1. OVERVIEW OF THE 2010 *MARYLAND SCHOOL ASSESSMENT-MATHEMATICS*

In 2002, the Maryland State Department of Education (MSDE), in order to conform to the requirements of the new Federal program "No Child Left Behind," retired its award-winning *Maryland School Performance Assessment Program* and adopted a testing program known as the *Maryland School Assessment* (*MSA*). The new program, like its predecessor, was based on the *Voluntary State Curriculum*, which set reasonable academic standards for what teachers were expected to teach and for what students were expected to learn in schools.

In 2003, the MSA-Math was introduced in grades 3, 5, and 8, with grades 4, 6, and 7 being added to the program in 2004. A Bookmark standard setting was conducted in 2003 to set proficiency level cut scores for grades 3, 5, and 8. Because 2004 was the first testing year for grades 4, 6, and 7, a second Bookmark standard setting was held in summer 2004 to set cut scores for these additional grades. The performance level cut scores were used to assign students to three proficiency levels (Basic, Proficient, and Advanced) for AYP reporting under the "No Child Left Behind" act. Information about the Bookmark procedures and results can be obtained from MSDE. It should be noted that these cut scores have been applied since 2003 (grades 3, 5, and 8) and 2004 (grades 4, 6, and 7).

It should be noted that in 2007, the MSA-Math was administered using a new vendor and applying a different IRT method (e.g., the Rasch model); therefore, a transformation of scale scores using the equipercentile method was conducted with the 2006 population data. Detailed information on the scale score transformation and its results can be found in Appendix C, *Year 2006 MSA-Math Recalibration Results from 3-PL IRT to the Rasch Model Using the Equipercentile Method* in the 2007 *MSA-Math Technical Report*.

In 2007, MSDE implemented an important action plan on the MSA-Math test: dropping all of the SAT10 items from the 2008 assessment. Consequently, several SAT10 items which contributed to the 2007 criterion-referenced test (CRT) were replaced by Maryland-specific items in 2008.

For the purposes of year-to-year linking and equating, we first constructed in 2010 a linking pool: only operational selected-response (SR) items (i.e., multiple-choice items) were included in the linking pool. It should be noted that these SR items appeared both in current and previous years' assessments and were used as either core or core link items in previous years' assessments (i.e., in any assessment before 2010). After setting up the linking pool, we conducted a stability check of linking items and decided which items should be excluded from or which items remain in the linking pool. During the calibration and equating processes, we kept and fixed the original operational Rasch item difficulty parameters of any linking items that remained through the stability check to put the 2010 assessment on a common scale. Accordingly, scale scores of the 2010 assessment were linked back to the 2006 assessment and all the scale scores of different years were comparable within each content and grade.

1.1 Purposes/Uses of the 2010 MSA-Math

By measuring students' achievement against the new academic standards, the 2010 MSA-Math fulfills two main purposes. First, the MSA-Math was designed to inform parents, teachers, and educators of what students actually learned in schools by providing specific feedback that can be

used to improve the quality of schools, classrooms, and individualized instructional programs, and to model effective assessment approaches that can be used in classrooms. Second, the MSA-Math serves as an accountability tool to measure performance levels of individual students, schools, and districts against the new academic standards.

1.2 The Voluntary State Curriculum

Federal law requires that states align their tests with their state content standards. MSDE worked carefully and rigorously to construct new tests to provide a strong alignment as defined by the U.S. Department of Education.

The *Voluntary State Curriculum* (*VSC*), which defined what students should know and be able to do at each grade level, helped schools understand the standards more clearly, and included more specificity with indicators and objectives. The format of the *VSC* specified standards statements, indicators, and objectives. Standards are broad, measurable statements of what students should know and be able to do. Indicators and objectives provide more specific content knowledge and skills that are unique at each grade level.

The objectives assessed by the MSA at each grade level are embedded in the *VSC*. In addition, they are identified with the notation, ***assessment limit***. Assessment limits provide clarification about the specific skills and content that students are expected to have learned for each assessed objective. Even though some objectives in the VSC may not have an Assessment limit at a given grade-level, these non-assessed objectives still must be included in instruction. They introduce important concepts in preparation for assessed skills and content at subsequent grade levels.

The following provides one example of assessment limit of Grade 3 MSA-Math:

STANDARD 1.0

Knowledge of Algebra, Patterns, and Functions

TOPIC:

A. PATTERNS AND FUNCTIONS

INDICATOR:

1. Identify, describe, extend, and create numeric patterns and functions

OBJECTIVES:

a. Represent and analyze numeric patterns using skip counting

Assessment limits:

Use 2, 5, 10, or 100 starting with any whole number (0 – 1000)

It should be noted that it was not the case that every indicator would necessarily be tested each year even if 100% of the standards should be tested. Consequently, the *VSC* specified curricular indicators and objectives that contributed directly to measuring content standards, which were aligned to the MSA. More information on assessment limits and standards can be found in Appendix D, *The 2010 MSA-Math Blueprint*.

1.3 Development and Review of the 2010 MSA-Math Items and Test

As seen in Table 1.1, the development of the 2010 MSA-Math test required the involvement of four groups in addition to MSDE and Pearson. These groups are as follows:

National Psychometric Council

The National Psychometric Council (NPC) took a major role in reviewing and making recommendations to MSDE on the development and implementation of the 2010 MSA-Math program. For example, they made recommendations to MSDE on issues, such as test blueprints, operational form construction, field test design, item analysis, item selection for scoring purposes, linking, equating and scaling issues, and other relevant statistical and psychometric issues.

Content Review Committee

Content Review Committee members ensured that the MSA-Math was appropriately difficult and fair. Committee members were either specialists in math for test items, or experts in test construction and measurement. They represented all levels of education as well as the ethnic and social diversity of Maryland students. Committee members were from different areas of the state.

The educators' understanding of Maryland curriculum and extensive classroom experience made them a valuable source of information. They reviewed test items and forms and took a holistic approach to ensure that tests were fair and balanced across reporting categories.

Bias Review Committee

In addition to the Content Review Committee, a separate Bias Review Committee examined each item on math tests. They looked for indications of bias that would impact the performance of an identifiable group of students. Committee members discussed and, if necessary, rejected items based on gender, ethnic, religious, or geographical bias.

Vision Review Committee

A Vision Review Committee reviewed the items and any associated art for bias to the visually impaired. The committee makes their recommendations to NOT put any item they had a concern with on Form A since this form is usually used for large print and braille forms.

Table 1.1 The 2010 MSA-Math Responsibility for Test Development

Development of the 2010 MSA-Math	Primary Responsibility
Development of Preliminary Blueprints and Item Specifications	Pearson; MSDE; NPC
Development of Operational Form Requirement and Session Blueprint	MSDE
Item Writing	MSDE; Pearson
Item Review	Pearson; MSDE; NPC; Content Review Committee
Bias Review	Pearson; MSDE; Bias Review Committee
Vision Review	Pearson: MSDE; Vision Review Committee
Modification of Special Forms	Pearson; MSDE
Review of Special Forms	MSDE
Construction of Operational Test Forms	Pearson; MSDE; NPC
Construction of Field Test Forms	Pearson; MSDE
Review of Operational Test Forms	MSDE
Final Construction of Test Forms	Pearson; MSDE

1.4 Test Form Design, Specifications, Item Types, and Item Roles

Test Form Design

The MSA-Math test had two forms of operational items at each grade. Field test items were embedded within the operational items resulting in a total of 10 test forms at each grade. As can be seen from Table 1.2, Forms A, B, C, D and E are identical with respect to operational items (designated as operational Form A) and differ only with respect to field test items. This is also true for Forms F, G, H, J, and K (designated as operational Form F).

Table 1.2 The 2010 MSA-Math Test Form Design: Grades 3 through 8

	Operational Item Sets		Field test Item Sets									
	A	F	A	B	C	D	E	F	G	H	J	K
Form A	X		X									
Form B	X			X								
Form C	X				X							
Form D	X					X						
Form E	X						X					
Form F		X						X				
Form G		X							X			
Form H		X								X		
Form J		X									X	
Form K		X										X

Note. Forms A, B, C, D, and E (Form A) are identical, and Forms F, G, H, J, and K (Form F) are identical in terms of operational test items.

Test Form Specifications and Reporting Category

Tables 1.3, 1.4, and 1.5 provide information on the total number of operational items included in each operational test form and how these items were broken down based on each content standard. It should be noted that the test specifications in these tables represent the targeted test design for each grade and show the targeted distribution of each content standard.

Mathematics has a total of seven content standards (Algebra, Geometry, Measurement, Statistics, Probability, Numbers and Computation, and Process). It should be acknowledged that some standards were combined for purposes of reporting subscale. Specifically, Geometry and Measurement standards and Statistics and Probability standards were combined to produce a total of five subscale reporting categories. Tables 1.6 through 1.23 provide information on the actual distribution of score points by standard and reporting category. The number of items and score points for each reporting standard were identical across forms within each grade.

Item Types

The 2010 MSA-Math included four types of items: *selected response (SR), student-produced response (SPR), brief constructed response (BCR), and extended constructed response (ECR)*.

SR items require students to select a correct answer from several alternatives. For the 2010 MSA-Math, students selected an answer from four options. Each *SR* item was scored dichotomously (i.e., 0 or 1).

SPR items require students to record their answers on a grid by shading in circles corresponding to the numbers in their answer. For the 2010 MSA-Math, only grade 7 and 8 tests included *SPR* items. Each *SPR* item was scored dichotomously.

BCR items require students to provide a short answer using words, numbers, and/or symbols, while *ECR* items require students to write an answer that consists of more information than is required for a brief constructed response item.

Both *BCR* and *ECR* items consist of Step A and Step B. Step A contributes to the content score while Step B contributes to the process score. Each step was considered as an independent item and separately scored;

All *BCR* and *ECR* Step A items received a 0-1 score point range from two independent scorers; all *BCR* Step B items received a 0-2 score point range; all *ECR* Step B items received a 0-3 score point range from two independent scorers. The score given was the higher of the first and the second Reader's scores, provided they were adjacent. A resolution reader's score was used if two non-adjacent initial scores were received. That is, the resolution reader's score was used in place of both the first and second Reader's scores. It should be noted that grade 3 and 4 tests did not include *ECR* items.

The Role of Operational SR Item

Most SR items were used for both form-to-form and year-to-year calibration and linking. As a result, operational SR items fell into one of the following four categories: unique core, common core, unique core linking, and common core linking items. First of all, it should be noted that form-to-form linking was conduced with both the common core and the common core linking items. Form-to-form calibration and linking procedures can be found in section of chapter 1.9, *Form-to-Form Linking Procedures*. More importantly, however, year-to-year linking was conduced with only the core linking items and year-to-year calibration and linking procedures can be found in section of chapter 1.9, *Year-to-Year Linking Procedures*.

While unique core items appeared on either operational form A or F, common core items appeared on both forms. As a result, only the common core items were used for form-to-form linking. Because the core items were not included into the possible 2010 linking pool, on the other hand, item parameters of these items were recalibrated with the 2010 live, operational data (i.e., stratified random sample) and then reserved in the 2010 Maryland item bank for the possible use as core linking items in the future. Classical and Rasch analyses on these core items can be found in section of chapter 1.8, *Validation Check with the 2010 Core Items*.

While a few core linking items appeared only on either operational form A or F (i.e., unique core linking), most core linking items (i.e., common core linking) appeared on both operational forms. As a result, the common core linking items appearing on both forms were used for both form-to-form and year-to-year linking. The unique core linking items were used only for year-to-year linking.

The role of the core linking items was to place the 2010 scale on the 2006 scale. Because these core linking items carried their operational item parameters on the 2006 scale, they were included in the 2010 year-to-year linking pool. Classical analysis on these items can be found in section of chapter 1.8, *P-Value Check with Year-to-Year Core Linking Items*, and calibration, linking and equating procedures on these core linking items can be found in chapter 1.9, *Linking, Equating, and Scaling Procedures of the 2010 MSA-Math*.

The Role of Operational SPR, BCR, and ECR Items

SPR, BCR, and ECR items were divided into one of the following two categories: unique core or common core items. Only the common core items appearing on both operational forms were used for form-to-form calibration and linking. Because these items were not included in the 2010 year-to-year linking pool, new Rasch item and step difficulty parameters were estimated with the 2010 live, operational data set (i.e., stratified random sample). These new item and step difficulty parameters were used to produce each student's theta estimate. More detailed information about how much these items changed across years in terms of classical and Rasch item difficulty can be found in section of chapter 1.8, *Validation Check with the 2010 Core Items*.

Table 1.3 Item Type of Content Standard for the 2010 MSA-Math: Grades 3 and 4

Grade	Standard	Item Type	No. of Items of Each Form	
			A	F
3			65	65
	Algebra	*SR*, *BCR*	13	13
	Geometry	*SR*, *BCR*	8	8
	Measurement	*SR*, *BCR*	7	7
	Statistics	*SR*, *BCR*	12	12
	Probability	*SR*	2	2
	Number Computation	*SR*, *BCR*	16	16
	Process	*BCR*	7	7
4			64	64
	Algebra	*SR*, *BCR*	14	14
	Geometry	*SR*, *BCR*	7	7
	Measurement	*SR*, *BCR*	7	7
	Statistics	*SR*, *BCR*	8	8
	Probability	*SR*, *BCR*	7	7
	Number Computation	*SR*, *BCR*	14	14
	Process	*BCR*	7	7

Note. *SR* items are selected response items, and *BCR* items are brief constructed response items. Form A designates the forms A, B, C, D, and E. Form F designates the forms F, G, H, J, and K.

Table 1.4 Item Type of Content Standard for the 2010 MSA-Math: Grades 5 and 6

Grade	Standard	Item Type	No. of Items of Each Form	
			A	F
5			65	65
	Algebra	*SR*, *BCR, ECR*	15	15
	Geometry	*SR*, *BCR*	6	6
	Measurement	*SR*, *BCR*	8	8
	Statistics	*SR*, *BCR*	9	9
	Probability	*SR*, *BCR*	4	4
	Number Computation	*SR*, *BCR*	15	15
	Process	*BCR, ECR*	8	8
6			62	62
	Algebra	*SR*, *BCR, ECR*	14	14
	Geometry	*SR*, *BCR*	8	8
	Measurement	*SR*, *BCR*	6	6
	Statistics	*SR*, *BCR*	9	9
	Probability	*SR*, *BCR*	4	4
	Number Computation	*SR*, *BCR*	14	14
	Process	*BCR, ECR*	7	7

Note. SR items are selected response items, *BCR* items are brief constructed response items, and *ECR* items are Extended Constructed Response. Form A designates the forms A, B, C, D, and E. Form F designates the forms F, G, H, J, and K.

Table 1.5 Item Type of Content Standard for the 2010 MSA-Math: Grades 7 and 8

Grade	Standard	Item Type	No. of Items of Each Form	
			A	F
7			62	62
	Algebra	*SR,SPR, BCR, ECR*	14	14
	Geometry	*SR, SPR, ECR*	7	7
	Measurement	*SR, SPR, BCR*	6	6
	Statistics	*SR, SPR, BCR, ECR*	9	9
	Probability	*SR, SPR, BCR*	5	5
	Number Computation	*SR, SPR*	14	14
	Process	*BCR, ECR*	7	7
8			62	62
	Algebra	*SR,SPR, BCR, ECR*	15	15
	Geometry	*SR, SPR, ECR*	8	8
	Measurement	*SR, SPR, BCR*	5	5
	Statistics	*SR, SPR, BCR, ECR*	9	9
	Probability	*SR, SPR, BCR*	5	5
	Number Computation	*SR, SPR*	12	12
	Process	*BCR, ECR*	8	8

Note. SR items are selected response items, *SPR* items are student-produced response, *BCR* items are brief constructed response items, and *ECR* items are extended constructed response. Form A designates the forms A, B, C, D, and E. Form F designates the forms F, G, H, J, and K.

Table 1.6 Item Distribution of Each Content Standard for the 2010 MSA-Math: Grade 3

Form	Total Item Number of Each Standard							**Total # of Item**
	1*	2*	3*	4*	5*	6*	7*	
A	13	8	7	12	2	16	7	**65**
F	13	8	7	12	2	16	7	**65**

Note. 1*. Algebra; 2*. Geometry; 3*. Measurement; 4*. Statistics; 5*. Probability; 6*. Numbers and Computation; 7*. Process

Table 1.7 Total and Reporting Content Standard Scores for the 2010 MSA-Math: Grade 3

Form	Total and Reporting Standard Scores					
	1	2&3	4&5	6	7	**Total Score**
A	13	15	14	16	14	**72**
F	13	15	14	16	14	**72**

Table 1.8 Item Type and Score Point Distribution for the 2010 MSA-Math: Grade 3

Form	# of SR Item	# of BCR Item		**Total # of Item**	Scores of SR	Scores of BCR		**Total Score**
		Step A	Step B			Step A	Step B	
A	51	7	7	**65**	51	7	14	**72**
F	51	7	7	**65**	51	7	14	**72**

Table 1.9 Item Distribution of Each Content Standard for the 2010 MSA-Math: Grade 4

Form	Total Item Number of Each Standard							**Total # of Item**
	1*	2*	3*	4*	5*	6*	7*	
A	14	7	7	8	7	14	7	**64**
F	14	7	7	8	7	14	7	**64**

Note. 1*. Algebra; 2*. Geometry; 3*. Measurement; 4*. Statistics; 5*. Probability; 6*. Numbers and Computation; 7*. Process

Table 1.10 Total and Reporting Content Standard Scores for the 2010 MSA-Math: Grade 4

Form	Total and Reporting Standard Scores					
	1	2&3	4&5	6	7	**Total Score**
A	14	14	15	14	14	**71**
F	14	14	15	14	14	**71**

Table 1.11 Item Type and Score Point Distribution for the 2010 MSA-Math: Grade 4

Form	# of SR Item	# of BCR item		**Total # of Item**	Scores of SR Item	Scores of BCR		**Total Score**
		Step A	Step B			Step A	Step B	
A	50	7	7	**64**	50	7	14	**71**
F	50	7	7	**64**	50	7	14	**71**

Table 1.12 Item Distribution of Each Content Standard for the 2010 MSA-Math: Grade 5

Form	Total Item Number of Each Standard							**Total # of Item**
	1*	2*	3*	4*	5*	6*	7*	
A	15	6	8	9	4	15	8	**65**
F	15	6	8	9	4	15	8	**65**

Note. 1*. Algebra; 2*. Geometry; 3*. Measurement; 4*. Statistics; 5*. Probability; 6*. Numbers and Computation; 7*. Process

Table 1.13 Total and Reporting Content Standard Scores for the 2010 MSA-Math: Grade 5

Form	Total and Reporting Standard Scores					
	1	2&3	4&5	6	7	**Total Score**
A	15	14	13	15	17	**74**
F	15	14	13	15	17	**74**

Table 1.14 Item Type and Score Point Distribution for the 2010 MSA-Math: Grade 5

Form	# of SR Item	# of BCR Item		# of ECR Item		**Total # of Item**	Scores of SR	Scores of BCR		Scores of ECR		**Total Score**
		Step A	Step B	Step A	Step B			Step A	Step B	Step A	Step B	
A	49	7	7	1	1	**65**	49	7	14	1	3	**74**
F	49	7	7	1	1	**65**	49	7	14	1	3	**74**

Table 1.15 Item Distribution of Each Content Standard for the 2010 MSA-Math: Grade 6

Form	Total Item Number of Each Standard							**Total # of Item**
	1*	2*	3*	4*	5*	6*	7*	
A	14	8	6	9	4	14	7	**62**
F	14	8	6	9	4	14	7	**62**

Note. 1*. Algebra; 2*. Geometry; 3*. Measurement; 4*. Statistics; 5*. Probability; 6*. Numbers and Computation; 7*. Process

Table 1.16 Total and Reporting Content Standard Scores for the 2010 MSA-Math: Grade 6

Form	Total and Reporting Standard Scores					
	1	2&3	4&5	6	7	**Total Score**
A	14	14	13	14	15	**70**
F	14	14	13	14	15	**70**

Table 1.17 Item Type and Score Point Distribution for the 2010 MSA-Math: Grade 6

Form	# of SR Item	# of BCR Item		# of ECR Item		**Total # of Item**	Scores of SR	Scores of BCR		Scores of ECR		**Total Score**
		Step A	Step B	Step A	Step B			Step A	Step B	Step A	Step B	
A	48	6	6	1	1	**62**	48	6	12	1	3	**70**
F	48	6	6	1	1	**62**	48	6	12	1	3	**70**

