.44	0.13		
941	9	OE		0.78	-0.04		
941	10	OE		6.21	0.15		
942	1	MC	0.35	0.79	0.05		
942	2	MC	-0.53	1.94	-0.10		
942	3	MC	-0.01	0.00	0.01		

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
942	4	MC	0.52	1.88	0.09		
942	5	MC	0.42	0.87	0.06		
942	6	MC	-0.24	0.28	-0.02		
942	7	MC	-0.27	0.45	-0.04		
942	8	OE		0.56	0.03		
942	9	OE					
942	10	OE		1.92	-0.06		
943	1	MC	0.54	1.57	0.06		
943	2	MC	-0.42	1.05	-0.07		
943	3	MC	0.76	2.95	0.12		
943	4	MC	-0.01	0.00	-0.01		
943	5	MC	0.26	0.31	0.04		
943	6	MC	-0.14	0.13	-0.02		
943	7	MC	0.17	0.21	0.04		
943	8	OE		0.18	-0.03		
943	9	OE		1.24	-0.04		
943	10	OE					
944	1	MC	-0.01	0.00	0.00		
944	2	MC	-1.20	10.24	-0.22	B	M
944	3	MC	0.00	0.00	0.00		
944	4	MC	-0.23	0.24	-0.03		
944	5	MC	0.26	0.44	0.06		
944	6	MC	-0.28	0.48	-0.04		

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
944	7	MC	0.22	0.23	0.02		
944	8	OE		0.59	0.05		
944	9	OE		0.96	0.05		
944	10	OE		0.02	-0.01		
945	1	MC	0.32	0.75	0.07		
945	2	MC	-0.14	0.13	-0.03		
945	3	MC	0.18	0.27	0.03		
945	4	MC	-0.04	0.01	0.00		
945	5	MC	-0.51	1.95	-0.09		
945	6	MC	0.69	3.58	0.12		
945	7	MC	-0.92	6.27	-0.16		
945	8	OE		0.70	0.04		
945	9	OE		0.22	-0.01		
945	10	OE		0.14	0.03		
946	1	MC	-0.88	5.81	-0.16		
946	2	MC	-0.96	6.01	-0.16		
946	3	MC	0.66	3.34	0.12		
946	4	MC	-1.01	5.65	-0.14	B	M
946	5	MC	0.52	2.04	0.09		
946	6	MC	0.89	5.40	0.15		
946	7	MC	0.36	1.03	0.07		
946	8	OE		2.47	-0.09		
946	9	OE		1.97	0.07		

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
946	10	OE		0.12	-0.02		
947	1	MC	0.10	0.07	0.02		
947	2	MC	-0.29	0.56	-0.05		
947	3	MC	0.01	0.00	0.01		
947	4	MC	-0.38	0.61	-0.04		
947	5	MC	1.09	7.43	0.17	B	F
947	6	MC	-0.01	0.00	-0.02		
947	7	MC	-0.66	2.35	-0.10		
947	8	OE		3.98	-0.11		
947	9	OE					
947	10	OE		4.25	0.10		
948	1	MC	-0.46	1.36	-0.08		
948	2	MC	0.05	0.02	0.01		
948	3	MC	0.62	2.10	0.09		
948	4	MC	-0.61	2.93	-0.11		
948	5	MC	-0.82	3.94	-0.15		
948	6	MC	-0.11	0.08	-0.01		
948	7	MC	0.33	0.77	0.06		
948	8	OE		1.41	-0.07		
948	9	OE		2.23	0.07		
948	10	OE		0.19	0.02		
949	1	MC	0.15	0.16	0.04		
949	2	MC	-0.15	0.17	-0.03		

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
949	3	MC	-0.53	1.94	-0.08		
949	4	MC	0.25	0.54	0.06		
949	5	MC	-0.34	0.63	-0.04		
949	6	MC	0.05	0.02	0.01		
949	7	MC	0.37	0.99	0.05		
949	8	OE		0.77	-0.05		
949	9	OE		0.05	-0.01		
949	10	OE		0.32	0.03		
950	1	MC	1.12	7.41	0.17	B	F
950	2	MC	0.72	3.34	0.11		
950	3	MC	-0.90	4.23	-0.15		
950	4	MC	0.20	0.28	0.03		
950	5	MC	-0.99	6.09	-0.14		
950	6	MC	-0.83	4.59	-0.15		
950	7	MC	-0.04	0.01	0.00		
950	8	OE		0.02	0.01		
950	9	OE		1.57	0.05		
950	10	OE		0.32	-0.02		
951	1	MC	-0.26	0.43	-0.04		
951	2	MC	-1.34	13.79	-0.25	B	M
951	3	MC	0.36	0.77	0.05		
951	4	MC	-0.05	0.02	-0.02		
951	5	MC	-0.86	4.78	-0.15		