Table 1.18 Item Distribution of Each Content Standard for the 2010 MSA-Math: Grade 7

Form	Total Item Number of Each Standard							**Total # of Item**
	1*	2*	3*	4*	5*	6*	7*	
A	14	7	6	9	5	14	7	**62**
F	14	7	6	9	5	14	7	**62**

Note. 1*. Algebra; 2*. Geometry; 3*. Measurement; 4*. Statistics; 5*. Probability; 6*. Numbers and Computation; 7*. Process

Table 1.19 Total and Reporting Content Standard Scores for the 2010 MSA-Math: Grade 7

Form	Total and Reporting Standard Scores					
	1	2&3	4&5	6	7	**Total Score**
A	14	13	14	14	17	**72**
F	14	13	14	14	17	**72**

Table 1.20 Item Type and Score Point Distribution for the 2010 MSA-Math: Grade 7

Form	# of SR Item	# of SPR Item	# of BCR Item		# of ECR Item		**Total # of Item**	Scores of SR	Scores of SPR	Scores of BCR		Scores of ECR		**Total Score**
			Step A	Step B	Step A	Step B				Step A	Step B	Step A	Step B	
A	36	12	4	4	3	3	**62**	36	12	4	8	3	9	**72**
F	36	12	4	4	3	3	**62**	36	12	4	8	3	9	**72**

Table 1.21 Item Distribution of Each Content Standard for the 2010 MSA-Math: Grade 8

Form	Total Item Number of Each Standard							**Total # of Item**
	1*	2*	3*	4*	5*	6*	7*	
A	15	8	5	9	5	12	8	**62**
F	15	8	5	9	5	12	8	**62**

Note. 1*. Algebra; 2*. Geometry; 3*. Measurement; 4*. Statistics; 5*. Probability; 6*. Numbers and Computation; 7*. Process

Table 1.22 Total and Reporting Content Standard Scores for the 2010 MSA-Math: Grade 8

Form	Total and Reporting Standard Scores					
	1	2&3	4&5	6	7	**Total Score**
A	15	13	14	12	19	**73**
F	15	13	14	12	19	**73**

Table 1.23 Item Type and Score Point Distribution for the 2010 MSA-Math: Grade 8

Form	# of SR Item	# of SPR Item	# of BCR Item		# of ECR Item		**Total # of Item**	Scores of SR	Scores of SPR	Scores of BCR		Scores of ECR		**Total Score**
			Step A	Step B	Step A	Step B				Step A	Step B	Step A	Step B	
A	34	12	5	5	3	3	**62**	34	12	5	10	3	9	**73**
F	34	12	5	5	3	3	**62**	34	12	5	10	3	9	**73**

1.5 Operational Test Form Construction Using the Rasch Model

The selection of items to be included in the final operational test forms of the 2010 MSA-Math required a careful consideration based on test blueprints, classical item analyses, *DIF* analyses, and IRT analyses. Specifically, the Rasch model (i.e., 1-Parameter Logistic IRT) played a major role in constructing the 2010 operational forms. First, Pearson suggested the following guidelines:

- Do not include items that are too easy or too hard.
- Do not include *BCR* items with score distributions that do not elicit the full range of rubric scores.
- Do not include items with *DIF* classifications "C" for the *SR* items and "CC" for the *BCR* items *unless* they have been deemed acceptable by the external review of content experts.
- Finally, do not include items which have Rasch *Infit* and *Outfit* mean-squares lower than .5 or higher than 1.5. More specific information on Rasch *Infit* and *Outfit* mean-squares can be found in the third part of the 2010 technical report, *Overview of Statistical Summaries*.

A procedure for using IRT methods to build tests that meet any desired set of test specifications was outlined by Lord (1977). The procedure utilizes an item bank with item parameter estimates available for the IRT model of choice, with accompanying information functions. The steps in the procedure suggested by Lord (1977) are as follows:

- First, the shape of desired test information needs to be decided. This was termed as the "target information function" by Lord (1977).
- Second, specific items need to be selected from the item bank with item information functions that will fill up hard-to-fill areas under the target information function.
- Third, the test information function after test items are added needs to be recalculated.
- Fourth, until the test information function approximates the target information function to a satisfactory degree, test items need to keep on being selected.

It should be noted that these steps were implemented within a framework defined by the content specification of the test. In addition, math content specialists from MSDE reviewed the final test forms of the 2010 MSA-Math. The following table and figure show an example of the 2010 MSA-Math operational test form construction using the Rasch (i.e., 1-PL IRT) method. Detailed information about constructing operational forms using the Rasch method can be obtained from either MSDE or Pearson.

Table 1.24 The 2010 Math Operational Test Construction Using the Rasch Model: Grade 4 Form A

Item Type	P-value	A	D_{i1}	D_{i2}
BCR_A	0.50	1.00	1.3917	
BCR_B	0.43	1.00	2.1573	-0.5064
SR	0.98	1.00	-3.5593	
BCR_A	0.88	1.00	-1.1338	
BCR_B	0.45	1.00	1.8916	-0.6131
SR	0.79	1.00	-0.3698	
SR	0.89	1.00	-1.2509	
BCR_A	0.82	1.00	-0.5527	
BCR_B	0.44	1.00	1.8966	-0.2409
SR	0.83	1.00	-0.7461	
SR	0.81	1.00	-0.5617	
SR	0.79	1.00	-0.2781	
SR	0.94	1.00	-2.7781	
SR	0.79	1.00	-0.4437	
BCR_A	0.44	1.00	1.6468	
BCR_B	0.40	1.00	2.0517	0.3771
SR	0.55	1.00	0.9747	
SR	0.84	1.00	-0.7274	
SR	0.94	1.00	-2.1097	
SR	0.80	1.00	-0.7990	
SR	0.66	1.00	0.1763	
SR	0.83	1.00	-1.0550	
SR	0.84	1.00	-0.7925	
SR	0.51	1.00	1.2932	
SR	0.80	1.00	-0.4685	
SR	0.79	1.00	-0.3646	
BCR_A	0.37	1.00	2.1316	
BCR_B	0.47	1.00	1.6184	0.4623
SR	0.53	1.00	1.1949	
SR	0.66	1.00	0.4883	
BCR_A	0.75	1.00	-0.0713	
BCR_B	0.70	1.00	0.1552	
BCR_A	0.73	1.00	0.1264	-0.2255
BCR_B	0.48	1.00	1.5957	-0.7421
SR	0.40	1.00	1.4979	
SR	0.60	1.00	0.3940	
SR	0.50	1.00	0.9009	
SR	0.97	1.00	-2.8436	
SR	0.84	1.00	-0.9014	

Table 1.24 (Continued)

Item Type	P-value	A	D_{i1}	D_{i2}
SR	0.65	1.00	0.0796	
SR	0.71	1.00	0.1774	
SR	0.83	1.00	-0.9767	
SR	0.48	1.00	1.0327	
SR	0.51	1.00	0.9291	
SR	0.75	1.00	-0.4674	
SR	0.68	1.00	-0.0118	
SR	0.85	1.00	-1.2169	
SR	0.55	1.00	0.6901	
SR	0.50	1.00	0.6281	
SR	0.90	1.00	-1.3494	
SR	0.68	1.00	0.3863	
SR	0.68	1.00	0.1618	
SR	0.81	1.00	-0.8522	
SR	0.75	1.00	-0.2435	
SR	0.93	1.00	-2.3000	
SR	0.35	1.00	1.7570	
SR	0.70	1.00	-0.1077	
SR	0.60	1.00	0.4796	
SR	0.69	1.00	0.1230	
SR	0.63	1.00	0.3118	
SR	0.76	1.00	-0.3619	
SR	0.81	1.00	-0.8156	
SR	0.73	1.00	-0.1831	
SR	0.70	1.00	-0.1060	

Note. A: item discrimination; D_{i1} : first structure measure estimate; D_{i2} : second structure measure estimate.

Note. Please refer to section 3.3 of this technical report to get detailed information about how to estimate structure measure estimate ($D_{ij} = D_i + F_{ij}$)

Note. BCR_A: Step A item; BCR_B: Step B item

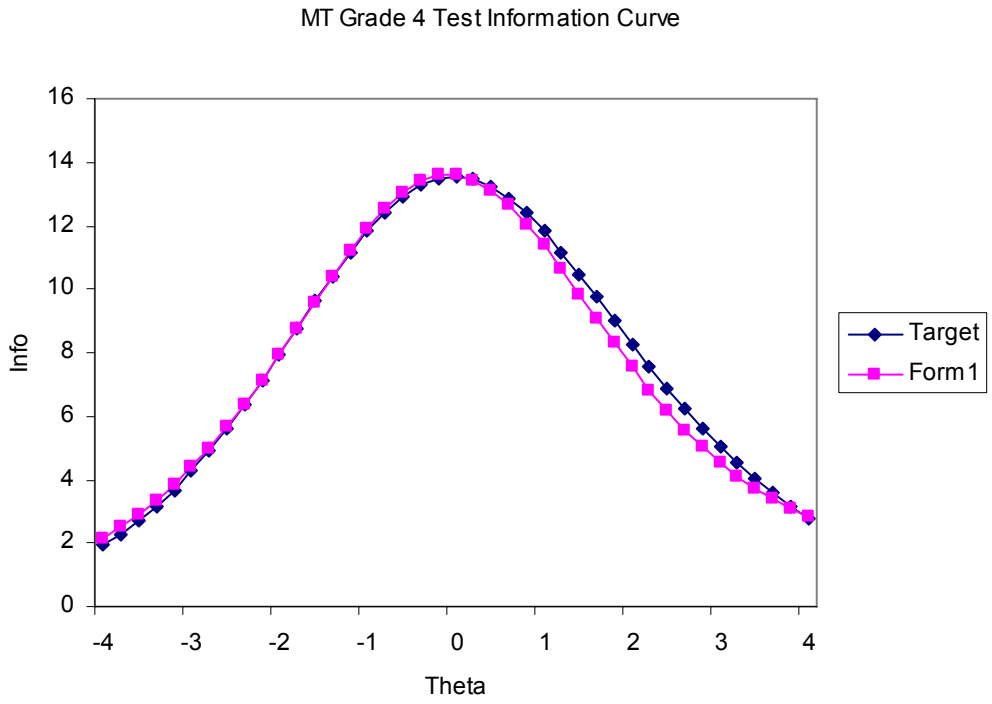

Figure 1.1 Test Information Curves of Target Form vs. Current Year's Math Operational Test Form

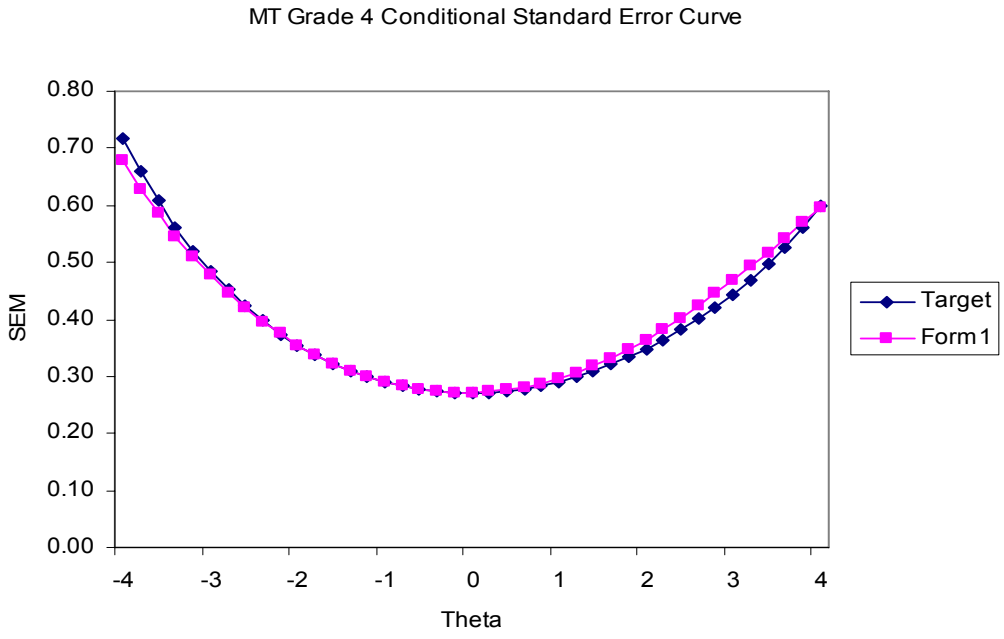

Figure 1.2 Standard Errors of Target Form vs. Current Year's Math Operational Test Form

1.6 Test Administration of the 2010 MSA-Math

The 2010 MSA-Math test was administered to all students in grades 3 through 8 except for students taking the Alt-MSA-Math or the Mod-MSA-Math. Pearson coordinated the test administration procedures with MSDE prior to implementation. This chapter was prepared to provide general information about the 2010 test administration. Detailed information about the 2010 test administration can be obtained from the 2010 Test Administration and Coordination Manual (TACM) and Examiners Manual (EM) which are available from either MSDE or Pearson.

Test Materials

All test materials had to be stored in a secure location prior to test administration. The School Test Coordinator (STC) provided test administration training and test materials to the test examiners. The Daily Testing Materials Tracking Record (or an equivalent form designed by the LEA) was used to track the distribution and return of Test Books.

Before testing began, the Test Examiners (TEs) carefully inventoried all test materials given to them, as they were accountable for the return of all secure materials at the end of testing. The TEs checked to ensure they have all the materials they needed for testing.

For the Test Examiner, Pearson provided the following materials:

- MSA Examiner's Manual for grades 3 through 8- Math
- Pre-printed and generic labels
- Scoring Service Identification (SSID) sheets

For each student, the following materials were provided by Pearson:

- Test/Answer Book
- Special accommodations testing materials, if necessary

For each student, the following additional materials were provided by school or student:

- Two No. 2 pencils with erasers
- Blank scratch paper
- Classroom Calculator for Day 1 (all grades)
- Classroom ruler with both U.S. customary and metric measurements (all grades)
- Classroom protractor for grades 5 through 8
- Classroom compass for grades 7 and 8 only

Each classroom used for the assessment also needed the following additional materials:

- A sign for the door that reads "Testing: Do not Disturb"
- A digital clock or a watch, or clock with a second hand

Two test-related Examiners Manuals (EMs) were developed for the 2010 MSA: one version for reading and the other for mathematics for use in all grades 3-8. Developed in partnership with

MSDE, the EMs contained instructions for preparation and administration of the test. In addition to the EMs, one Test Administration and Coordination Manual (TACM) was developed for use by the Local Accountability Coordinators (LAC) and building-level School Test Coordinators (STC). Included in this manual were instructions for preparation of materials for testing, monitoring of testing, and packaging of materials for return to Pearson for scoring. The TACM was distributed and reviewed during a workshop in January for STCs and LACs, with duplicates sent to each school along with its testing materials.

Test Administration Schedule

The primary test window for MSA was established by MSDE (March 8-17, 2010, with make-up testing held March 18-23, 2010). However, each LEA (Local Education Agency) set a specific schedule for administration of the MSA within that window for their district. For a given test, grade, content area, and test format, all testing (with the exception of the make-up administration) had to take place on the same schedule. Each LEA schedule was submitted to MSDE in advance and approved for each district by the state. For example, all Grade 3 MSA-Math must be administered on the same days throughout the LEA. In addition, each content area in each grade was tested on two days during the window. In any given grade, one content area's primary testing window was completed before beginning the second content area's primary testing window.

The MSA-Math testing schedule allowed approximately 2 hours and 30 minutes for testing on Day 1 and 1 hour and 45 minutes on Day 2 (including preparation time and breaks).

For the 2010 MSA-Math, the primary testing days were as follows:

- Test materials delivered to schools (Examiner's Manuals, Test/Answer Books, and Test Coordinator's Kits) — February 16-22, 2010
- Mathematics Primary Testing Window — March 8 – March 17, 2010
- Make-up Testing Window — March 18 – March 23, 2010

If a student was absent on the testing days, a make-up test was administered on any two consecutive days within the testing window. If a school had an unscheduled closing or delayed opening that prohibited the administration from occurring on the scheduled testing dates, the STCs were consulted by LACs to determine the testing schedule to be followed.

During the administration of the 2010 MSA-Math, MSDE had testing monitors in selected schools observing administration procedures and testing conditions. All monitors had identification cards for security purposes. There was no prior notification of which schools would be monitored, but monitors followed local procedures for reporting to the school's main office and giving proper notification that an MSDE monitor was in the building.

Student Participation

MSDE calculates actual participation of students who took the test. This means that the schools are held accountable not only for student achievement on MSA or Mod-MSA testing, but also they are accountable to ensure that at least 95% of students participate in testing. Accordingly, schools should do all they can to test all students on MSA, Mod-MSA, or Alt-MSA, as applicable.

All students in grades 3 through 5 had to participate in the 2010 MSA-Math, and all students in grades 6 through 8 had to participate in either the 2010 MSA-Math or Mod-MSA-Math. All students in grade 6 through 8 had to participate in the 2010 Mod-MSA-Math, if determined to be eligible by the student's IEP. The only exception was that students with severe cognitive disabilities were assessed by the *Alternate Maryland School Assessment* (Alt-MSA) instead of the regular MSA-Math or Mod-MSA-Math. The criteria that students need in order to be tested in the Alt-MSA program instead of the MSA-Math can be viewed in section 5, Appendix A of the TACM.

Participation of English Language Learners (ELLs) in the MSA-Math or the Mod-MSA-Math

There are special rules that apply to the participation of English Language Learners (ELLs) in the MSA-Math and the Mod-MSA-Math, as follows:

For the MSA-Math and Mod-MSA-Math, ELL students must participate in MSA-Math or Mod-MSA-Math regardless of how recently they entered the U.S. educational system. For ELL students in their first year of enrollment in a U.S. school, "participation" in the MSA-Math or the Mod-MSA-Math is defined as allowing the student to attempt the test for at least 20 minutes. If, after 20 minutes, the TE determines in his or her professional judgment that the student does not possess sufficient English fluency to be able to continue testing, the test administration for that student may be concluded at that time.

Accommodations for Assessment

Accommodations for assessment of students with disabilities (i.e., students having an Individualized Education Program or a Section 504 Plan) and students for English Language Learners (ELL) had to be approved and documented according to the procedures and requirements outlined in the document entitled "Maryland Accommodations Manual: A Guide to Selecting, Administrating, and Evaluating the Use of Accommodations for Instruction and Assessment" (MAM). A copy of the most recent edition of this document is available electronically on the LAC and STC web pages at https://docushare.msde.state.md.us/docushare.

No accommodations could be made for students merely because they were members of an instructional group. Any accommodation had to be based on individual needs and not on a category of disability area, level of instruction, environment, or other group characteristics. Responsibility for confirming the need and appropriateness of an accommodation rested with the LAC and school-based staff involved with each student's instructional program. A master list of all students and their accommodations had to be maintained by the principal and submitted to the LAC, who provided a copy to MSDE upon request. Please refer to Section 1 of the 2010 TACM for further information regarding testing accommodations.

Large-Print and Braille Test Books and Kurzweil™ Test Forms on CD

The MSA-Math was administered to those requiring (1) large-print Student Test/Answer Books or (2) Braille Test Books, or (3) Kurzweil™ Test Forms on CD for a verbatim reading

accommodation. For large-print Test/Answer Books, Braille Test Books, and Kurzweil™ Test Forms on CD, student responses were transcribed into the standard-size Test/Answer Book following testing.

The student's name, LEA number, and school number were written on the large-print Test/Answer Book for proper transcription into the standard-size Test/Answer Book.

The pre-printed student ID label was affixed to the standard-size Test/Answer Book containing the transcribed responses, and not to the large-print Test/Answer Book or Braille books. The bubbles on the demographic page of the standard-size Test/Answer Book were not filled in if there was a pre-printed student ID label for the student.

A certified Test Examiner (TE) transcribed the student responses into a standard-size Test/Answer Book exactly as given by the student. The standard-size Test/Answer Book with the pre-printed or general label attached was returned to Pearson with all other Test/Answer Books.

Large-Print Test/Answer Books and Braille Test/Answer Books containing the original student responses prior to transcription are to be returned with Non-Scorable materials. Any Test/Answer Books which were used as source documents for transcription were invalidated by drawing a large slash across the student demographic page with a black permanent marker.

Once the student responses had been transcribed, the transcribed Test/Answer Book was returned for scoring with the standard-size materials. Specific packing instructions are provided in the 2010 TACM in sections 2 and 3.

Verbatim Reading Accommodation and Kurzweil™ Test Form on CD

Students who had a verbatim reading accommodation documented in their Individual Education Plan (IEP), ELL Plan, or Section 504 Plan, and who received that accommodation in regular instruction, received the accommodation on the 2010 MSA-Math. The accommodation was provided by a live reader or through technology. Appendix L of the 2010 TACM provided information on verbatim reading instruction. Technology used to provide the verbatim reading accommodation was Kurzweil™ reading software. Official, secure electronic copies of the test were ordered through the LAC. MSDE encouraged (but did not require) the use of the Kurzweil™ software to ensure uniformity in the delivery of the verbatim reading accommodation throughout the state.

Students using Kurzweil™ software had to familiarize themselves with its operation prior to the test administration. When there were technical difficulties with Kurzweil™ a certified staff member was used instead. Kurzweil™ Test Form CDs were shipped by Pearson. After testing, schools returned the CDs to Pearson with the non-scorable secure materials.

Administration Procedures for Students with IEP, 504 Plan, or ELL Plan Permitting a Dictated Responses or Use of Word Processor

A student whose IEP, 504 Plan, or ELL Plan permitted a dictated response had his/her responses transcribed at the school level by an eligible TE, or by a staff member working under the direct supervision of a certified TE, into the student's Test/Answer Book with a pre-printed or generic ID label attached.

A student whose IEP, 504 Plan, or ELL plan permitted the use of a word processor had his/her responses transcribed by hand or under the direct supervision of an eligible TE or STC exactly as the student entered his/her responses on the word processor. The student's responses were always

transcribed at the school level into the student's Test/Answer Book with the pre-printed or generic ID label attached. After the student's responses had been transcribed, the memory of the word processor was cleared. The original word-processed print-out was returned to Pearson with the non-scorable materials.

Test Format

All grade levels of the MSA-Math used a Test Book format in which students wrote their answers directly in the Test Book. There were 10 forms of MSA-Math. Different test forms were administered to students in each classroom participating in math tests, and each test form was identified by color and form number/letter. All forms of the MSA Test/Answer Books for each grade had the same grade designation and picture on the front cover. The Test/Answer Books were spiraled within a classroom, and each student used a combined Test/Answer Book.

Since the Test/Answer Books were scanned for scoring, students were encouraged not to use highlights in any part of the book. Although students might be accustomed to using highlighters in daily instruction, highlighting in the Test/Answer Book could obliterate information in a student's book when it was scanned for scoring. As an alternative to highlighting, students were allowed to lightly circle or underline information in test items or perform calculations to help them in responding, as long as markings did not interfere with the bubbled answer choice area and/or the track marks along the outside margins of each page.

Security of Test Materials

The following code of ethics conforms to the Standards for Educational and Psychological Testing developed by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (Pearson, 2010):

> It is breach of professional ethics for school personnel to provide verbal or nonverbal clues or answers, teach items on the test, share writing prompts, coach, hint, or in any way influence a student's performance during the testing situation. A breach of ethics may result in invalidation of test results and local education agency (LEA) or MSDE disciplinary action. (p. 11)

The Test/Answer Books for the 2010 MSA-Math were confidential and kept secure at all times. Unauthorized use, duplication, or reproduction of any or all portions of the assessment was prohibited, which is reflected by the following statement (Pearson, 2010):

> Violation of security can result in prosecution and/or penalties as imposed by the Maryland State Board of Education and/or State Superintendent of Schools in accordance with the COMAR 13A.03.04 and 13A.12.05. (p. 11)

All materials were treated as confidential and placed in locked areas. Secure and non-secure test materials were as follows:

- Secure materials: Test/Answer Books (including large-print and Braille), Kurzweil™ test forms on CD, and used scratch paper
- Non-secure materials: TACM, Examiner's Manuals, unused pre-printed student and generic ID labels, unused FedEx return shipping labels, and unused green/orange shipping labels

1.7 Hand Scoring Procedures of the 2010 MSA-Math

Students' responses to *SR* and *SPR* items were machine-scored, and their responses to *BCR* and *ECR* items were individually read and scored by Pearson.

Hand Scoring Staff

The PSC Project Manager (PSC PM), Content Specialist (CS), and Scoring Directors (SD) participated in the rangefinding sessions in Maryland. (Detailed information about rangefinding procedures can be found in the following portion of this section: *Development Procedures for Rangefinding*.) The SD was responsible for maintaining annotations and meeting minutes from all sessions. These notes were a record of the comments and decisions made by the MSDE personnel and members of the Maryland teacher committee. These notes were utilized by the SD responsible for training the Scoring Supervisors and Scorers for the respective Maryland prompts.

1) Scorer

A graduate of a four-year accredited college or university who had completed the Maryland-specific domain training. The scorers were eligible to score items for which they had been trained and successfully qualified.

2) Scoring Supervisor

A reader who directly monitored the scoring of a team of Scorers and retrained as needed. The reader had successfully completed the PSC Scoring Supervisor training.

3) Scoring Director (SD)

An experienced and knowledgeable PSC team leader who was responsible for selecting a wide variety of student responses for such activities as rangefinding and building training materials. Selected papers were then submitted to MSDE for comment and approval. Scoring directors remained on the project as rangefinding participants and trainers. Scoring directors worked with scoring supervisors and the Content Specialist to oversee the scoring of several items. An SD's main duty during scoring was to rule on validity of questionable papers and to maintain consistency in scoring decisions.

4) Content Specialist (CS)

Experienced content/training personnel who had served as SDs and were selected by the Scoring Resources staff and Project Manager to train and support Scoring Directors for Maryland.

Scorer Recruitment and Qualifications

All Scorers for MSDE had to provide Pearson their résumé and documentation of a four-year college degree. Human Resources made every effort to recruit Scorers with a teaching background and to match Scorers to projects which suited their educational background and previous scoring experience. Regardless of educational background, applicants then participated in a one-day general introductory training workshop presented by a PSC staff member. These workshops allowed Pearson to introduce potential Scorers and Scoring Supervisors to large-scale scoring in general and to the Maryland rubric specifically. The PSC staff member who presented the workshop evaluated potential Supervisors and submitted these evaluations to the PSC Site Manager with his/her recommendations. Those who successfully completed the workshop were

added to Pearson's general pool of potential Scorers and Supervisors of MSA Math. This addition to the scoring pool did not qualify these Scorers for scoring the MSDE program.

Scoring Supervisor Selection

The training for new Scoring Supervisors consisted of a two-day course focusing on the duties and responsibilities necessary to successfully manage a team of Scorers. The workshop was led by the PSC Site Manager and Scoring Directors. The instruction included a review of PSC policies and procedures, sessions on use of ePEN and the monitoring reports to track a Scorer's speed and accuracy, role playing activities which explored various situations that could occur with Scorers during the scoring of a project, and Scorer counseling and retraining guidelines. Upon completion of the workshop, the PSC Site Manager and Scoring Directors in conjunction with the Content Specialist reviewed each participant's performance, making sure that each had a complete understanding of the Scoring Supervisor role and its responsibilities. Any participant they found who did not perform to their satisfaction was not added to the qualified Supervisor list.

Scoring Supervisor Project Training and Qualification

Project-specific Supervisor training for MSDE was conducted in the days immediately preceding Scorer training. This training began with the SD reading the rubrics aloud and answering any questions the Supervisor might have regarding the rubric. The SD then read each anchor paper aloud to the Supervisors. Each response in the anchor set was thoroughly explained, including the notes and comments of the rangefinding committee. Practice Set 1 was reviewed next. The Supervisors scored the practice set individually in the electronic scoring system (ePEN) as well as recorded their scores on a paper copy of the practice set, and then waited for all Supervisors to complete scoring the set. When everyone had completed scoring the training set, the SD discussed the responses one by one, focusing on why each received that score and not another. The SD reviewed with the group the reason for assigning each score point and discussed each paper in its entirety. The Supervisors were then ready to score Practice Set 2. Practice Set 2 was scored and reviewed exactly as Practice Set 1.

Having thoroughly discussed both practice sets with the group, the SD explained that in order for a participant to qualify as a Scoring Supervisor, it was required that the Supervisor should score at least 90% perfect agreement on Step A and 80% perfect agreement on Step B on two of three qualifying sets or one of two qualifying sets, depending on the number of sets available for each item (Qualification Rules, Attachment M). The Supervisors scored the first qualifying set individually and recorded their scores in ePEN. As each Supervisor finished scoring, the SD reviewed the qualifying reports before allowing the Supervisor to proceed to the next qualifying set. Each response was reviewed and any questions the Supervisor had were addressed before the Supervisor attempted the next qualifying set. The Supervisor followed the same procedure with Qualifying set 2 (and set 3 if available). Supervisors had to pass one of two or two of three sets (depending on the number of qualifying sets available per item) with 80% agreement for Math Step B and 90% in Math Step A as specified in the qualification rules or they would be released from the MSDE project.

Scoring Supervisor Duties

Scoring Supervisors were responsible for monitoring the training and qualifying of the Scorers assigned to their team. The Supervisors assisted the SD, if requested, during the training of the Scorers. The Supervisor was responsible for monitoring Scorers' progress through the qualifying sets. The Supervisor was also responsible for monitoring each Scorer's assignment of scores to

the responses. Additionally, the Supervisor reviewed the statistical reports with each individual on the team. The Supervisor consulted the SD regarding variations by the team members from the acceptable standards (95% for Math Step A, and 85% for Math Step B). The Supervisor had the initial responsibility to see that the Scorer maintained the set standards through individual retraining. The SD monitored the Supervisor by reviewing team statistics and working one-on-one with the Supervisor.

Scoring Director Selection and Qualification

The candidates for Scoring Director had been recommended by the Content Specialist, PSC Resource Staffing Managers or Site Manager. The recommendations were based upon the evaluations the candidates received as Scorers and Supervisors and were part of their personnel file. The candidates generally had been Supervisors on large-scale projects for multiple teams, and/or they had served as Supervisors on small-scale projects where Supervisors trained their individual teams. They had been evaluated on their ability to train Scorers as well as their ability to monitor the scoring accuracy and consistency of Scorers. These evaluations were submitted in writing at the end of each scoring project by the Site Managers and SDs that had observed the work of the SD candidates.

Scoring Director Project Training

The SDs familiarized themselves with the rubric. Any questions regarding the rubric were addressed by the PSC Content Specialist or MSDE. The next step was for the SD to become familiar with all their items and all training materials and scoring decisions/issues associated with their items prior to Supervisor training.

Scoring Director Duties

The SD's job was to conduct the training of the Supervisors and Scorers, oversee the actual scoring of the papers, monitor the work of the Supervisor, and act as the decision-maker for situations or questions that may arise during the scoring process. For example, all condition code (foreign language, off-topic, off-mode, etc.) responses were reviewed by the SD, who had to confirm any such decision and ensure consistency of decisions. (Blank condition codes were assigned at the Scorer level and did not require SD confirmation.) Additionally the SD and Supervisor conducted all resolution readings. The resolution score became the reported score.

The SD also reviewed any potential questionable content responses and forwarded those to the Content Specialist to consult with MSDE before processing.

The SD was also responsible for daily statistical review and analysis of all monitoring reports to ensure the quality of the scoring. Review of the data allowed the SD not only to monitor the Scorer but also to provide the Supervisor with additional input. Available data included 1) individual Scorer agreement rates between two independent scorings; 2) score point distributions by Scorer and trend review; 3) prompt statistics for agreement rates and score point distributions; 4) Resolution data; 5) scorer-level and item-level agreement on validity papers pre-scored by MSDE.

Scorer Training

Scorer training was led by the SD, and each SD was responsible for training the items he/she monitored throughout scoring. After sufficient student responses were scored for equating purposes for the first item, the SD reconvened the group and trained the second item. Training began with the definition and an overview of holistic scoring. Training continued with a reading

and discussion of the generic rubric and item, and then the student responses in the anchor set were read and discussed. In the anchor set the scores had been recorded on the student responses and were arranged in ascending point-scale order. Each annotated anchor response was read aloud and discussed thoroughly. Emphasis was placed on the Scorers' understanding of how the responses differed from one another in incremental quality, how each response reflected the description of its score point as generalized in the scoring rubric, and how each reflected the MSDE's standard for application of each score point.

Once Scorers had all their questions answered and the discussion of the anchor set was finished, the Scorers began to assign scores to the first practice set. Each Scorer independently read and scored the responses in the practice set in the electronic scoring system (ePEN). The correct scores were then read to the group when everyone had completed the scoring. In addition, each practice paper was discussed as to reasons for applying each given score. At this point, Scorers interacted with the SD in discussing the characteristics of each response that earned the assigned score point. The same format was followed for each practice set. During this process, the job of the Scorer was to internalize the scoring scale and adjust his or her individual scoring to conform to that scale. Once all practice papers had been scored and fully discussed, Scorers began the qualifying process.

For MSA Math, there were two or three qualifying sets, depending on the particular item. MSDE informed PSC in writing for each specific administration how many qualifying sets were approved and were available to the Scorers. Scorers had to achieve at least 90% perfect agreement on Step A and 80% perfect agreement on Step B on two of three qualifying sets or one of two qualifying sets, depending on the number of sets available for each item.

Scoring Rules for MSA-Math

The following scoring rules were applied to MSA-Math BCR and ECR items:

- Math BCR (Brief Constructed Response) items were scored:
  Step A: 0, 1 with two readings
  Step B: 0, 1, 2 with two readings
- Math ECR (Extended Constructed Response) items were scored:
  Step A: 0, 1 with two readings
  Step B: 0,1,2,3 with two readings
- Scores given were the higher of the 1st and 2nd Scorer's scores provided they were adjacent.
- For example:

1st Scorer	2nd Scorer	Final Score
1	2	2
2	3	3

- A resolution scorer was used if two non-adjacent initial scores were received.

- The resolution scorer's score was used in place of both the 1st and 2nd Scorers' scores.
- For example:

1st Scorer	2nd Scorer	Resolution Scorer	Final Score
0	2	1	1
0	3	2	2
1	3	3	3
2	0	1	1
3	0	2	2

Inter-Rater Agreement

Pearson's scoring system generated many kinds of internal monitoring reports that enabled the project leadership to monitor the accuracy and consistency of scoring. These reports were compiled by prompt, listed the entire prompt's Scorers, and provided the results of their scoring for each day. Information on these reports included the number of responses read by the Scorers during the period, the number and percent of condition code responses, and the number of responses for which there had been a second reading. The number of responses with second readings provided data that allowed for reporting of the number and percent of responses with perfect agreement; the number and percent of responses on which the first Scorer was a point lower than the second Scorer; the number and percent of responses on which the first Scorer was a point higher than the second Scorer (Adjacent); and the number and percent of responses differing by more than one score point (Non-Adjacent). The Scoring Director also reviewed the daily statistical reports to identify individuals or teams who might need retraining in order to provide continuous scoring consistency on the project. MSDE received data summary reports. Statistical summaries of inter-rater reliability can be found in section 3.4, *Inter-Rater Reliability*.

Scorer Retraining

When a Scorer's performance fell below acceptable parameters for a project, the Scorer was retrained. Retraining was the process by which the SD or Supervisor utilized a number of methods such as individual tutoring on problem score points, individual review of selected responses, and anchor and rubric review to get a Scorer back on track with the guidelines provided by a specific program. Group retraining was conducted by the SD every Monday (or following any extended break) during the scoring project. In addition, daily retraining occurred as deemed necessary by the MSDE representative and CS.

Backreading

Pearson's ePEN system allowed Supervisors and/or SDs to conduct backreads as an additional monitoring method. When conducting backreads, the Supervisor or SD received images of student responses and the scores assigned by the Scorer. Responses selected for backreads might be randomly selected or might be targeted backreads (e.g., responses receiving specific scores, etc.). These backreads were very useful in tracking specific areas of confusion for a given Scorer or group of Scorers and assisted the Supervisor and SD in knowing just how to direct retraining activities for individual Scorers or teams. The initial backreading percentage was set at 3%. This

percentage might be adjusted either higher or lower by the Supervisor based upon the performance of the Scorer.

Development Procedures for Rangefinding

Scoring Directors were selected by the PSC Scoring Resource Manager and Content Specialist to prepare sets of papers for client approval. These experienced SDs were judged by the CS for their ability to recognize and assemble a wide variety of responses. The SD also participated with the clients as a facilitator during the rangefinding session in order to make notes and be prepared to assemble the finished sets to the client's specifications. For a given math prompt, the SD had the following responsibilities:

1) To know the prompt and the rubric thoroughly
2) To read responses
    - Looked for responses that seemed to represent the full range of quality as described in the rubric.
    - Searched all orders for responses, with particular emphasis on the state's high-performing districts.
    - Included not only papers that were homogeneous in their level of quality but also papers that differed in quality from variable to variable but which could be given an overall classification of High, Medium, or Low.
    - Marked High, Medium, and Low papers—marked especially good ones that might potentially receive top scores.
3) To sort copies
    - Copies were sorted into piles, reflecting the nature of the flag—all potential high papers were together, all potential medium papers were together, etc., with all problem papers grouped together.
    - For problem or decision papers, duplicates of types of problems were culled. The best example of each problem type was retained; the rest were set aside for possible future use.
4) To develop sets for rangefinding
    - Decided which particular papers from the sorted piles should go into sets for rangefinding. Each paper selected went into a rangefinding set arranged in performance from low to high performance.

Rangefinding Procedures

The objective of rangefinding sessions was for the team members to arrive at a consensus as to the score of each paper in the proposed training materials. These sessions were attended by Maryland educators, MSDE, and PSC Project Manager, Content Specialists, and Scoring Directors, who selected and prepared all of the papers that would be reviewed. These papers and their corresponding scores formed the basis of selecting final Anchor Sets, Practice Sets, and Qualifying Sets. Discussions among the team members were important, as they revealed what kinds of qualities characterized certain score points. The most difficult aspects involved balancing widely discrepant qualities found in the same paper and defining the line between adjacent scores.

During formal rangefinding, the procedure for assigning scores to the papers in each set was as follows:

- The item was reviewed by the committee and criteria were discussed for receiving full credit.
- Selected "grounding" papers that represented the full range of scores were read aloud and discussed by the rangefinding panel. Reading aloud focused attention on the ideas presented—or what the student had to say—allowing the panel members to divorce themselves from how the paper looked or how well it had been edited.
- After each response was read, each panel member independently assigned a score. An overall tentative score was assigned to each response on which there seemed to be consensus. However, all assigned scores at this point, even those on responses for which there were complete agreement, were provisional and subject to change based on later considerations.
- All subsequent responses were read and scored by each panel member independently, using the tentative scores on the previous sets as guidelines. After each set had been read, the results were recorded on a consensus sheet and discussed after each committee member had already recorded tentative scoring decisions. There might be frequent reference to previous responses to make sure that decisions on score points were consistent.

This iterative process of reading, charting, and discussing successive responses had three results:

- It established scores for papers for which there was virtually unanimous agreement.
- It identified papers that were on the line between two adjacent scores, necessitating the clarification of that line.
- It contributed to understanding the rationale behind scoring decisions.

During this process, the tentative scores assigned to earlier responses became firm.

1.8 The 2010 Operational Item Analyses

Classical Analysis with Common Items Used for Form-to-Form Linking

As mentioned in chapter1.4, two operational forms were randomly distributed to students and linked using common items appearing on both forms (i.e., operational forms A and F). As a result, classical analysis of these common items was conducted to check if the two groups taking different operational forms were equivalent. The following descriptive statistics were calculated based on a raw, number-right score of the common items: mean (*M*) and standard deviation (*SD*). The results indicated that the students taking the two operational forms were statistically close and equivalent across all grades, as seen from Table 1.25.