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
951	6	MC	0.39	0.76	0.07		
951	7	MC	-0.49	1.40	-0.07		
951	8	OE		3.36	0.10		
951	9	OE		4.99	0.09		
951	10	OE		0.05	-0.02		
952	1	MC	-0.27	0.50	-0.05		
952	2	MC	0.25	0.40	0.05		
952	3	MC	0.06	0.02	0.01		
952	4	MC	0.04	0.01	0.00		
952	5	MC	-0.29	0.50	-0.05		
952	6	MC	-0.20	0.28	-0.03		
952	7	MC	-0.66	2.76	-0.10		
952	8	OE					
952	9	OE		7.12	0.11		
952	10	OE		1.04	-0.05		
953	1	MC	0.15	0.17	0.03		
953	2	MC	0.13	0.08	0.02		
953	3	MC	-1.42	14.71	-0.26	B	M
953	4	MC	-1.67	15.31	-0.23	C	M
953	5	MC	-0.06	0.01	-0.01		
953	6	MC	1.13	9.93	0.23	B	F
953	7	MC	0.02	0.00	0.02		
953	8	OE		0.78	0.06		

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
953	9	OE		1.68	-0.07		
953	10	OE		7.45	0.13		
954	1	MC	0.09	0.06	0.03		
954	2	MC	1.46	16.19	0.28	B	F
954	3	MC	-0.64	2.32	-0.10		
954	4	MC	-0.13	0.11	-0.02		

*DIF Category meanings: A/AA=negligible, B/BB=moderate, C/CC=large

Table 10. DIF Statistics (continued)

Form	Item	Item Type	MH Delta	MH Chi-Sq	Effect Size	DIF Category	Favored Group
954	5	MC	-0.44	1.50	-0.09		
954	6	MC	-0.75	3.92	-0.13		
954	7	MC	-0.16	0.17	-0.04		
954	8	OE		4.02	-0.13		
954	9	OE		1.08	0.05		
954	10	OE		0.80	0.05		

Appendix D: Operational Test Maps

Table 11. Operational Test Map for January 2010

Position	Item Type	Max Points	Weight	Strand	Mean	PBS	Rasch	S1	S2	S3	S4	S5	S6
1	MC	1	2	Informal and Formal Proofs	0.85	0.32	-1.82						
2	MC	1	2	Informal and Formal Proofs	0.85	0.49	-1.79						
3	MC	1	2	Transformational Geometry	0.52	0.24	0.09						
4	MC	1	2	Constructions	0.82	0.39	-1.57						
5	MC	1	2	Informal and Formal Proofs	0.77	0.47	-1.22						
6	MC	1	2	Transformational Geometry	0.68	0.32	-0.64						
7	MC	1	2	Informal and Formal Proofs	0.75	0.41	-1.10						
8	MC	1	2	Informal and Formal Proofs	0.67	0.41	-0.56						
9	MC	1	2	Constructions	0.64	0.45	-0.50						
10	MC	1	2	Coordinate Geometry	0.61	0.43	-0.28						
11	MC	1	2	Locus	0.60	0.16	-0.24						
12	MC	1	2	Geometric Relationships	0.56	0.25	-0.01						
13	MC	1	2	Informal and Formal Proofs	0.50	0.47	0.16						
14	MC	1	2	Coordinate Geometry	0.49	0.43	0.19						

Table 11. Operational Test Map for January 2010 (continued)