Table 1.25 Descriptive Statistics for Form-to-Form Common Items

Grade	Form	No. of Items	N	M	SD
3	A	43	29,508	34.86	7.76
	F	43	29,279	35.19	7.54
4	A	39	29,778	29.87	7.24
	F	39	29,797	30.03	7.12
5	A	41	29,532	30.18	7.32
	F	41	28,934	30.44	7.21
6	A	37	29,928	26.32	8.50
	F	37	29,384	26.81	8.36
7	A	36	30,039	22.07	8.69
	F	36	29,140	22.24	8.59
8	A	30	30,331	17.56	6.78
	F	30	29,748	17.79	6.69

Note. Form A designates the identical operational portion of Forms A, B, C, D, and E. Form F designates the identical operational portion of Forms F, G, H, J, and K.
Note. Analysis was conducted with a statewide population.

P-Value Check with Year-to-Year Core Linking Items

As mentioned in chapter 1.4, different years' assessments were linked using core linking items. This section was prepared to provide information about how much p-values (i.e., classical item difficulty) of the 2010 core linking items varied from previous years.

First, only SR items were used for the purpose of year-to-year linking. Second, classical analysis (e.g., p-value) on these items was conducted with a statewide population, and item sequence numbers on the tables were assigned based on the 2010 assessment. Finally it should be noted that detailed information about Rasch analysis on these core linking items can be found in chapter 1.9, *Linking, Equating, Scaling Procedures of the 2010 MSA-Math*.

As seen in Tables 1.26 through 1.37, we could conclude that most of the 2010 p-values were almost the same or slightly increased compared to those of previous years across all grades.

Table 1.26 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 3 Form A

Item Seq. No.	Item CID	Previous Year	Y10 FA	Item Seq. No.	Item CID	Previous Year	Y10 FA
1	3509931	0.66	0.70	48	3510065	0.96	0.95
2	3548059	0.71	0.78	49	3509961	0.92	0.93
5	3510009	0.79	0.84	51	3510018	0.76	0.82
6	100000044159	0.57	0.61	52	3510035	0.87	0.90
7	3548057	0.73	0.82	56	3510058	0.87	0.90
13	100000366020	0.59	0.68	62	3510347	0.68	0.75
14	3510017	0.91	0.94	63	3510053	0.83	0.85
15	3509926	0.40	0.49	64	100000044162	0.83	0.85
16	3509960	0.76	0.82	65	3496700	0.87	0.84
17	100000044152	0.86	0.89	66	3510036	0.85	0.87
18	100000025225	0.85	0.85	67	3510329	0.55	0.67
19	3510022	0.47	0.57	68	3510033	0.79	0.83
20	3509983	0.91	0.93	69	3510043	0.80	0.78
23	3509927	0.75	0.82	71	100000025211	0.78	0.78
31	3548507	0.85	0.90	72	3509950	0.72	0.78
32	3509988	0.73	0.75	81	3510176	0.67	0.69
33	3488123	0.60	0.56	82	3509929	0.52	0.57
41	3510063	0.73	0.80				

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 3 Form A

Form	Year	No. of Items	M	SD
A	Previous Year	35	0.75	0.14
	2010	35	0.79	0.12

Table 1.27 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 3 Form F

Item Seq. No.	Item CID	Previous Year	Y10 FF	Item Seq. No.	Item CID	Previous Year	Y10 FF
1	3509931	0.66	0.71	48	3510065	0.96	0.95
2	3548059	0.71	0.78	49	3509961	0.92	0.92
5	3510009	0.79	0.85	51	3510018	0.76	0.83
6	100000044159	0.57	0.62	52	3510035	0.87	0.90
7	3548057	0.73	0.83	56	3510058	0.87	0.92
13	100000366020	0.59	0.69	62	3510347	0.68	0.75
14	3510017	0.91	0.95	63	3510053	0.83	0.85
15	3509926	0.40	0.52	64	100000044162	0.83	0.86
16	3509960	0.76	0.82	65	3496700	0.87	0.84
17	100000044152	0.86	0.91	66	3510036	0.85	0.88
18	100000025225	0.85	0.86	67	3510329	0.55	0.67
19	3510022	0.47	0.57	68	3510033	0.79	0.84
20	3509983	0.91	0.93	69	3510043	0.80	0.77
23	3509927	0.75	0.82	71	100000025211	0.78	0.82
31	3548507	0.85	0.90	72	3509950	0.72	0.79
32	3509988	0.73	0.76	81	3510176	0.67	0.69
33	3488123	0.60	0.60	82	3509929	0.52	0.58
41	3510063	0.73	0.82				

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 3 Form F

Form	Year	No. of Items	M	SD
F	Previous Year	35	0.75	0.14
	2010	35	0.79	0.12

Table 1.28 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 4 Form A

Item Seq. No.	Item CID	Previous Year	Y10 FA	Item Seq. No.	Item CID	Previous Year	Y10 FA
2	3515407	0.81	0.87	49	3515471	0.83	0.86
3	3515643	0.36	0.49	50	3515630	0.51	0.57
6	3515408	0.66	0.81	54	3515635	0.56	0.62
7	3515641	0.81	0.83	55	3515631	0.76	0.79
8	3515410	0.84	0.88	62	100000201857	0.48	0.50
10	3515605	0.50	0.64	63	3515634	0.75	0.79
18	3488166	0.79	0.82	64	3515853	0.71	0.83
19	3515447	0.41	0.55	66	3548078	0.50	0.49
22	3515604	0.60	0.71	67	3515933	0.76	0.82
23	3515737	0.83	0.83	68	3548079	0.94	0.96
24	3488190	0.56	0.52	69	100000201852	0.85	0.85
25	3515470	0.66	0.75	70	100000011489	0.97	0.97
26	3490562	0.53	0.54	71	3515592	0.82	0.87
30	100000018336	0.79	0.80	78	3515506	0.86	0.91
32	3551599	0.82	0.87	80	3515632	0.69	0.70
33	100000044144	0.94	0.96	81	3548088	0.74	0.76
47	3515575	0.71	0.89				

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 4 Form A

Form	Year	No. of Items	M	SD
A	Previous Year	33	0.71	0.16
	2010	33	0.76	0.15

Table 1.29 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 4 Form F

Item Seq. No.	Item CID	Previous Year	Y10 FF	Item Seq. No.	Item CID	Previous Year	Y10 FF
2	3515407	0.81	0.88	49	3515471	0.83	0.86
3	3515643	0.36	0.50	50	3515630	0.51	0.57
6	3515408	0.66	0.81	54	3515635	0.56	0.61
7	3515641	0.81	0.83	55	3515631	0.76	0.79
8	3515410	0.84	0.89	62	100000201857	0.48	0.48
10	3515605	0.50	0.65	63	3515634	0.75	0.81
18	3488166	0.79	0.81	64	3515853	0.71	0.83
19	3515447	0.41	0.56	66	3548078	0.50	0.51
22	3515604	0.60	0.72	67	3515933	0.76	0.81
23	3515737	0.83	0.83	68	3548079	0.94	0.96
24	3488190	0.56	0.52	69	100000201852	0.85	0.86
25	3515470	0.66	0.75	70	100000011489	0.97	0.97
26	3490562	0.53	0.56	71	3515592	0.82	0.88
30	100000018336	0.79	0.80	78	3515506	0.86	0.91
32	3551599	0.82	0.87	80	3515632	0.69	0.71
33	100000044144	0.94	0.97	81	3548088	0.74	0.78
47	3515575	0.71	0.89				

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 4 Form F

Form	Year	No. of Items	M	SD
F	Previous Year	33	0.71	0.16
	2010	33	0.76	0.15

Table 1.30 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 5 Form A

Item Seq. No.	Item CID	Previous Year	Y10 FA	Item Seq. No.	Item CID	Previous Year	Y10 FA
2	3511269	0.83	0.88	43	3511513	0.85	0.85
8	3511203	0.86	0.94	47	3488391	0.89	0.88
16	3511196	0.56	0.57	48	3488431	0.74	0.75
18	3488373	0.66	0.71	49	3492140	0.91	0.90
19	3511467	0.81	0.83	55	3512595	0.78	0.78
20	3512529	0.56	0.60	56	3488509	0.82	0.77
21	3512606	0.63	0.65	58	100000043857	0.82	0.85
23	100000043853	0.67	0.70	60	3512712	0.91	0.91
26	3492117	0.96	0.95	61	3511429	0.75	0.77
27	3512638	0.64	0.72	64	3511626	0.81	0.87
28	3492126	0.88	0.88	70	3488326	0.68	0.66
37	100000366318	0.75	0.87	71	3488251	0.61	0.65
38	3511246	0.73	0.79	72	3512691	0.52	0.58
39	3488443	0.43	0.49	82	3488328	0.71	0.73
42	3511566	0.66	0.68	83	3511448	0.76	0.79

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 5 Form A

Form	Year	No. of Items	M	SD
A	Previous Year	30	0.74	0.13
	2010	30	0.77	0.12

Table 1.31 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 5 Form F

Item Seq. No.	Item CID	Previous Year	Y10 FF	Item Seq. No.	Item CID	Previous Year	Y10 FF
2	3511269	0.83	0.90	43	3511513	0.85	0.86
8	3511203	0.86	0.94	47	3488391	0.89	0.89
16	3511196	0.56	0.59	48	3488431	0.74	0.77
18	3488373	0.66	0.71	49	3492140	0.91	0.91
19	3511467	0.81	0.83	55	3512595	0.78	0.78
20	3512529	0.56	0.59	56	3488509	0.82	0.80
21	3512606	0.63	0.66	58	100000043857	0.82	0.86
23	100000043853	0.67	0.71	60	3512712	0.91	0.92
26	3492117	0.96	0.95	61	3511429	0.75	0.78
27	3512638	0.64	0.73	64	3511626	0.81	0.89
28	3492126	0.88	0.89	70	3488326	0.68	0.68
37	100000366318	0.75	0.88	71	3488251	0.61	0.64
38	3511246	0.73	0.80	72	3512691	0.52	0.57
39	3488443	0.43	0.47	82	3488328	0.71	0.71
42	3511566	0.66	0.69	83	3511448	0.76	0.78

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 5 Form F

Form	Year	No. of Items	M	SD
F	Previous Year	30	0.74	0.13
	2010	30	0.77	0.12

Table 1.32 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 6 Form A

Item Seq. No.	Item CID	Previous Year	Y10 FA	Item Seq. No.	Item CID	Previous Year	Y10 FA
1	3516257	0.79	0.89	35	3516241	0.82	0.86
3	3516291	0.46	0.52	36	3516247	0.52	0.63
4	100000022483	0.55	0.57	37	3516329	0.50	0.61
6	3516243	0.64	0.75	38	3516623	0.72	0.75
9	100000078832	0.55	0.81	45	3492095	0.80	0.82
10	3516559	0.84	0.91	50	3516565	0.42	0.52
11	3516255	0.70	0.77	51	3488358	0.70	0.68
12	3516258	0.53	0.64	54	3516906	0.60	0.61
19	3516240	0.55	0.68	55	3516332	0.48	0.52
20	3516909	0.59	0.60	56	3516301	0.60	0.77
25	3516351	0.51	0.51	57	100000022470	0.54	0.52
26	3516290	0.60	0.70	58	3488489	0.73	0.74
27	100000043862	0.61	0.64	61	3492120	0.65	0.65
30	3492143	0.77	0.77	68	3516613	0.51	0.53
34	3516331	0.38	0.52	80	3516303	0.47	0.54

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 6 Form A

Form	Year	No. of Items	M	SD
A	Previous Year	30	0.60	0.12
	2010	30	0.67	0.12

Table 1.33 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 6 Form F

Item Seq. No.	Item CID	Previous Year	Y10 FF	Item Seq. No.	Item CID	Previous Year	Y10 FF
1	3516257	0.79	0.89	35	3516241	0.82	0.87
3	3516291	0.46	0.54	36	3516247	0.52	0.66
4	100000022483	0.55	0.57	37	3516329	0.50	0.67
6	3516243	0.64	0.76	38	3516623	0.72	0.77
9	100000078832	0.55	0.82	45	3492095	0.80	0.80
10	3516559	0.84	0.91	50	3516565	0.42	0.51
11	3516255	0.70	0.79	51	3488358	0.70	0.69
12	3516258	0.53	0.66	54	3516906	0.60	0.59
19	3516240	0.55	0.71	55	3516332	0.48	0.54
20	3516909	0.59	0.61	56	3516301	0.60	0.80
25	3516351	0.51	0.53	57	100000022470	0.54	0.54
26	3516290	0.60	0.71	58	3488489	0.73	0.75
27	100000043862	0.61	0.66	61	3492120	0.65	0.64
30	3492143	0.77	0.78	68	3516613	0.51	0.56
34	3516331	0.38	0.54	80	3516303	0.47	0.55

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 6 Form F

Form	Year	No. of Items	M	SD
F	Previous Year	30	0.60	0.12
	2010	30	0.68	0.12

Table 1.34 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 7 Form A

Item Seq. No.	Item CID	Previous Year	Y10 FA	Item Seq. No.	Item CID	Previous Year	Y10 FA
1	3517604	0.34	0.34	49	100000012796	0.58	0.49
2	3517601	0.44	0.54	50	3517602	0.39	0.50
3	3517678	0.88	0.95	51	3517687	0.56	0.58
7	3517616	0.55	0.65	52	3517692	0.77	0.84
8	3517634	0.61	0.69	63	3517712	0.41	0.51
10	3517677	0.70	0.66	64	3517714	0.50	0.60
12	100000026796	0.85	0.84	65	3517885	0.35	0.48
18	3517652	0.63	0.70	66	100000018106	0.61	0.61
19	3547473	0.77	0.83	69	3517721	0.41	0.51
20	3517739	0.85	0.86	70	3517691	0.61	0.67
30	3517639	0.28	0.35	72	3555858	0.39	0.49
31	3517665	0.35	0.41	79	3555859	0.74	0.78
32	3517742	0.50	0.65	80	100000363463	0.44	0.61
43	3517656	0.59	0.67	81	3488830	0.58	0.54

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 7 Form A

Form	Year	No. of Items	M	SD
A	Previous Year	28	0.56	0.17
	2010	28	0.62	0.16

Table 1.35 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 7 Form F

Item Seq. No.	Item CID	Previous Year	Y10 FF	Item Seq. No.	Item CID	Previous Year	Y10 FF
1	3517604	0.34	0.33	49	100000012796	0.58	0.48
2	3517601	0.44	0.55	50	3517602	0.39	0.51
3	3517678	0.88	0.96	51	3517687	0.56	0.60
7	3517616	0.55	0.67	52	3517692	0.77	0.84
8	3517634	0.61	0.70	63	3517712	0.41	0.49
10	3517677	0.70	0.67	64	3517714	0.50	0.58
12	100000026796	0.85	0.84	65	3517885	0.35	0.46
18	3517652	0.63	0.74	66	100000018106	0.61	0.61
19	3547473	0.77	0.84	69	3517721	0.41	0.53
20	3517739	0.85	0.86	70	3517691	0.61	0.66
30	3517639	0.28	0.34	72	3555858	0.39	0.48
31	3517665	0.35	0.40	79	3555859	0.74	0.79
32	3517742	0.50	0.64	80	100000363463	0.44	0.60
43	3517656	0.59	0.67	81	3488830	0.58	0.55

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 7 Form F

Form	Year	No. of Items	M	SD
F	Previous Year	28	0.56	0.17
	2010	28	0.62	0.16

Table 1.36 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 8 Form A

Item Seq. No.	Item CID	Previous Year	Y10 FA	Item Seq. No.	Item CID	Previous Year	Y10 FA
1	3514015	0.23	0.26	47	3487901	0.85	0.86
2	3514014	0.52	0.57	48	3514056	0.74	0.82
5	100000018156	0.68	0.70	50	3487525	0.50	0.53
7	3514053	0.70	0.77	51	100000018153	0.66	0.63
8	100000043330	0.45	0.47	52	3514074	0.42	0.43
14	3500150	0.47	0.51	53	3514075	0.60	0.66
22	3514595	0.68	0.71	62	3514095	0.27	0.33
27	100000043320	0.47	0.48	65	3487902	0.82	0.85
32	3500154	0.73	0.75	66	100000018151	0.56	0.58
33	3514062	0.39	0.44	72	3487712	0.63	0.58
38	100000049037	0.67	0.66	73	3492047	0.34	0.27
41	100000043323	0.36	0.49	78	3487912	0.53	0.54
42	3514291	0.73	0.82	79	3547536	0.49	0.52
46	3514055	0.53	0.59				

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 8 Form A

Form	Year	No. of Items	M	SD
A	Previous Year	27	0.56	0.16
	2010	27	0.59	0.17

Table 1.37 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2010: Grade 8 Form F

Item Seq. No.	Item CID	Previous Year	Y10 FF	Item Seq. No.	Item CID	Previous Year	Y10 FF
1	3514015	0.23	0.27	47	3487901	0.85	0.87
2	3514014	0.52	0.58	48	3514056	0.74	0.85
5	100000018156	0.68	0.71	50	3487525	0.50	0.55
7	3514053	0.70	0.78	51	100000018153	0.66	0.63
8	100000043330	0.45	0.49	52	3514074	0.42	0.45
14	3500150	0.47	0.50	53	3514075	0.60	0.67
22	3514595	0.68	0.73	62	3514095	0.27	0.36
27	100000043320	0.47	0.45	65	3487902	0.82	0.86
32	3500154	0.73	0.74	66	100000018151	0.56	0.55
33	3514062	0.39	0.47	72	3487712	0.63	0.61
38	100000049037	0.67	0.65	73	3492047	0.34	0.29
41	100000043323	0.36	0.52	78	3487912	0.53	0.56
42	3514291	0.73	0.80	79	3547536	0.49	0.51
46	3514055	0.53	0.61				

Note. Analysis was conducted with a statewide population.
Note. Item sequence numbers were assigned based on the 2010 assessment.

Descriptive Statistics for Year-Year Linking Common Items: Grade 8 Form F

Form	Year	No. of Items	M	SD
F	Previous Year	27	0.56	0.16
	2010	27	0.59	0.16

Validation Check with the 2010 MSA-Math Core Items

As mentioned in chapter 1.4, operational items fell into one of two categories: core and core linking items. Because the core items were not included into the 2010 year-to-year linking pool, Rasch item and step difficulty parameters of the core items were reestimated with the 2010 stratified random samples during calibration and equating. (Please see section 1.9 and Appendix A for stratified random sampling procedures) As a result, this section was prepared to provide detailed information about how much the core items changed in terms of item difficulty, both classical item p-value and Rasch item difficulty. Detailed information about the roles of the 2010 core and core linking items can be found in section 1.4, *Test Form Design, Specifications, Item Type, and Item Roles*.

First of all, it should be noted that a smaller number of cases (i.e., about 2,500) in the table indicates that it is a field-test item. P-values of both BCR and ECR items were calculated by dividing the item mean score by the item score range (i.e., score point 2 for BCR and 3 for ECR). The percentage of "Omits" for each CR item was low and indicated that a small number of students did not respond at all. In general, item p-value analysis results indicated that most of the 2010 p-values were almost the same or somewhat increased compared to those in previous years across all grades.

With respect to the Rasch item calibration and equating, it should be noted that we coded "Omit" of each item as "missing" before we ran the data with the Rasch model. In general, the level of the 2010 item difficulties stayed almost the same or became a little lower compared to that of previous years across all grades. It should be noted that all of the Rasch item and step difficulty parameters were on a common scale (i.e., linked to the 2006 assessment).

In conclusion, both p-value and Rasch item difficulty results reflected the same phenomenon, indicating that the level of item difficulty stayed the same or became a little lower.

Table 1.38 P-Value Comparison of Core Items for Previous Year vs. Year 2010: Grade 3 Form A

Item CID	Previous Year	Y10 FA	Item CID	Previous Year	Y10 FA
100000011186	**0.75**	**0.73**	**3564080**	**0.64**	**0.61**
3595529	**0.61**	**0.62**	3519712	0.94	0.93
100000232540	0.80	0.84	100000067839	0.84	0.85
3509941	**0.61**	**0.64**	3488003	0.45	0.45
3595501	**0.57**	**0.55**	100000004262	0.71	0.72
100000063515	0.88	0.87	100000063549	0.59	0.65
3519708	0.37	0.47	**100000011176**	**0.88**	**0.91**
100000025196	**0.90**	**0.90**	**3595524**	**0.64**	**0.64**
3595519	**0.65**	**0.67**	100000025212	0.86	0.89
100000044158	0.88	0.86	100000067773	0.91	0.93
100000067832	0.52	0.57	**100000004277**	**0.81**	**0.82**
100000063513	0.84	0.83	**3595518**	**0.58**	**0.56**
3488139	**0.49**	**0.48**	100000011208	0.84	0.84
3564095	**0.47**	**0.47**	100000067837	0.68	0.65
3510072	**0.88**	**0.85**	100000067778	0.94	0.92

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item.