Position	Item Type	Max Points	Weight	Strand	Mean	PBS	Rasch	S1	S2	S3	S4	S5	S6
15	MC	1	2	Informal and Formal Proofs	0.49	0.49	0.20						
16	MC	1	2	Informal and Formal Proofs	0.49	0.49	0.36						
17	MC	1	2	Coordinate Geometry	0.44	0.51	0.46						
18	MC	1	2	Coordinate Geometry	0.58	0.43	-0.17						
19	MC	1	2	Informal and Formal Proofs	0.37	0.48	0.78						
20	MC	1	2	Coordinate Geometry	0.38	0.46	0.83						
21	MC	1	2	Informal and Formal Proofs	0.34	0.38	0.95						
22	MC	1	2	Informal and Formal Proofs	0.33	0.22	0.99						
23	MC	1	2	Informal and Formal Proofs	0.34	0.33	1.01						
24	MC	1	2	Geometric Relationships	0.32	0.28	1.05						
25	MC	1	2	Coordinate Geometry	0.33	0.32	1.10						
26	MC	1	2	Transformational Geometry	0.33	0.40	1.09						
27	MC	1	2	Geometric Relationships	0.27	0.39	1.31						
28	MC	1	2	Informal and Formal Proofs	0.61	0.23	-0.35						
29	CR	2	1	Informal and Formal Proofs	1.47	0.62	-0.65	0.69	-0.69				
30	CR	2	1	Geometric Relationships	1.00	0.67	0.21	0.40	-0.40				

Table 11. Operational Test Map for January 2010 (continued)

Position	Item Type	Max Points	Weight	Strand	Mean	PBS	Rasch	S1	S2	S3	S4	S5	S6
31	CR	2	1	Coordinate Geometry	0.99	0.66	0.23	-0.09	0.09				
32	CR	2	1	Constructions	0.70	0.53	0.70	2.74	-2.74				
33	CR	2	1	Informal and Formal Proofs	0.65	0.62	0.90	0.64	-0.64				
34	CR	2	1	Informal and Formal Proofs	0.39	0.34	2.14	-1.26	1.26				
35	CR	4	1	Transformational Geometry	1.55	0.71	0.58	1.86	-2.29	1.54	-1.10		
36	CR	4	1	Informal and Formal Proofs	0.67	0.58	1.42	0.17	0.06	0.00	-0.23		
37	CR	4	1	Locus	0.56	0.54	1.49	0.05	0.77	0.85	-1.67		
38	CR	6	1	Coordinate Geometry	2.25	0.75	0.59	0.22	-0.92	2.81	-2.84	1.97	-1.24

Table 12. Operational Test Map for June 2010

Position	Item Type	Max Points	Weight	Strand	Mean	PBS	Rasch	S1	S2	S3	S4	S5	S6
1	MC	1	2	Informal and Formal Proofs	0.89	0.35	-2.05						
2	MC	1	2	Informal and Formal Proofs	0.87	0.34	-1.98						
3	MC	1	2	Geometric Relationships	0.81	0.44	-1.35						
4	MC	1	2	Informal and Formal Proofs	0.82	0.44	-1.46						
5	MC	1	2	Transformational Geometry	0.77	0.5	-1.29						
6	MC	1	2	Geometric Relationships	0.68	0.42	-0.55						
7	MC	1	2	Informal and Formal Proofs	0.66	0.49	-0.5						
8	MC	1	2	Informal and Formal Proofs	0.63	0.41	-0.33						
9	MC	1	2	Informal and Formal Proofs	0.62	0.39	-0.29						
10	MC	1	2	Informal and Formal Proofs	0.58	0.49	-0.07						
11	MC	1	2	Coordinate Geometry	0.54	0.33	0.08						
12	MC	1	2	Constructions	0.53	0.41	0.22						
13	MC	1	2	Informal and Formal Proofs	0.51	0.46	0.15						
14	MC	1	2	Coordinate Geometry	0.51	0.47	0.26						

Table 12. Operational Test Map for June 2010 (continued)

Position	Item Type	Max Points	Weight	Strand	Mean	PBS	Rasch	S1	S2	S3	S4	S5	S6
15	MC	1	2	Transformational Geometry	0.5	0.35	0.34						
16	MC	1	2	Informal and Formal Proofs	0.5	0.4	0.17						
17	MC	1	2	Geometric Relationships	0.48	0.31	0.35						
18	MC	1	2	Transformational Geometry	0.47	0.46	0.34						
19	MC	1	2	Informal and Formal Proofs	0.47	0.43	0.34						
20	MC	1	2	Constructions	0.46	0.29	0.49						
21	MC	1	2	Coordinate Geometry	0.41	0.56	0.59						
22	MC	1	2	Coordinate Geometry	0.41	0.48	0.64						
23	MC	1	2	Informal and Formal Proofs	0.39	0.19	0.79						
24	MC	1	2	Informal and Formal Proofs	0.38	0.29	0.84						
25	MC	1	2	Locus	0.37	0.27	0.94						
26	MC	1	2	Informal and Formal Proofs	0.31	0.26	1.16						
27	MC	1	2	Coordinate Geometry	0.31	0.4	1.19						
28	MC	1	2	Coordinate Geometry	0.29	0.35	1.27						
29	CR	2	1	Geometric Relationships	1.33	0.62	-0.49	-0.14	0.14				
30	CR	2	1	Informal and Formal Proofs	1.18	0.62	0.04	1.16	-1.16				