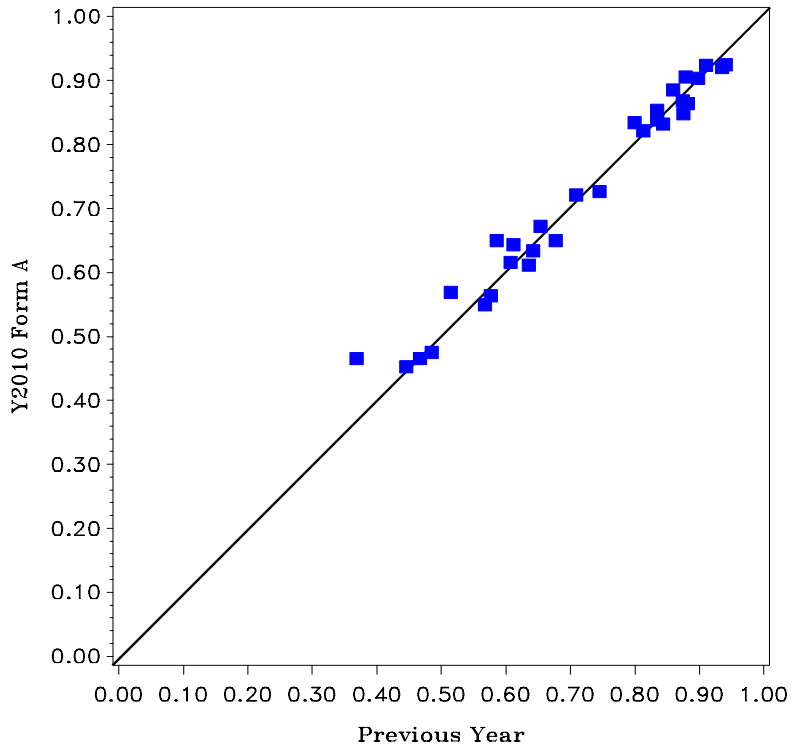

Table 1.39 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 3 Form A

						Score-Point Distribution (%)			
Year	Item CID	Item Type	N	Mean	SD	0	1	2	Omit
2009	100000011186	BCR	10,712	0.75	0.44	25.08	74.62		0.30
2009	3595529	BCR	10,712	1.22	0.51	4.28	68.84	26.46	0.42
2009	3509941	BCR	8,808	0.61	0.49	38.26	61.32		0.42
2009	3595501	BCR	8,808	1.14	0.58	10.23	64.23	24.73	0.82
2009	100000025196	BCR	8,808	0.90	0.30	9.79	89.84		0.37
2009	3595519	BCR	8,808	1.31	0.73	15.44	36.83	46.99	0.74
2009	3488139	BCR	8,808	0.49	0.50	50.89	48.63		0.49
2009	3564095	BCR	8,808	0.94	0.58	19.24	66.33	13.59	0.84
2009	3510072	BCR	8,808	0.88	0.33	12.10	87.64		0.26
2009	3564080	BCR	8,808	1.27	0.73	15.78	39.62	43.82	0.77
2008	100000011176	BCR	2,555	0.88	0.33	10.37	87.95		1.68
2008	3595524	BCR	2,555	1.29	0.60	6.07	55.66	36.52	1.76
2008	100000004277	BCR	2,607	0.81	0.39	17.84	81.40		0.77
2008	3595518	BCR	2,607	1.16	0.76	20.64	40.62	37.44	1.30
2010	100000011186	BCR	29,508	0.73	0.45	26.77	72.80		0.41
2010	3595529	BCR	29,508	1.23	0.53	4.38	66.10	28.64	0.77
2010	3509941	BCR	29,508	0.64	0.48	34.82	64.39		0.73
2010	3595501	BCR	29,508	1.10	0.58	11.07	65.23	22.45	1.14
2010	100000025196	BCR	29,508	0.90	0.29	8.77	90.45		0.75
2010	3595519	BCR	29,508	1.35	0.74	14.78	33.27	50.65	1.15
2010	3488139	BCR	29,508	0.48	0.50	51.75	47.59		0.61
2010	3564095	BCR	29,508	0.93	0.65	23.54	57.19	18.01	1.11
2010	3510072	BCR	29,508	0.85	0.36	14.39	84.95		0.63
2010	3564080	BCR	29,508	1.23	0.75	18.15	38.72	41.90	1.13
2010	100000011176	BCR	29,508	0.91	0.29	7.35	90.67		1.90
2010	3595524	BCR	29,508	1.27	0.52	2.69	65.82	30.60	0.59
2010	100000004277	BCR	29,508	0.82	0.38	16.89	82.25		0.80
2010	3595518	BCR	29,508	1.13	0.77	22.76	38.75	37.08	1.17

Table 1.40 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 3 Form A

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2009	3	100000011186	BCR_A	0.7532		
2009	4	3595529	BCR_B	1.0151	-2.4072	2.4072
2009	8	100000232540	SR	0.2621		
2009	21	3509941	BCR_A	1.4160		
2009	22	3595501	BCR_B	1.5639	-1.9549	1.9549
2009	24	100000063515	SR	-0.3353		
2007	25	3519708	SR	2.4481		
2009	26	100000025196	BCR_A	-0.7697		
2009	27	3595519	BCR_B	1.2125	-0.9262	0.9262
2009	28	100000044158	SR	-0.4990		
2009	29	100000067832	SR	1.9896		
2009	30	100000063513	SR	-0.0929		
2009	36	3488139	BCR_A	2.1303		
2009	37	3564095	BCR_B	2.4239	-2.0904	2.0904
2009	42	3510072	BCR_A	-0.3424		
2009	43	3564080	BCR_B	1.4331	-0.8085	0.8085
2008	44	3519712	SR	-1.3169		
2009	45	100000067839	SR	-0.0401		
2007	46	3488003	SR	2.1071		
2008	47	100000004262	SR	0.9094		
2009	50	100000063549	SR	1.6354		
2008	53	100000011176	BCR_A	-0.5748		
2008	54	3595524	BCR_B	0.9228	-1.7984	1.7984
2008	55	100000025212	SR	-0.3110		
2009	70	100000067773	SR	-0.9648		
2008	73	100000004277	BCR_A	0.0371		
2008	74	3595518	BCR_B	1.4885	-0.8352	0.8352
2008	75	100000011208	SR	-0.1032		
2009	76	100000067837	SR	1.1077		
2009	80	100000067778	SR	-1.2832		
2010	3	100000011186	BCR_A	0.8552		
2010	4	3595529	BCR_B	0.8995	-2.3179	2.3179
2010	8	100000232540	SR	0.0514		
2010	21	3509941	BCR_A	1.3031		
2010	22	3595501	BCR_B	1.6084	-2.0125	2.0125
2010	24	100000063515	SR	-0.2985		
2010	25	3519708	SR	2.2939		
2010	26	100000025196	BCR_A	-0.7483		

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2010	27	3595519	BCR_B	1.2431	-0.6761	0.6761
2010	28	100000044158	SR	-0.2871		
2010	29	100000067832	SR	1.8287		
2010	30	100000063513	SR	0.1375		
2010	36	3488139	BCR_A	2.2833		
2010	37	3564095	BCR_B	2.3841	-1.4396	1.4396
2010	42	3510072	BCR_A	-0.0411		
2010	43	3564080	BCR_B	1.5221	-0.8553	0.8553
2010	44	3519712	SR	-1.0473		
2010	45	100000067839	SR	-0.1399		
2010	46	3488003	SR	2.4137		
2010	47	100000004262	SR	0.8784		
2010	50	100000063549	SR	1.3353		
2010	53	100000011176	BCR_A	-0.8382		
2010	54	3595524	BCR_B	0.7295	-2.4568	2.4568
2010	55	100000025212	SR	-0.4987		
2010	70	100000067773	SR	-1.2267		
2010	73	100000004277	BCR_A	0.1937		
2010	74	3595518	BCR_B	1.8092	-0.7713	0.7713
2010	75	100000011208	SR	0.0471		
2010	76	100000067837	SR	1.3024		
2010	80	100000067778	SR	-1.1288		

Note. These Rasch difficulties were based on a common scale.

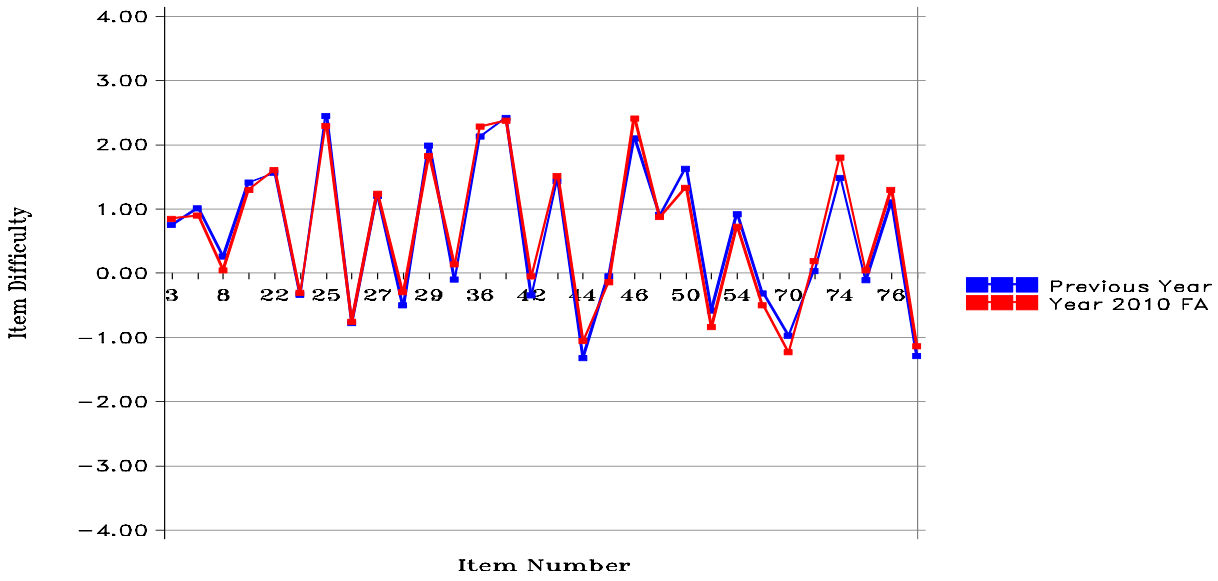

Figure 1.3 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 3 Form A

Table 1.41 P-Value Comparisons of Core Items for Previous Year vs. Year 2010: Grade 3 Form F

Item CID	Previous Year	Y10 FF	Item CID	Previous Year	Y10 FF
100000011186	**0.75**	**0.74**	**3564099**	**0.40**	**0.40**
3595529	**0.61**	**0.62**	100000025234	0.82	0.82
100000232428	0.95	0.96	100000067806	0.93	0.93
3509941	**0.61**	**0.65**	100000004264	0.70	0.69
3595501	**0.57**	**0.56**	100000025204	0.89	0.89
100000025201	0.77	0.80	100000004268	0.47	0.53
100000367550	0.77	0.75	**3509949**	**0.77**	**0.76**
100000025196	**0.90**	**0.91**	**3985609**	**0.69**	**0.70**
3595519	**0.65**	**0.67**	3487800	0.82	0.90
3511729	0.69	0.72	100000067777	0.91	0.93
100000067864	0.84	0.88	**3509932**	**0.98**	**0.98**
3497892	0.73	0.70	**3564086**	**0.47**	**0.44**
3488139	**0.49**	**0.47**	100000011210	0.87	0.87
3564095	**0.47**	**0.46**	100000018400	0.69	0.67
3488087	**0.44**	**0.43**	100000067804	0.86	0.85

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item.

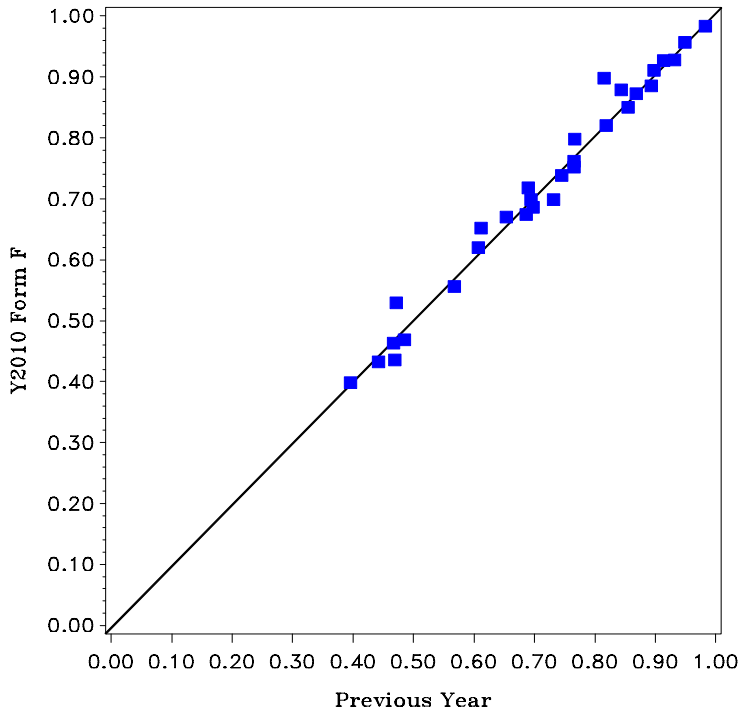

Table 1.42 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 3 Form F

						Score-Point Distribution (%)			
Year	Item CID	Item Type	N	Mean	SD	0	1	2	Omit
2009	100000011186	BCR	10,712	0.75	0.44	25.08	74.62		0.30
2009	3595529	BCR	10,712	1.22	0.51	4.28	68.84	26.46	0.42
2009	3509941	BCR	8,808	0.61	0.49	38.26	61.32		0.42
2009	3595501	BCR	8,808	1.14	0.58	10.23	64.23	24.73	0.82
2009	100000025196	BCR	8,808	0.90	0.30	9.79	89.84		0.37
2009	3595519	BCR	8,808	1.31	0.73	15.44	36.83	46.99	0.74
2009	3488139	BCR	8,808	0.49	0.50	50.89	48.63		0.49
2009	3564095	BCR	8,808	0.94	0.58	19.24	66.33	13.59	0.84
2009	3488087	BCR	10,712	0.44	0.50	54.85	44.35		0.79
2009	3564099	BCR	10,712	0.79	0.65	31.77	54.14	12.64	1.46
2009	3509949	BCR	8,808	0.77	0.42	22.23	76.70		1.07
2009	3985609	BCR	8,808	1.39	0.69	11.32	36.92	51.02	0.74
2009	3509932	BCR	10,712	0.98	0.13	1.29	98.38		0.34
2009	3564086	BCR	10,712	0.94	0.64	22.85	58.98	17.56	0.61
2010	100000011186	BCR	29,279	0.74	0.44	25.77	73.92		0.30
2010	3595529	BCR	29,279	1.24	0.53	4.39	65.77	29.22	0.55
2010	3509941	BCR	29,279	0.65	0.48	33.97	65.34		0.66
2010	3595501	BCR	29,279	1.11	0.58	10.49	65.28	23.07	1.08
2010	100000025196	BCR	29,279	0.91	0.28	8.21	91.17		0.57
2010	3595519	BCR	29,279	1.34	0.74	14.74	34.04	50.14	0.88
2010	3488139	BCR	29,279	0.47	0.50	52.35	46.94		0.62
2010	3564095	BCR	29,279	0.93	0.65	23.81	57.11	17.85	1.09
2010	3488087	BCR	29,279	0.43	0.50	55.41	43.36		1.20
2010	3564099	BCR	29,279	0.80	0.65	31.39	53.22	13.31	1.89
2010	3509949	BCR	29,279	0.76	0.43	22.38	76.26		1.33
2010	3985609	BCR	29,279	1.40	0.68	10.67	37.05	51.51	0.60
2010	3509932	BCR	29,279	0.98	0.12	0.94	98.47		0.58
2010	3564086	BCR	29,279	0.87	0.63	25.77	58.98	14.13	0.99

Table 1.43 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 3 Form F

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2009	3	100000011186	BCR_A	0.7532		
2009	4	3595529	BCR_B	1.0151	-2.4072	2.4072
2009	8	100000232428	SR	-1.5988		
2009	21	3509941	BCR_A	1.4160		
2009	22	3595501	BCR_B	1.5639	-1.9549	1.9549
2008	24	100000025201	SR	0.5320		
2006	25	100000367550	SR	0.3600		
2009	26	100000025196	BCR_A	-0.7697		
2009	27	3595519	BCR_B	1.2125	-0.9262	0.9262
2009	28	3511729	SR	1.0809		
2009	29	100000067864	SR	-0.0926		
2008	30	3497892	SR	0.6640		
2009	36	3488139	BCR_A	2.1303		
2009	37	3564095	BCR_B	2.4239	-2.0904	2.0904
2009	42	3488087	BCR_A	2.4484		
2009	43	3564099	BCR_B	2.8035	-1.5459	1.5459
2008	44	100000025234	SR	0.0506		
2009	45	100000067806	SR	-1.2298		
2008	46	100000004264	SR	0.8815		
2008	47	100000025204	SR	-0.5741		
2008	50	100000004268	SR	2.2412		
2009	53	3509949	BCR_A	0.5146		
2009	54	3985609	BCR_B	0.9641	-0.9820	0.9820
2007	55	3487800	SR	-0.2004		
2009	70	100000067777	SR	-0.9625		
2009	73	3509932	BCR_A	-2.8170		
2009	74	3564086	BCR_B	2.3423	-1.5697	1.5697
2008	75	100000011210	SR	-0.4611		
2008	76	100000018400	SR	0.9724		
2009	80	100000067804	SR	-0.3530		
2010	3	100000011186	BCR_A	0.8552		
2010	4	3595529	BCR_B	0.8995	-2.3179	2.3179
2010	8	100000232428	SR	-1.7247		
2010	21	3509941	BCR_A	1.3031		
2010	22	3595501	BCR_B	1.6084	-2.0125	2.0125
2010	24	100000025201	SR	0.4111		
2010	25	100000367550	SR	0.7454		
2010	26	100000025196	BCR_A	-0.7483		
2010	27	3595519	BCR_B	1.2431	-0.6761	0.6761

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2010	28	3511729	SR	0.9775		
2010	29	100000067864	SR	-0.4308		
2010	30	3497892	SR	1.1154		
2010	36	3488139	BCR_A	2.2833		
2010	37	3564095	BCR_B	2.3841	-1.4396	1.4396
2010	42	3488087	BCR_A	2.5974		
2010	43	3564099	BCR_B	2.8239	-1.4658	1.4658
2010	44	100000025234	SR	0.1035		
2010	45	100000067806	SR	-1.1593		
2010	46	100000004264	SR	1.1323		
2010	47	100000025204	SR	-0.5377		
2010	50	100000004268	SR	2.0124		
2010	53	3509949	BCR_A	0.7528		
2010	54	3985609	BCR_B	1.1340	-0.7901	0.7901
2010	55	3487800	SR	-0.6059		
2010	70	100000067777	SR	-1.0684		
2010	73	3509932	BCR_A	-3.0569		
2010	74	3564086	BCR_B	2.6464	-1.6647	1.6647
2010	75	100000011210	SR	-0.2484		
2010	76	100000018400	SR	1.1695		
2010	80	100000067804	SR	-0.2424		

Note. These Rasch difficulties were based on a common scale.

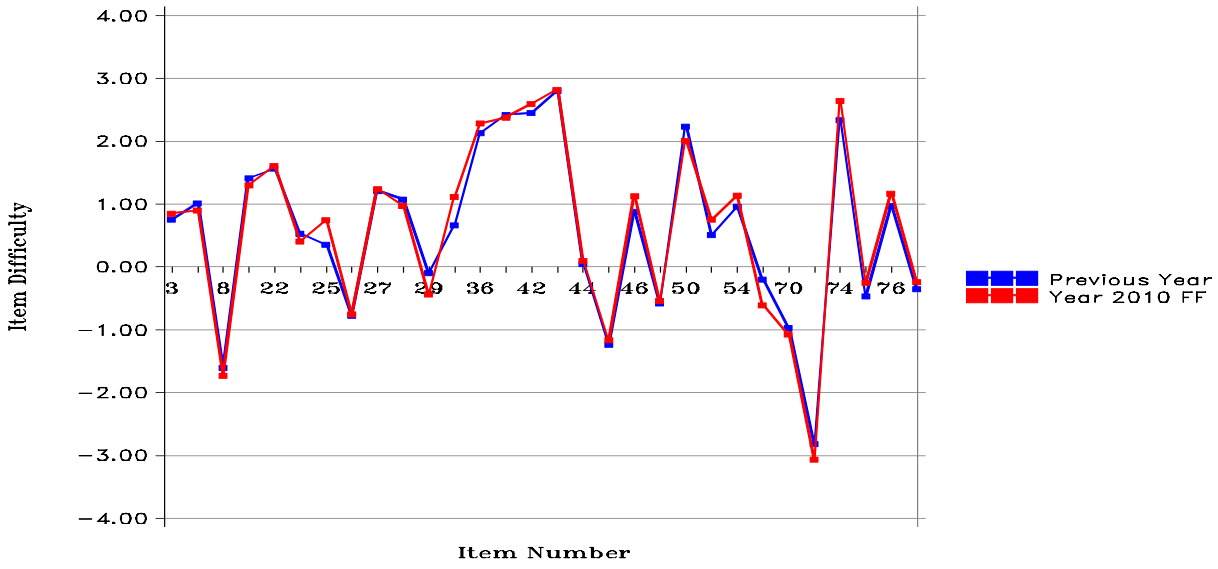

Figure 1.4 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 3 Form F

Table 1.44 P-Value Comparison of Core Items for Previous Year vs. Year 2010: Grade 4 Form A

Item CID	Previous Year	Y10 FA	Item CID	Previous Year	Y10 FA
3497865	0.67	0.70	**3595532**	**0.39**	**0.39**
3487819	**0.50**	**0.53**	**3488150**	**0.38**	**0.38**
3564186	**0.44**	**0.42**	**3564176**	**0.48**	**0.47**
100000069423	0.52	0.56	100000007114	0.95	0.94
100000069397	0.80	0.79	100000069424	0.89	0.90
100000044142	**0.83**	**0.84**	**3488145**	**0.71**	**0.70**
3595499	**0.47**	**0.47**	**3564189**	**0.48**	**0.49**
3487979	0.69	0.68	3497871	0.69	0.76
100000025175	**0.88**	**0.89**	100000252124	0.61	0.68
3595574	**0.45**	**0.41**	100000069399	0.72	0.79
100000069344	0.84	0.85	100000011499	0.91	0.90
100000069324	0.85	0.88	**3515783**	**0.75**	**0.74**
100000025187	0.80	0.83	**3595560**	**0.73**	**0.71**
100000025153	0.98	0.98	100000252111	0.64	0.78
100000069420	0.79	0.87	100000011497	0.68	0.74
3515823	**0.45**	**0.45**			

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item.

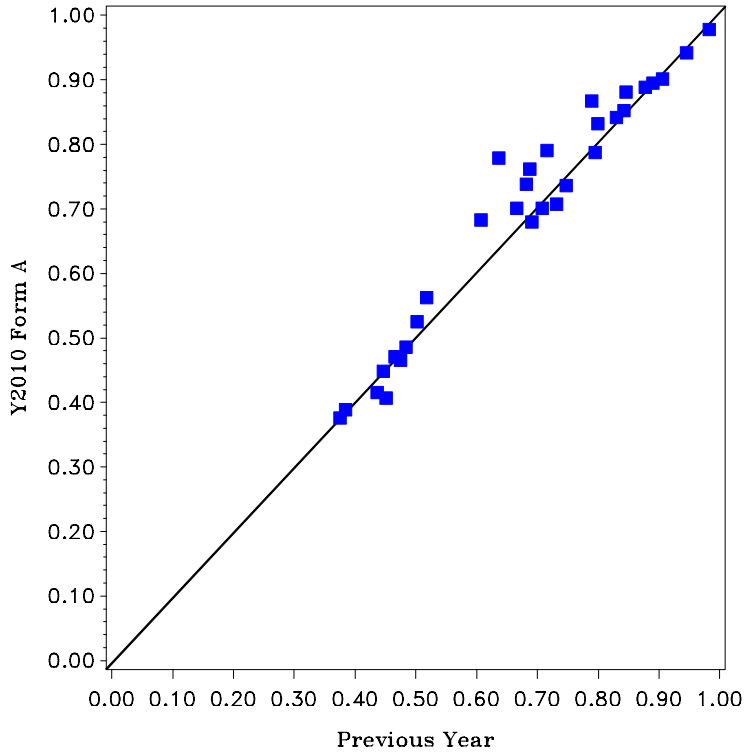

Table 1.45 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 4 Form A

						Score-Point Distribution (%)			
Year	Item CID	Item Type	N	Mean	SD	0	1	2	Omit
2009	3487819	BCR	15,855	0.50	0.50	49.12	50.26		0.62
2009	3564186	BCR	15,855	0.88	0.48	17.50	75.69	5.94	0.88
2009	100000044142	BCR	15,855	0.83	0.38	16.27	83.08		0.66
2009	3595499	BCR	15,855	0.93	0.55	17.54	69.52	11.91	1.02
2008	100000025175	BCR	2,255	0.88	0.33	11.53	87.89		0.58
2008	3595574	BCR	2,255	0.90	0.49	16.32	75.08	7.67	0.93
2009	3515823	BCR	15,855	0.45	0.50	54.42	44.71		0.88
2009	3595532	BCR	15,855	0.77	0.59	29.98	60.03	8.49	1.49
2009	3488150	BCR	15,855	0.38	0.48	61.71	37.58		0.71
2009	3564176	BCR	15,855	0.95	0.72	27.64	47.62	23.76	0.98
2009	3488145	BCR	15,855	0.71	0.45	28.77	70.84		0.40
2009	3564189	BCR	15,855	0.97	0.53	14.49	72.22	12.36	0.93
2009	3515783	BCR	15,855	0.75	0.43	24.74	74.82		0.45
2009	3595560	BCR	15,855	1.47	0.74	13.96	23.78	61.36	0.90
2010	3487819	BCR	29,778	0.53	0.50	46.71	52.66		0.59
2010	3564186	BCR	29,778	0.83	0.45	19.12	76.60	3.37	0.86
2010	100000044142	BCR	29,778	0.84	0.36	14.85	84.31		0.78
2010	3595499	BCR	29,778	0.94	0.54	16.54	70.43	11.93	0.94
2010	100000025175	BCR	29,778	0.89	0.31	10.23	88.98		0.80
2010	3595574	BCR	29,778	0.82	0.48	21.33	73.33	4.11	0.82
2010	3515823	BCR	29,778	0.45	0.50	54.29	44.89		0.75
2010	3595532	BCR	29,778	0.78	0.59	29.81	59.83	9.07	1.14
2010	3488150	BCR	29,778	0.38	0.48	61.58	37.65		0.72
2010	3564176	BCR	29,778	0.93	0.68	25.77	52.92	20.21	0.96
2010	3488145	BCR	29,778	0.70	0.46	29.15	70.19		0.63
2010	3564189	BCR	29,778	0.97	0.57	16.29	67.62	14.87	1.00
2010	3515783	BCR	29,778	0.74	0.44	25.67	73.71		0.60
2010	3595560	BCR	29,778	1.42	0.79	17.69	20.94	60.38	0.84

Table 1.46 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 4 Form A

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2009	1	3497865	SR	0.4883		
2009	4	3487819	BCR_A	1.3917		
2009	5	3564186	BCR_B	2.1573	-2.6637	2.6637
2009	9	100000069423	SR	1.2932		
2009	11	100000069397	SR	-0.3646		
2009	20	100000044142	BCR_A	-0.7600		
2009	21	3595499	BCR_B	1.7168	-2.1933	2.1933
2007	27	3487979	SR	0.1230		
2008	28	100000025175	BCR_A	-1.1338		
2008	29	3595574	BCR_B	1.8916	-2.5047	2.5047
2009	31	100000069344	SR	-0.7274		
2009	34	100000069324	SR	-0.7925		
2008	35	100000025187	SR	-0.4685		
2009	36	100000025153	SR	-3.5593		
2009	37	100000069420	SR	-0.3698		
2009	38	3515823	BCR_A	1.6944		
2009	39	3595532	BCR_B	2.3297	-1.8612	1.8612
2009	44	3488150	BCR_B	2.1316		
2009	45	3564176	BCR_B	1.6184	-1.1561	1.1561
2008	46	100000007114	SR	-2.1097		
2009	48	100000069424	SR	-1.2509		
2009	51	3488145	BCR_A	0.1552		
2009	52	3564189	BCR_B	1.5957	-2.3378	2.3378
2007	53	3497871	SR	0.1619		
2006	56	100000252124	SR	0.4796		
2009	57	100000069399	SR	0.1774		
2008	65	100000011499	SR	-1.3494		
2009	73	3515783	BCR_A	-0.0626		
2009	74	3595560	BCR_B	0.1566	-0.3225	0.3225
2006	77	100000252111	SR	0.3118		
2008	79	100000011497	SR	0.3863		
2010	1	3497865	SR	0.2341		
2010	4	3487819	BCR_A	1.4019		
2010	5	3564186	BCR_B	2.7967	-3.0944	3.0944
2010	9	100000069423	SR	1.2547		
2010	11	100000069397	SR	-0.1970		
2010	20	100000044142	BCR_A	-0.6437		
2010	21	3595499	BCR_B	1.7679	-2.2015	2.2015
2010	27	3487979	SR	0.5114		

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2010	28	100000025175	BCR_A	-1.2660		
2010	29	3595574	BCR_B	2.5138	-2.7223	2.7223
2010	31	100000069344	SR	-0.9015		
2010	34	100000069324	SR	-0.9820		
2010	35	100000025187	SR	-0.5744		
2010	36	100000025153	SR	-3.4861		
2010	37	100000069420	SR	-0.9062		
2010	38	3515823	BCR_A	1.7822		
2010	39	3595532	BCR_B	2.2823	-1.9034	1.9034
2010	44	3488150	BCR_B	2.1818		
2010	45	3564176	BCR_B	1.7181	-1.4235	1.4235
2010	46	100000007114	SR	-2.0205		
2010	48	100000069424	SR	-1.2966		
2010	51	3488145	BCR_A	0.3722		
2010	52	3564189	BCR_B	1.5344	-2.0477	2.0477
2010	53	3497871	SR	-0.1044		
2010	56	100000252124	SR	0.5293		
2010	57	100000069399	SR	-0.2536		
2010	65	100000011499	SR	-1.3258		
2010	73	3515783	BCR_A	0.1496		
2010	74	3595560	BCR_B	0.4463	-0.0430	0.0430
2010	77	100000252111	SR	-0.1004		
2010	79	100000011497	SR	0.1555		

Note. These Rasch difficulties were based on a common scale.

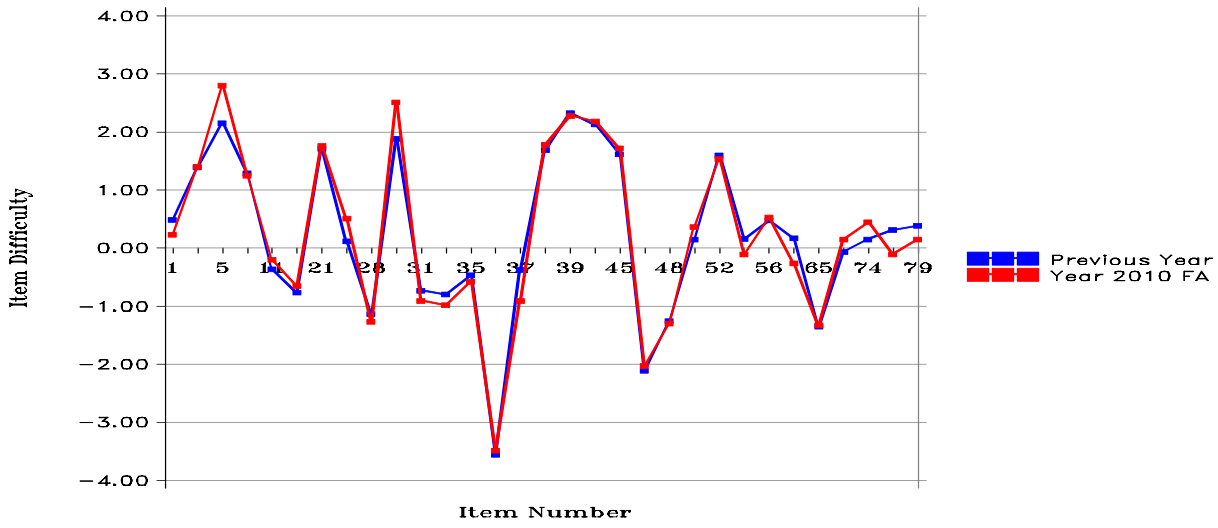

Figure 1.5 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 4 Form A

Table 1.47 P-Value Comparison of Core Items for Previous Year vs. Year 2010: Grade 4 Form F

Item CID	Previous Year	Y10 FF	Item CID	Previous Year	Y10 FF
3487811	0.63	0.70	**3985620**	**0.43**	**0.43**
3487819	**0.50**	**0.53**	**3488076**	**0.55**	**0.55**
3564186	**0.44**	**0.42**	**3595578**	**0.33**	**0.32**
100000069320	0.93	0.96	100000252099	0.72	0.82
100000069323	0.46	0.42	3487994	0.69	0.76
100000044142	**0.83**	**0.85**	**3488145**	**0.71**	**0.70**
3595499	**0.47**	**0.48**	**3564189**	**0.48**	**0.49**
100000069387	0.65	0.66	3551460	0.77	0.75
3548776	**0.51**	**0.50**	3488034	0.87	0.91
3595562	**0.58**	**0.61**	100000011498	0.81	0.85
100000007113	0.73	0.76	100000252117	0.53	0.61
3497864	0.80	0.80	**100000007087**	**0.71**	**0.77**
100000007120	0.69	0.69	**3595563**	**0.68**	**0.70**
3548767	0.71	0.70	100000011510	0.69	0.67
100000069321	0.90	0.94	100000011491	0.78	0.80
100000201938	**0.67**	**0.68**			

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item.

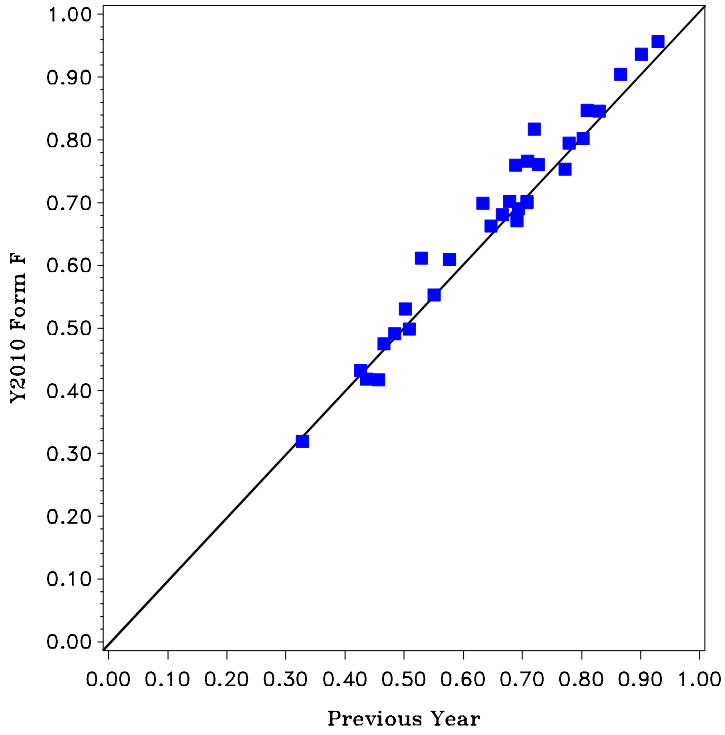

Table 1.48 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 4 Form F

						Score-Point Distribution (%)			
Year	Item CID	Item Type	N	Mean	SD	0	1	2	Omit
2009	3487819	BCR	15,855	0.50	0.50	49.12	50.26		0.62
2009	3564186	BCR	15,855	0.88	0.48	17.50	75.69	5.94	0.88
2009	100000044142	BCR	15,855	0.83	0.38	16.27	83.08		0.66
2009	3595499	BCR	15,855	0.93	0.55	17.54	69.52	11.91	1.02
2008	3548776	BCR	2,636	0.51	0.50	47.61	51.02		1.37
2008	3595562	BCR	2,636	1.16	0.75	19.42	41.88	36.87	1.82
2009	100000201938	BCR	23,861	0.67	0.47	32.77	66.69		0.54
2009	3985620	BCR	23,861	0.85	0.46	18.12	76.47	4.43	0.97
2008	3488076	BCR	2,282	0.55	0.50	43.56	55.13		1.31
2008	3595578	BCR	2,282	0.66	0.50	33.52	63.01	1.36	2.10
2009	3488145	BCR	15,855	0.71	0.45	28.77	70.84		0.40
2009	3564189	BCR	15,855	0.97	0.53	14.49	72.22	12.36	0.93
2008	100000007087	BCR	2,636	0.71	0.45	28.00	71.05		0.95
2008	3595563	BCR	2,636	1.36	0.64	7.74	45.98	44.99	1.29
2010	3487819	BCR	29,797	0.53	0.50	46.28	53.10		0.60
2010	3564186	BCR	29,797	0.84	0.45	18.69	77.16	3.35	0.73
2010	100000044142	BCR	29,797	0.85	0.36	14.38	84.73		0.83
2010	3595499	BCR	29,797	0.95	0.54	15.61	71.07	12.05	1.11
2010	3548776	BCR	29,797	0.50	0.50	49.17	49.96		0.82
2010	3595562	BCR	29,797	1.22	0.69	14.21	47.03	37.49	1.10
2010	100000201938	BCR	29,797	0.68	0.47	31.24	68.18		0.54
2010	3985620	BCR	29,797	0.87	0.47	17.76	76.26	5.21	0.74
2010	3488076	BCR	29,797	0.55	0.50	43.65	55.34		0.73
2010	3595578	BCR	29,797	0.64	0.50	35.39	61.89	1.10	1.31
2010	3488145	BCR	29,797	0.70	0.46	29.10	70.28		0.59
2010	3564189	BCR	29,797	0.98	0.57	15.91	67.39	15.49	0.94
2010	100000007087	BCR	29,797	0.77	0.42	22.65	76.72		0.56
2010	3595563	BCR	29,797	1.41	0.62	6.43	44.76	47.93	0.76

Table 1.49 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 4 Form F

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2007	1	3487811	SR	0.6277		
2009	4	3487819	BCR_A	1.3917		
2009	5	3564186	BCR_B	2.1573	-2.6637	2.6637
2009	9	100000069320	SR	-2.1790		
2009	11	100000069323	SR	1.6114		
2009	20	100000044142	BCR_A	-0.7600		
2009	21	3595499	BCR_B	1.7168	-2.1933	2.1933
2009	27	100000069387	SR	0.5288		
2008	28	3548776	BCR_A	1.2997		
2008	29	3595562	BCR_B	0.8907	-0.9659	0.9659
2008	31	100000007113	SR	0.0389		
2008	34	3497864	SR	-0.4664		
2008	35	100000007120	SR	0.3577		
2009	36	3548767	SR	0.0868		
2009	37	100000069321	SR	-1.6379		
2009	38	100000201938	BCR_A	0.4797		
2009	39	3985620	BCR_B	2.4328	-2.8528	2.8528
2008	44	3488076	BCR_B	1.0836		
2008	45	3595578	BCR_B	3.4522	-2.9019	2.9019
2006	46	100000252099	SR	-0.2084		
2007	48	3487994	SR	0.1102		
2009	51	3488145	BCR_A	0.1552		
2009	52	3564189	BCR_B	1.5957	-2.3378	2.3378
2007	53	3551460	SR	-0.5315		
2007	56	3488034	SR	-1.2641		
2008	57	100000011498	SR	-0.5633		
2006	65	100000252117	SR	0.9020		
2008	73	100000007087	BCR_A	0.1765		
2008	74	3595563	BCR_B	0.1138	-1.4376	1.4376
2008	77	100000011510	SR	0.3409		
2008	79	100000011491	SR	-0.2631		
2010	1	3487811	SR	0.3353		
2010	4	3487819	BCR_A	1.4019		
2010	5	3564186	BCR_B	2.7967	-3.0944	3.0944
2010	9	100000069320	SR	-2.2702		
2010	11	100000069323	SR	2.0168		
2010	20	100000044142	BCR_A	-0.6437		
2010	21	3595499	BCR_B	1.7679	-2.2015	2.2015
2010	27	100000069387	SR	0.6341		
2010	28	3548776	BCR_A	1.5300		
2010	29	3595562	BCR_B	0.7810	-1.1908	1.1908
2010	31	100000007113	SR	-0.0095		

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2010	34	3497864	SR	-0.2837		
2010	35	100000007120	SR	0.5750		
2010	36	3548767	SR	0.4355		
2010	37	100000069321	SR	-1.8118		
2010	38	100000201938	BCR_A	0.5229		
2010	39	3985620	BCR_B	2.3075	-2.5494	2.5494
2010	44	3488076	BCR_B	1.2332		
2010	45	3595578	BCR_B	3.6555	-2.9024	2.9024
2010	46	100000252099	SR	-0.4699		
2010	48	3487994	SR	0.0297		
2010	51	3488145	BCR_A	0.3722		
2010	52	3564189	BCR_B	1.5344	-2.0477	2.0477
2010	53	3551460	SR	-0.0075		
2010	56	3488034	SR	-1.3871		
2010	57	100000011498	SR	-0.6789		
2010	65	100000252117	SR	0.9730		
2010	73	100000007087	BCR_A	-0.0833		
2010	74	3595563	BCR_B	-0.0165	-1.4651	1.4651
2010	77	100000011510	SR	0.6061		
2010	79	100000011491	SR	-0.2494		

Note. These Rasch difficulties were based on a common scale.