Table 12. Operational Test Map for June 2010 (continued)

Position	Item Type	Max Points	Weight	Strand	Mean	PBS	Rasch	S1	S2	S3	S4	S5	S6
31	CR	2	1	Informal and Formal Proofs	1.01	0.7	0.15	0.26	-0.26				
32	CR	2	1	Transformational Geometry	0.89	0.66	0.53	-0.24	0.24				
33	CR	2	1	Locus	0.77	0.67	0.73	0.05	-0.05				
34	CR	2	1	Geometric Relationships	1.13	0.66	0.04	0.06	-0.06				
35	CR	4	1	Informal and Formal Proofs	1.95	0.74	0.29	-0.76	0.98	0.25	-0.47		
36	CR	4	1	Coordinate Geometry	1.09	0.71	1.23	-0.03	-0.32	2.13	-1.77		
37	CR	4	1	Coordinate Geometry	0.81	0.63	1.29	0.23	0.67	0.40	-1.30		
38	CR	6	1	Informal and Formal Proofs	1.96	0.8	0.98	0.90	-2.14	2.74	-2.51	1.74	-0.73

Table 13. Operational Test Map for August 2010

Position	Item Type	Max Points	Weight	Strand	Mean	PBS	Rasch	S1	S2	S3	S4	S5	S6
1	MC	1	2	Informal and Formal Proofs	0.9	0.47	-2.43						
2	MC	1	2	Geometric Relationships	0.78	0.44	-1.21						
3	MC	1	2	Informal and Formal Proofs	0.77	0.42	-1.28						
4	MC	1	2	Coordinate Geometry	0.6	0.46	-0.36						
5	MC	1	2	Constructions	0.52	0.41	0.20						
6	MC	1	2	Informal and Formal Proofs	0.79	0.35	-1.29						
7	MC	1	2	Informal and Formal Proofs	0.74	0.45	-0.96						
8	MC	1	2	Geometric Relationships	0.56	0.46	-0.08						
9	MC	1	2	Coordinate Geometry	0.63	0.5	-0.38						
10	MC	1	2	Coordinate Geometry	0.54	0.5	0.03						
11	MC	1	2	Informal and Formal Proofs	0.55	0.5	-0.02						
12	MC	1	2	Informal and Formal Proofs	0.53	0.34	0.17						
13	MC	1	2	Coordinate Geometry	0.48	0.55	0.45						
14	MC	1	2	Coordinate Geometry	0.52	0.48	0.21						
15	MC	1	2	Transformational Geometry	0.49	0.39	0.26						
16	MC	1	2	Informal and Formal Proofs	0.52	0.25	0.04						

Table 13. Operational Test Map for August 2010 (continued)

Position	Item Type	Max Points	Weight	Strand	Mean	PBS	Rasch	S1	S2	S3	S4	S5	S6
17	MC	1	2	Informal and Formal Proofs	0.47	0.39	0.44						
18	MC	1	2	Informal and Formal Proofs	0.46	0.37	0.52						
19	MC	1	2	Coordinate Geometry	0.45	0.54	0.55						
20	MC	1	2	Geometric Relationships	0.44	0.43	0.51						
21	MC	1	2	Transformational Geometry	0.43	0.44	0.62						
22	MC	1	2	Informal and Formal Proofs	0.42	0.38	0.58						
23	MC	1	2	Transformational Geometry	0.4	0.43	0.72						
24	MC	1	2	Coordinate Geometry	0.43	0.47	0.65						
25	MC	1	2	Informal and Formal Proofs	0.45	0.37	0.62						
26	MC	1	2	Informal and Formal Proofs	0.38	0.42	0.91						
27	MC	1	2	Informal and Formal Proofs	0.59	0.32	-0.12						
28	MC	1	2	Locus	0.54	0.37	0.04						
29	CR	2	1	Informal and Formal Proofs	1.24	0.62	-0.13	0.68	-0.68				
30	CR	2	1	Geometric Relationships	1.23	0.63	-0.23	0.13	-0.13				
31	CR	2	1	Informal and Formal Proofs	0.79	0.56	0.66	0.18	-0.18				
32	CR	2	1	Constructions	0.92	0.6	0.39	2.64	-2.64				