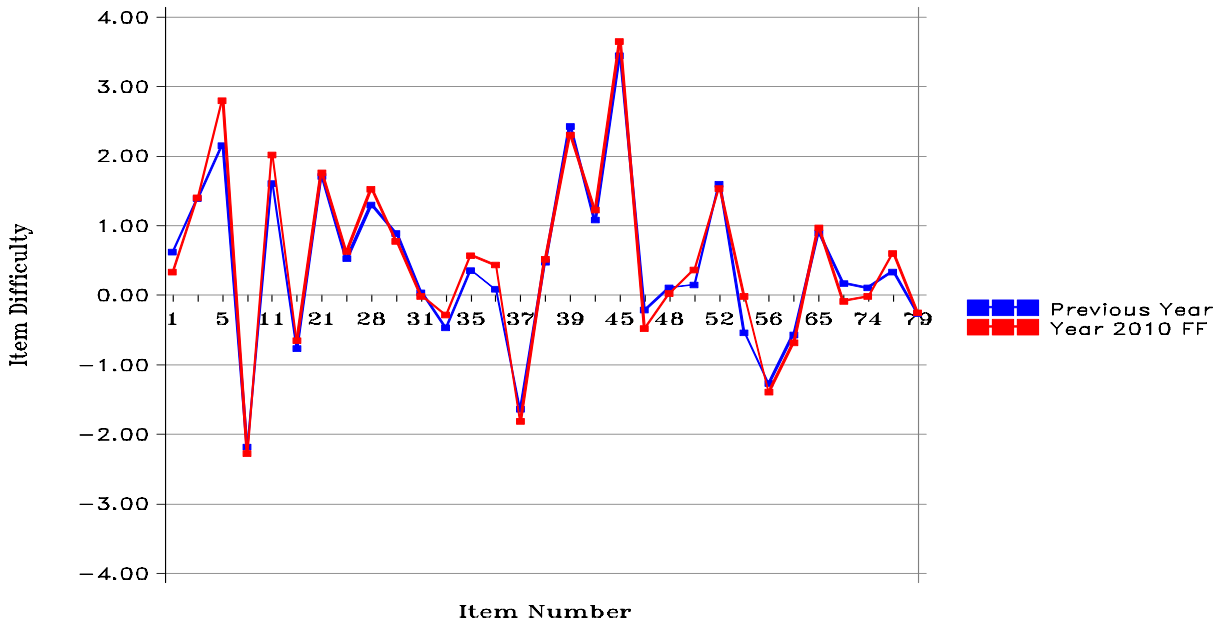

Figure 1.6 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 4 Form F

Table 1.50 P-Value Comparison of Core Items for Previous Year vs. Year 2010: Grade 5 Form A

Item CID	Previous Year	Y10 FA	Item CID	Previous Year	Y10 FA
3488503	0.78	0.79	**3488402**	**0.77**	**0.79**
3488420	0.64	0.65	**3564000**	**0.55**	**0.57**
3488471	**0.31**	**0.31**	3492133	0.55	0.57
3564052	**0.42**	**0.45**	3488485	0.77	0.89
100000065210	0.78	0.80	**100000022528**	**0.38**	**0.43**
100000065234	0.88	0.92	**3595465**	**0.24**	**0.27**
100000028266	0.93	0.95	100000086836	0.55	0.59
100000065176	0.79	0.83	3511636	0.49	0.46
100000366297	0.76	0.70	100000065196	0.76	0.80
3488376	0.78	0.80	**3512644**	**0.37**	**0.44**
100000028212	**0.65**	**0.68**	**3595445**	**0.47**	**0.53**
3595468	**0.45**	**0.47**	100000366304	0.85	0.91
100000143144	0.57	0.53	**3488530**	**0.36**	**0.36**
3511483	**0.38**	**0.37**	**3564054**	**0.31**	**0.32**
3563992	**0.34**	**0.39**	3488237	0.64	0.63
100000022541	0.74	0.74	**100000022532**	**0.36**	**0.42**
100000030431	0.45	0.48	**3595471**	**0.37**	**0.42**
3488370	0.86	0.87			

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item.

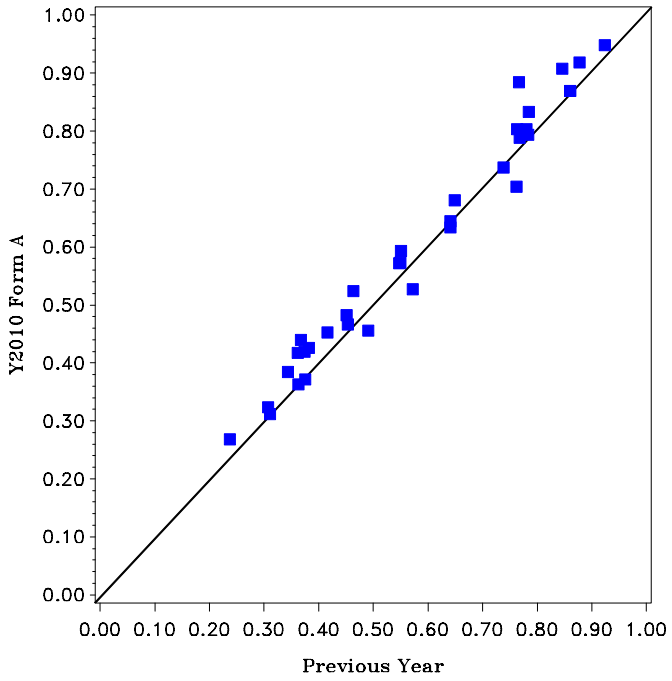

Table 1.51 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 5 Form A

						Score-Point Distribution (%)				
Year	Item CID	Item Type	N	Mean	SD	0	1	2	3	Omit
2009	3488471	BCR	6,870	0.31	0.46	67.38	31.18			1.44
2009	3564052	BCR	6,870	0.83	0.58	24.09	63.78	9.84		2.29
2008	100000028212	BCR	2,634	0.65	0.48	33.79	64.96			1.25
2008	3595468	BCR	2,634	0.91	0.63	23.27	60.10	15.41		1.21
2007	3511483	BCR	31,083	0.38	0.48	60.64	37.64			1.72
2007	3563992	BCR	31,083	0.69	0.68	41.32	43.49	12.70		2.49
2007	3488402	ECR	2,164	0.77	0.42	20.84	76.94			2.22
2007	3564000	ECR	2,164	1.65	0.98	12.71	23.66	39.42	20.93	3.28
2008	100000022528	BCR	2,590	0.38	0.49	57.76	38.26			3.98
2008	3595465	BCR	2,590	0.48	0.57	51.39	40.19	3.71		4.71
2008	3512644	BCR	30,537	0.37	0.48	58.24	36.89			4.88
2008	3595445	BCR	30,537	0.93	0.76	26.55	42.24	25.38		5.83
2007	3488530	BCR	2,125	0.36	0.48	57.36	36.38			6.26
2007	3564054	BCR	2,125	0.62	0.73	46.21	31.58	15.06		7.15
2008	100000022532	BCR	2,633	0.36	0.48	59.06	36.23			4.71
2008	3595471	BCR	2,633	0.75	0.92	52.79	8.36	33.31		5.55
2010	3488471	BCR	29,532	0.31	0.46	66.81	31.26			1.85
2010	3564052	BCR	29,532	0.91	0.64	23.37	57.77	16.45		2.15
2010	100000028212	BCR	29,532	0.68	0.47	31.00	68.21			0.76
2010	3595468	BCR	29,532	0.93	0.62	21.46	61.53	15.96		0.95
2010	3511483	BCR	29,532	0.37	0.48	61.16	37.31			1.49
2010	3563992	BCR	29,532	0.77	0.68	35.22	48.34	14.33		1.85
2010	3488402	ECR	29,532	0.79	0.41	19.43	78.96			1.58
2010	3564000	ECR	29,532	1.72	0.96	11.04	23.67	40.23	22.65	2.30
2010	100000022528	BCR	29,532	0.43	0.49	54.35	42.74			2.76
2010	3595465	BCR	29,532	0.54	0.60	48.42	43.11	5.36		2.91
2010	3512644	BCR	29,532	0.44	0.50	52.68	44.08			3.12
2010	3595445	BCR	29,532	1.05	0.78	24.94	38.22	33.44		3.21
2010	3488530	BCR	29,532	0.36	0.48	59.40	36.43			3.90
2010	3564054	BCR	29,532	0.65	0.75	47.42	31.72	16.58		4.05
2010	100000022532	BCR	29,532	0.42	0.49	55.67	41.81			2.40
2010	3595471	BCR	29,532	0.84	0.94	49.66	10.05	37.09		2.90

Table 1.52 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 5 Form A

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	1	3488503	SR	-0.4590			
2007	3	3488420	SR	0.2073			
2009	4	3488471	BCR_A	2.2862			
2009	5	3564052	BCR_B	1.8914	-2.0216	2.0216	
2009	6	100000065210	SR	-0.4799			
2009	7	100000065234	SR	-1.4366			
2008	9	100000028266	SR	-1.9932			
2009	10	100000065176	SR	-0.5990			
2006	17	100000366297	SR	-0.4641			
2008	22	3488376	SR	-0.4234			
2008	24	100000028212	BCR_A	0.4540			
2008	25	3595468	BCR_B	1.5778	-1.6677	1.6677	
2009	34	100000143144	SR	0.8647			
2007	35	3511483	BCR_A	1.7333			
2007	36	3563992	BCR_B	1.9616	-1.2852	1.2852	
2008	40	100000022541	SR	-0.2067			
2008	41	100000030431	SR	1.3829			
2007	44	3488370	SR	-1.1987			
2007	45	3488402	ECR_A	-0.6294			
2007	46	3564000	ECR_B	0.6751	-1.2482	-0.3864	1.6345
2008	50	3492133	SR	0.8787			
2007	51	3488485	SR	-0.4193			
2008	52	100000022528	BCR_A	1.7756			
2008	53	3595465	BCR_B	3.1227	-1.6243	1.6243	
2009	54	100000086836	SR	0.8230			
2008	57	3511636	SR	1.2623			
2009	59	100000065196	SR	-0.4534			
2008	62	3512644	BCR_A	1.8531			
2008	63	3595445	BCR_B	1.3274	-0.9532	0.9532	
2006	65	100000366304	SR	-1.1281			
2007	73	3488530	BCR_A	1.6774			
2007	74	3564054	BCR_B	1.9860	-0.6652	0.6652	
2008	79	3488237	SR	0.3823			
2008	80	100000022532	BCR_A	1.8775			
2008	81	3595471	BCR_B	1.6734	1.1626	-1.1626	
2010	1	3488503	SR	-0.4592			
2010	3	3488420	SR	0.4222			
2010	4	3488471	BCR_A	2.2200			
2010	5	3564052	BCR_B	1.5853	-1.5919	1.5919	
2010	6	100000065210	SR	-0.4266			
2010	7	100000065234	SR	-1.7803			
2010	9	100000028266	SR	-2.3060			

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2010	10	100000065176	SR	-0.8064			
2010	17	100000366297	SR	0.0688			
2010	22	3488376	SR	-0.5510			
2010	24	100000028212	BCR_A	0.3107			
2010	25	3595468	BCR_B	1.4747	-1.6789	1.6789	
2010	34	100000143144	SR	1.0908			
2010	35	3511483	BCR_A	1.9029			
2010	36	3563992	BCR_B	1.9325	-1.2674	1.2674	
2010	40	100000022541	SR	-0.1671			
2010	41	100000030431	SR	1.3624			
2010	44	3488370	SR	-1.0980			
2010	45	3488402	ECR_A	-0.4879			
2010	46	3564000	ECR_B	0.8000	-1.3484	-0.2459	1.5943
2010	50	3492133	SR	0.7803			
2010	51	3488485	SR	-1.2799			
2010	52	100000022528	BCR_A	1.6254			
2010	53	3595465	BCR_B	2.9001	-1.5156	1.5156	
2010	54	100000086836	SR	0.6849			
2010	57	3511636	SR	1.4392			
2010	59	100000065196	SR	-0.5374			
2010	62	3512644	BCR_A	1.6118			
2010	63	3595445	BCR_B	1.0439	-0.7480	0.7480	
2010	65	100000366304	SR	-1.6268			
2010	73	3488530	BCR_A	1.8673			
2010	74	3564054	BCR_B	2.2051	-0.6519	0.6519	
2010	79	3488237	SR	0.4508			
2010	80	100000022532	BCR_A	1.6379			
2010	81	3595471	BCR_B	1.5127	0.8480	-0.8480	

Note. These Rasch difficulties were based on a common scale.

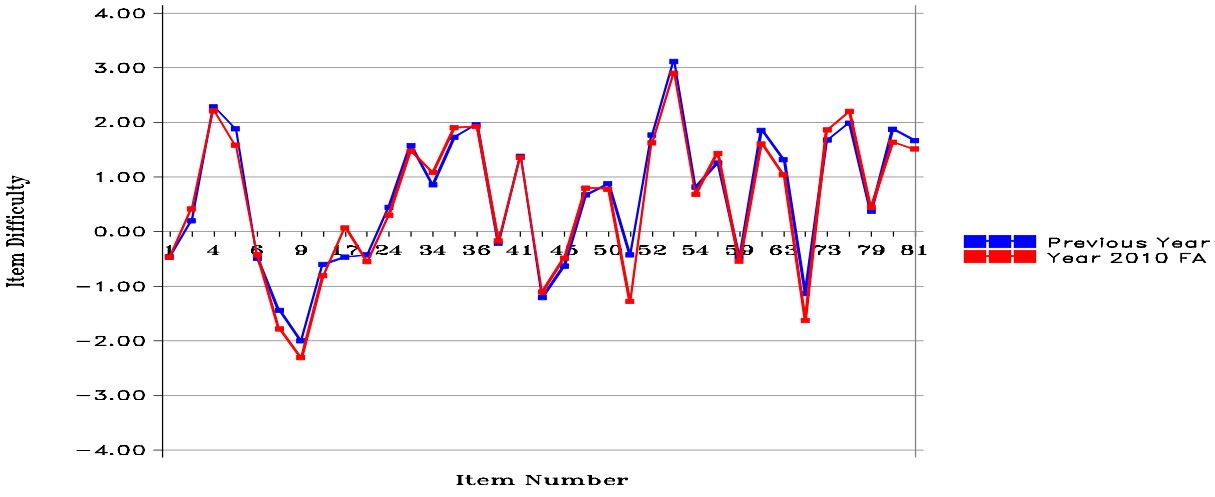

Figure 1.7 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 5 Form A

Table 1.53 P-Value Comparisons of Core Items for Previous Year vs. Year 2010: Grade 5 Form F

Item CID	Previous Year	Y10 FF	Item CID	Previous Year	Y10 FF
3488494	0.62	0.76	**100000028203**	**0.70**	**0.65**
3488475	0.72	0.69	**3595467**	**0.41**	**0.38**
3488471	**0.31**	**0.32**	3492133	0.55	0.58
3564052	**0.42**	**0.46**	3488485	0.77	0.85
100000065277	0.86	0.89	**3556476**	**0.50**	**0.51**
100000028272	0.69	0.73	**3563990**	**0.44**	**0.47**
100000028266	0.93	0.96	100000086836	0.55	0.61
3492129	0.51	0.63	3511636	0.49	0.49
3488511	0.44	0.40	100000065196	0.76	0.81
100000028227	0.25	0.26	**3511455**	**0.79**	**0.83**
3488459	**0.63**	**0.65**	**3563993**	**0.67**	**0.74**
3595470	**0.50**	**0.53**	100000143146	0.72	0.76
100000143144	0.57	0.52	**3512564**	**0.36**	**0.40**
3511483	**0.38**	**0.38**	**3595444**	**0.32**	**0.29**
3563992	**0.34**	**0.40**	100000022545	0.77	0.75
100000366311	0.78	0.82	**3488346**	**0.42**	**0.50**
100000022540	0.77	0.74	**3595473**	**0.42**	**0.47**
3488249	0.49	0.54			

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item or Extended Constructed Response (ECR) item.