Table 13. Operational Test Map for August 2010 (continued)

Position	Item Type	Max Points	Weight	Strand	Mean	PBS	Rasch	S1	S2	S3	S4	S5	S6
33	CR	2	1	Locus	0.76	0.6	0.76	0.38	-0.38				
34	CR	2	1	Coordinate Geometry	0.54	0.51	1.23	0.27	-0.27				
35	CR	4	1	Informal and Formal Proofs	1.22	0.66	1.10	0.14	-1.27	0.73	0.40		
36	CR	4	1	Transformational Geometry	1.45	0.66	0.64	0.19	-0.16	0.25	-0.28		
37	CR	4	1	Informal and Formal Proofs	0.86	0.59	1.19	1.38	-1.19	2.78	-2.97		
38	CR	6	1	Coordinate Geometry	0.66	0.64	2.09	0.44	-1.43	-0.40	0.60	-0.15	0.94

Appendix E: Scoring Tables

Table 14. Scoring Tables for January 2010

Raw Score	Ability	Scale Score		Raw Score	Ability	Scale Score		Raw Score	Ability	Scale Score		Raw Score	Ability	Scale Score
0	-5.1900	0.000		22	-0.7600	45.244		44	0.4610	68.033		66	1.5190	82.241
1	-4.4640	2.187		23	-0.6890	46.712		45	0.5060	68.779		67	1.5760	82.876
2	-3.7380	4.958		24	-0.6210	48.144		46	0.5510	69.486		68	1.6360	83.569
3	-3.3000	7.300		25	-0.5540	49.486		47	0.5960	70.167		69	1.6980	84.180
4	-2.9800	9.629		26	-0.4890	50.734		48	0.6410	70.871		70	1.7630	84.872
5	-2.7250	11.989		27	-0.4270	51.960		49	0.6860	71.507		71	1.8300	85.550
6	-2.5120	14.341		28	-0.3650	53.179		50	0.7310	72.200		72	1.9020	86.242
7	-2.3270	16.650		29	-0.3050	54.325		51	0.7760	72.877		73	1.9790	86.982
8	-2.1630	18.893		30	-0.2470	55.460		52	0.8210	73.495		74	2.0610	87.737
9	-2.0150	21.116		31	-0.1900	56.563		53	0.8670	74.123		75	2.1510	88.475
10	-1.8790	23.333		32	-0.1340	57.609		54	0.9130	74.750		76	2.2500	89.292
11	-1.7550	25.495		33	-0.0790	58.658		55	0.9590	75.357		77	2.3600	90.131
12	-1.6390	27.622		34	-0.0260	59.629		56	1.0060	76.048		78	2.4840	91.077
13	-1.5300	29.648		35	0.0270	60.588		57	1.0540	76.685		79	2.6270	92.046
14	-1.4280	31.626		36	0.0780	61.524		58	1.1020	77.288		80	2.7930	93.035
15	-1.3310	33.560		37	0.1290	62.413		59	1.1510	77.880		81	2.9910	94.139
16	-1.2390	35.368		38	0.1780	63.247		60	1.2010	78.549		82	3.2340	95.167
17	-1.1510	37.154		39	0.2270	64.096		61	1.2510	79.142		83	3.5480	96.403
18	-1.0660	38.951		40	0.2750	64.947		62	1.3020	79.771		84	3.9860	97.524
19	-0.9850	40.658		41	0.3220	65.782		63	1.3550	80.375		85	4.7210	98.857
20	-0.9080	42.200		42	0.3690	66.548		64	1.4080	80.980		86	5.4560	100.092
21	-0.8320	43.781		43	0.4150	67.281		65	1.4630	81.616				

Table 15. Scoring Tables for June 2010

Raw Score	Ability	Scale Score		Raw Score	Ability	Scale Score		Raw Score	Ability	Scale Score		Raw Score	Ability	Scale Score
0	-5.2690	0.000		22	-0.8030	44.370		44	0.4410	67.705		66	1.4530	81.503
1	-4.5400	1.889		23	-0.7320	45.821		45	0.4880	68.483		67	1.5070	82.109
2	-3.8110	4.682		24	-0.6640	47.242		46	0.5350	69.251		68	1.5620	82.712
3	-3.3700	6.873		25	-0.5970	48.631		47	0.5810	69.917		69	1.6200	83.394
4	-3.0470	9.087		26	-0.5330	49.886		48	0.6270	70.671		70	1.6810	84.013
5	-2.7890	11.330		27	-0.4690	51.128		49	0.6730	71.326		71	1.7450	84.680
6	-2.5730	13.620		28	-0.4080	52.340		50	0.7180	71.983		72	1.8130	85.380
7	-2.3850	15.883		29	-0.3480	53.502		51	0.7630	72.695		73	1.8870	86.101
8	-2.2190	18.098		30	-0.2890	54.633		52	0.8080	73.318		74	1.9660	86.854
9	-2.0690	20.255		31	-0.2310	55.780		53	0.8520	73.918		75	2.0520	87.663
10	-1.9320	22.450		32	-0.1740	56.859		54	0.8960	74.518		76	2.1470	88.442
11	-1.8060	24.572		33	-0.1190	57.892		55	0.9400	75.109		77	2.2520	89.307
12	-1.6880	26.715		34	-0.0640	58.946		56	0.9850	75.728		78	2.3700	90.207
13	-1.5790	28.735		35	-0.0100	59.912		57	1.0290	76.366		79	2.5050	91.222
14	-1.4750	30.696		36	0.0430	60.882		58	1.0730	76.925		80	2.6610	92.254
15	-1.3770	32.640		37	0.0950	61.834		59	1.1180	77.484		81	2.8470	93.339
16	-1.2840	34.489		38	0.1460	62.702		60	1.1640	78.041		82	3.0750	94.504
17	-1.1960	36.233		39	0.1970	63.574		61	1.2090	78.659		83	3.3700	95.704
18	-1.1110	37.995		40	0.2470	64.449		62	1.2560	79.198		84	3.7830	97.009
19	-1.0300	39.711		41	0.2960	65.320		63	1.3030	79.784		85	4.4860	98.434
20	-0.9510	41.340		42	0.3450	66.158		64	1.3520	80.343		86	5.1890	99.644
21	-0.8760	42.867		43	0.3930	66.930		65	1.4020	80.908				

Table 16. Scoring Tables for August 2010

Raw Score	Ability	Scale Score		Raw Score	Ability	Scale Score		Raw Score	Ability	Scale Score		Raw Score	Ability	Scale Score
0	-5.186	0.000		22	-0.725	45.965		44	0.456	67.950		66	1.416	81.075
1	-4.451	2.238		23	-0.657	47.391		45	0.500	68.681		67	1.469	81.684
2	-3.716	5.042		24	-0.591	48.753		46	0.544	69.386		68	1.524	82.295
3	-3.270	7.484		25	-0.527	50.000		47	0.587	70.017		69	1.582	82.947
4	-2.944	9.920		26	-0.465	51.207		48	0.630	70.714		70	1.643	83.638
5	-2.684	12.411		27	-0.404	52.420		49	0.673	71.326		71	1.707	84.275
6	-2.466	14.885		28	-0.345	53.559		50	0.715	71.933		72	1.776	85.010
7	-2.279	17.284		29	-0.288	54.653		51	0.757	72.612		73	1.850	85.750
8	-2.113	19.603		30	-0.232	55.760		52	0.799	73.195		74	1.931	86.515
9	-1.963	21.945		31	-0.177	56.804		53	0.840	73.755		75	2.018	87.363
10	-1.828	24.183		32	-0.123	57.817		54	0.882	74.327		76	2.113	88.161
11	-1.703	26.437		33	-0.070	58.831		55	0.923	74.886		77	2.219	89.042
12	-1.587	28.586		34	-0.018	59.771		56	0.965	75.435		78	2.338	89.963
13	-1.479	30.617		35	0.033	60.698		57	1.007	76.063		79	2.472	90.985
14	-1.378	32.620		36	0.083	61.616		58	1.049	76.621		80	2.627	92.046
15	-1.282	34.528		37	0.132	62.464		59	1.092	77.163		81	2.810	93.133
16	-1.192	36.314		38	0.180	63.281		60	1.135	77.688		82	3.034	94.333
17	-1.105	38.123		39	0.228	64.113		61	1.179	78.247		83	3.323	95.519
18	-1.023	39.858		40	0.275	64.947		62	1.224	78.836		84	3.729	96.870
19	-0.944	41.480		41	0.321	65.764		63	1.270	79.364		85	4.422	98.318
20	-0.868	43.033		42	0.367	66.516		64	1.317	79.962		86	5.115	99.519
21	-0.795	44.532		43	0.412	67.233		65	1.366	80.493				