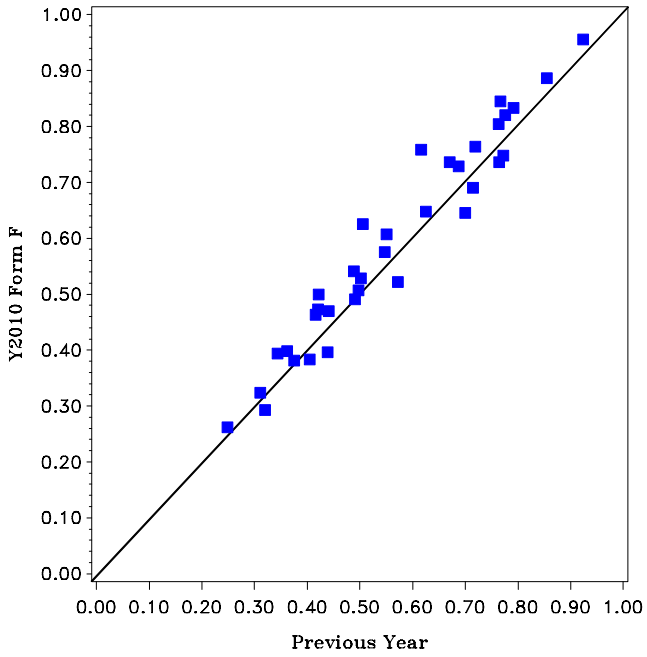

Table 1.54 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 5 Form F

						Score-Point Distribution (%)				
Year	Item CID	Item Type	N	Mean	SD	0	1	2	3	Omit
2009	3488471	BCR	6,870	0.31	0.46	67.38	31.18			1.44
2009	3564052	BCR	6,870	0.83	0.58	24.09	63.78	9.84		2.29
2008	3488459	BCR	2,633	0.63	0.48	36.04	62.63			1.33
2008	3595470	BCR	2,633	1.01	0.57	14.09	67.57	16.48		1.86
2007	3511483	BCR	31,083	0.38	0.48	60.64	37.64			1.72
2007	3563992	BCR	31,083	0.69	0.68	41.32	43.49	12.70		2.49
2008	100000028203	ECR	2,634	0.70	0.46	28.85	70.12			1.03
2008	3595467	ECR	2,634	1.22	0.80	15.45	52.54	23.65	7.33	1.03
2008	3556476	BCR	30,537	0.50	0.50	48.14	49.87			1.98
2008	3563990	BCR	30,537	0.88	0.91	45.22	15.45	36.46		2.86
2007	3511455	BCR	31,083	0.79	0.41	18.57	79.25			2.18
2007	3563993	BCR	31,083	1.34	0.81	18.53	23.07	55.57		2.83
2008	3512564	BCR	30,537	0.36	0.48	60.11	36.30			3.59
2008	3595444	BCR	30,537	0.64	0.60	39.26	51.16	6.54		3.04
2008	3488346	BCR	2,620	0.42	0.49	53.36	42.25			4.39
2008	3595473	BCR	2,620	0.84	0.92	45.69	13.28	35.53		5.50
2010	3488471	BCR	28,934	0.32	0.47	65.91	32.43			1.58
2010	3564052	BCR	28,934	0.93	0.64	22.51	57.97	17.41		1.86
2010	3488459	BCR	28,934	0.65	0.48	33.47	64.89			1.60
2010	3595470	BCR	28,934	1.06	0.55	10.66	69.23	18.37		1.62
2010	3511483	BCR	28,934	0.38	0.49	59.92	38.21			1.81
2010	3563992	BCR	28,934	0.79	0.68	33.75	48.72	15.16		2.14
2010	100000028203	ECR	28,934	0.65	0.48	32.72	64.65			2.61
2010	3595467	ECR	28,934	1.15	0.77	15.49	54.81	22.40	5.23	1.97
2010	3556476	BCR	28,934	0.51	0.50	47.57	50.84			1.54
2010	3563990	BCR	28,934	0.94	0.90	41.84	18.35	37.90		1.83
2010	3511455	BCR	28,934	0.83	0.37	15.05	83.40			1.48
2010	3563993	BCR	28,934	1.47	0.77	15.59	17.96	64.72		1.63
2010	3512564	BCR	28,934	0.40	0.49	57.21	39.93			2.87
2010	3595444	BCR	28,934	0.59	0.54	41.86	53.22	2.73		2.08
2010	3488346	BCR	28,934	0.50	0.50	46.58	50.04			3.30
2010	3595473	BCR	28,934	0.95	0.94	42.66	12.09	41.32		3.76

Table 1.55 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 5 Form F

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2007	1	3488494	SR	0.4231			
2007	3	3488475	SR	-0.0522			
2009	4	3488471	BCR_A	2.2862			
2009	5	3564052	BCR_B	1.8914	-2.0216	2.0216	
2009	6	100000065277	SR	-1.0922			
2008	7	100000028272	SR	0.1762			
2008	9	100000028266	SR	-1.9932			
2007	10	3492129	SR	0.9999			
2007	17	3488511	SR	1.3558			
2008	22	100000028227	SR	2.7266			
2008	24	3488459	BCR_A	0.5092			
2008	25	3595470	BCR_B	1.1507	-2.0643	2.0643	
2009	34	100000143144	SR	0.8647			
2007	35	3511483	BCR_A	1.7333			
2007	36	3563992	BCR_B	1.9616	-1.2852	1.2852	
2006	40	100000366311	SR	-0.6851			
2008	41	100000022540	SR	-0.3064			
2007	44	3488249	SR	1.1138			
2008	45	100000028203	ECR_A	0.1543			
2008	46	3595467	ECR_B	1.7799	-2.3755	0.5596	1.8159
2008	50	3492133	SR	0.8787			
2007	51	3488485	SR	-0.4193			
2008	52	3556476	BCR_A	1.2085			
2008	53	3563990	BCR_B	1.3852	0.6408	-0.6408	
2009	54	100000086836	SR	0.8230			
2008	57	3511636	SR	1.2623			
2009	59	100000065196	SR	-0.4534			
2007	62	3511455	BCR_A	-0.7457			
2007	63	3563993	BCR_B	0.1392	-0.1147	0.1147	
2009	65	100000143146	SR	-0.0741			
2008	73	3512564	BCR_A	1.7934			
2008	74	3595444	BCR_B	2.3824	-1.8916	1.8916	
2008	79	100000022545	SR	-0.3479			
2008	80	3488346	BCR_A	1.5033			
2008	81	3595473	BCR_B	1.4359	0.6485	-0.6485	
2010	1	3488494	SR	-0.1633			
2010	3	3488475	SR	0.2377			
2010	4	3488471	BCR_A	2.2200			
2010	5	3564052	BCR_B	1.5853	-1.5919	1.5919	
2010	6	100000065277	SR	-1.2857			
2010	7	100000028272	SR	0.0484			
2010	9	100000028266	SR	-2.3060			

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2010	10	3492129	SR	0.5872			
2010	17	3488511	SR	1.8688			
2010	22	100000028227	SR	2.6325			
2010	24	3488459	BCR_A	0.5314			
2010	25	3595470	BCR_B	1.0467	-2.1176	2.1176	
2010	34	100000143144	SR	1.0908			
2010	35	3511483	BCR_A	1.9029			
2010	36	3563992	BCR_B	1.9325	-1.2674	1.2674	
2010	40	100000366311	SR	-0.6239			
2010	41	100000022540	SR	-0.0859			
2010	44	3488249	SR	1.0221			
2010	45	100000028203	ECR_A	0.5215			
2010	46	3595467	ECR_B	2.0262	-2.4394	0.4672	1.9722
2010	50	3492133	SR	0.7803			
2010	51	3488485	SR	-1.2799			
2010	52	3556476	BCR_A	1.2350			
2010	53	3563990	BCR_B	1.3785	0.2627	-0.2627	
2010	54	100000086836	SR	0.6849			
2010	57	3511636	SR	1.4392			
2010	59	100000065196	SR	-0.5374			
2010	62	3511455	BCR_A	-0.8030			
2010	63	3563993	BCR_B	0.1427	0.1393	-0.1393	
2010	65	100000143146	SR	-0.1540			
2010	73	3512564	BCR_A	1.8176			
2010	74	3595444	BCR_B	3.0464	-2.0643	2.0643	
2010	79	100000022545	SR	-0.1281			
2010	80	3488346	BCR_A	1.1973			
2010	81	3595473	BCR_B	1.3196	0.8754	-0.8754	

Note. These Rasch difficulties were based on a common scale.

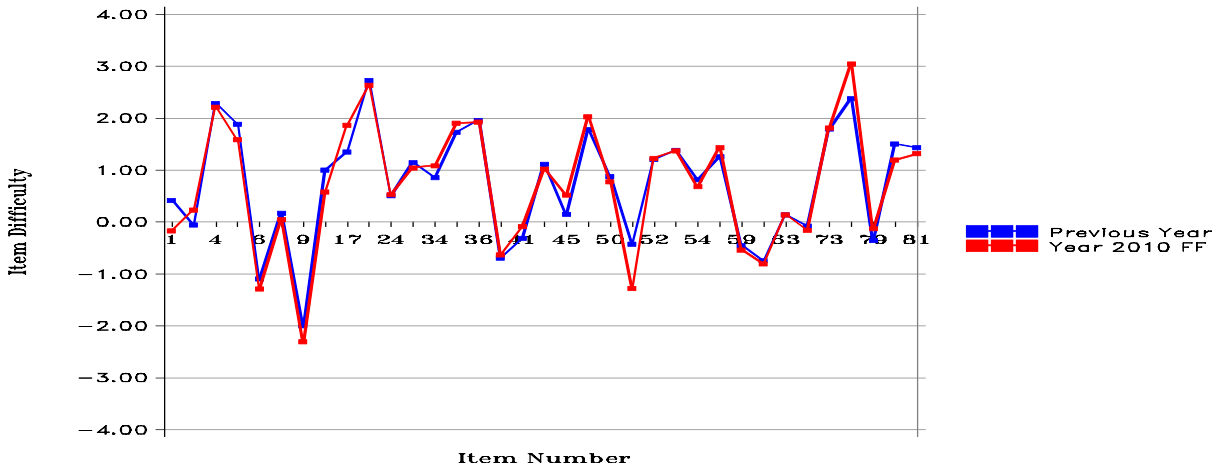

Figure 1.8 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 5 Form F

Table 1.56 P-Value Comparisons of Core Items for Previous Year vs. Year 2010: Grade 6 Form A

Item CID	Previous Year	Y10 FA	Item CID	Previous Year	Y10 FA
100000065103	0.65	0.67	**100000022498**	**0.64**	**0.61**
3488501	0.57	0.73	**3595494**	**0.60**	**0.54**
100000028374	**0.68**	**0.64**	3492073	0.78	0.80
3595487	**0.56**	**0.61**	100000022473	0.68	0.59
100000064636	0.79	0.78	100000065061	0.83	0.80
100000064615	0.81	0.81	**100000022501**	**0.42**	**0.44**
3548404	**0.55**	**0.53**	**3595485**	**0.72**	**0.77**
3564013	**0.48**	**0.43**	100000012861	0.86	0.85
100000065117	0.56	0.52	**3488399**	**0.58**	**0.63**
100000065065	0.44	0.40	**3595482**	**0.22**	**0.22**
100000022476	0.63	0.64	100000004449	0.51	0.48
100000028367	**0.36**	**0.37**	3488362	0.48	0.52
3595489	**0.52**	**0.53**	100000065076	0.52	0.51
100000094478	0.83	0.83	**3516327**	**0.44**	**0.49**
3488359	0.86	0.91	**3564005**	**0.61**	**0.62**
100000065053	0.41	0.38	100000064631	0.79	0.78

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item or Extended Constructed Response (ECR) item.

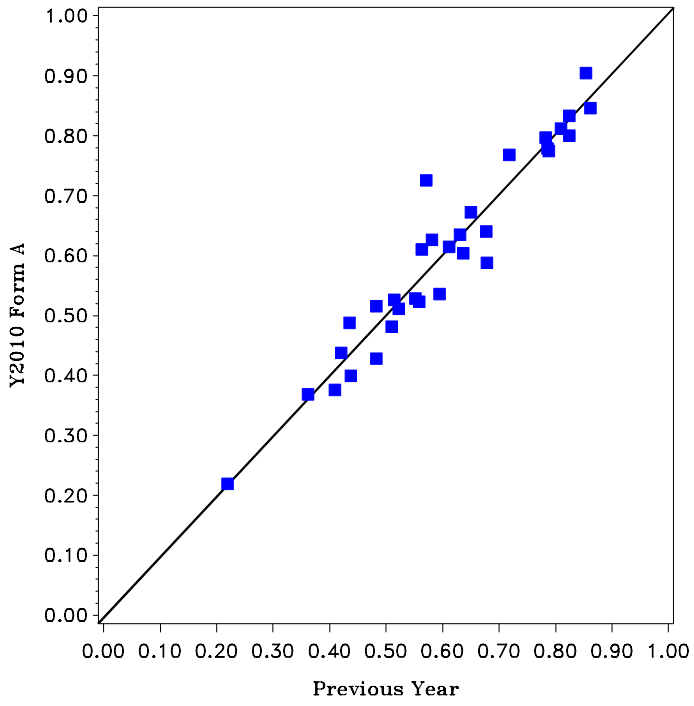

Table 1.57 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 6 Form A

						Score-Point Distribution (%)				
Year	Item CID	Item Type	N	Mean	SD	0	1	2	3	Omit
2008	100000028374	ECR	2,286	0.68	0.47	30.66	67.85			1.49
2008	3595487	ECR	2,286	1.69	1.00	12.82	24.28	36.79	23.75	2.36
2009	3548404	BCR	18,303	0.55	0.50	42.69	55.29			2.03
2009	3564013	BCR	18,303	0.97	0.41	7.60	82.79	7.04		2.57
2008	100000028367	BCR	2,280	0.36	0.48	62.54	36.27			1.18
2008	3595489	BCR	2,280	1.03	0.85	32.76	27.98	37.59		1.67
2008	100000022498	BCR	2,287	0.64	0.48	34.89	63.71			1.40
2008	3595494	BCR	2,287	1.19	0.78	20.77	35.24	41.98		2.01
2008	100000022501	BCR	2,293	0.42	0.49	56.39	42.13			1.48
2008	3595485	BCR	2,293	1.44	0.66	7.37	37.64	53.03		1.96
2008	3488399	BCR	2,240	0.58	0.49	40.63	58.21			1.16
2008	3595482	BCR	2,240	0.44	0.73	68.93	15.04	14.46		1.56
2008	3516327	BCR	30,292	0.44	0.50	54.40	43.67			1.92
2008	3564005	BCR	30,292	1.23	0.74	16.34	39.59	41.47		2.60
2010	100000028374	ECR	29,928	0.64	0.48	35.03	64.14			0.82
2010	3595487	ECR	29,928	1.83	1.00	10.01	26.18	30.48	32.11	1.13
2010	3548404	BCR	29,928	0.53	0.50	44.46	52.94			2.57
2010	3564013	BCR	29,928	0.86	0.48	16.33	74.96	5.43		3.19
2010	100000028367	BCR	29,928	0.37	0.48	61.39	36.90			1.67
2010	3595489	BCR	29,928	1.05	0.84	30.32	29.89	37.73		1.99
2010	100000022498	BCR	29,928	0.61	0.49	37.23	60.54			2.17
2010	3595494	BCR	29,928	1.07	0.79	24.38	37.44	34.99		2.98
2010	100000022501	BCR	29,928	0.44	0.50	54.56	43.86			1.56
2010	3595485	BCR	29,928	1.54	0.61	4.39	33.94	59.92		1.70
2010	3488399	BCR	29,928	0.63	0.48	34.38	62.72			2.84
2010	3595482	BCR	29,928	0.44	0.74	68.04	14.04	14.95		2.90
2010	3516327	BCR	29,928	0.49	0.50	49.20	48.91			1.76
2010	3564005	BCR	29,928	1.23	0.74	15.58	40.35	41.42		2.44

Table 1.58 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 6 Form A

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2009	2	100000065103	SR	0.1362			
2007	5	3488501	SR	0.2783			
2008	7	100000028374	ECR_A	-0.0590			
2008	8	3595487	ECR_B	0.5948	-1.2603	-0.2661	1.5263
2009	13	100000064636	SR	-0.7442			
2009	21	100000064615	SR	-0.8981			
2009	22	3548404	BCR_A	0.7098			
2009	23	3564013	BCR_B	1.2037	-3.0576	3.0576	
2009	24	100000065117	SR	0.6628			
2009	28	100000065065	SR	1.2537			
2008	29	100000022476	SR	0.3026			
2008	31	100000028367	BCR_A	1.7298			
2008	32	3595489	BCR_B	0.8813	-0.2699	0.2699	
2009	33	100000094478	SR	-1.0807			
2007	44	3488359	SR	-1.6712			
2009	46	100000065053	SR	1.4304			
2008	47	100000022498	BCR_A	0.2274			
2008	48	3595494	BCR_B	0.4522	-0.6967	0.6967	
2007	49	3492073	SR	-0.9462			
2008	52	100000022473	SR	0.0812			
2009	53	100000065061	SR	-1.0696			
2008	59	100000022501	BCR_A	1.3577			
2008	60	3595485	BCR_B	-0.5202	-1.1544	1.1544	
2008	62	100000012861	SR	-1.3605			
2008	66	3488399	BCR_A	0.5306			
2008	67	3595482	BCR_B	2.3988	0.2503	-0.2503	
2009	69	100000004449	SR	0.9156			
2007	70	3488362	SR	0.8884			
2009	71	100000065076	SR	0.8417			
2008	77	3516327	BCR_A	1.1725			
2008	78	3564005	BCR_B	0.2939	-0.8570	0.8570	
2009	79	100000064631	SR	-0.7289			
2010	2	100000065103	SR	0.0740			
2010	5	3488501	SR	-0.2534			
2010	7	100000028374	ECR_A	0.2351			
2010	8	3595487	ECR_B	0.3640	-1.3201	0.1828	1.1373
2010	13	100000064636	SR	-0.6205			
2010	21	100000064615	SR	-0.9301			
2010	22	3548404	BCR_A	0.7819			
2010	23	3564013	BCR_B	1.7242	-2.6919	2.6919	
2010	24	100000065117	SR	0.8319			
2010	28	100000065065	SR	1.4833			

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2010	29	100000022476	SR	0.2578			
2010	31	100000028367	BCR_A	1.6930			
2010	32	3595489	BCR_B	0.8402	-0.3408	0.3408	
2010	33	100000094478	SR	-1.0731			
2010	44	3488359	SR	-1.8828			
2010	46	100000065053	SR	1.6667			
2010	47	100000022498	BCR_A	0.3903			
2010	48	3595494	BCR_B	0.7504	-0.7203	0.7203	
2010	49	3492073	SR	-0.7730			
2010	52	100000022473	SR	0.5012			
2010	53	100000065061	SR	-0.8511			
2010	59	100000022501	BCR_A	1.3128			
2010	60	3595485	BCR_B	-0.8338	-1.1925	1.1925	
2010	62	100000012861	SR	-1.2153			
2010	66	3488399	BCR_A	0.2229			
2010	67	3595482	BCR_B	2.3397	0.3430	-0.3430	
2010	69	100000004449	SR	1.0940			
2010	70	3488362	SR	0.9268			
2010	71	100000065076	SR	0.9253			
2010	77	3516327	BCR_A	1.0327			
2010	78	3564005	BCR_B	0.2748	-0.9394	0.9394	
2010	79	100000064631	SR	-0.6210			

Note. These Rasch difficulties were based on a common scale.

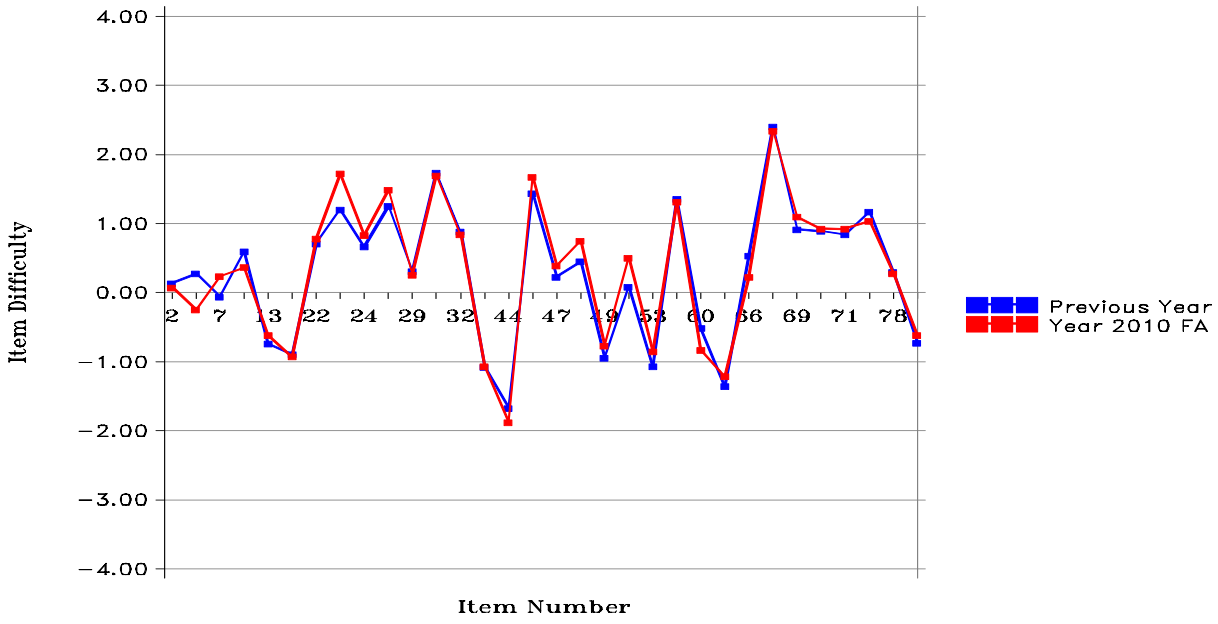

Figure 1.9 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 6 Form A

Table 1.59 P-Value Comparison of Core Items for Previous Year vs. Year 2010: Grade 6 Form F

Item CID	Previous Year	Y10 FF	Item CID	Previous Year	Y10 FF
100000065090	0.91	0.92	**100000022498**	**0.64**	**0.60**
100000028429	0.59	0.59	**3595494**	**0.60**	**0.56**
100000028374	**0.68**	**0.66**	3492094	0.73	0.72
3595487	**0.56**	**0.62**	100000064637	0.62	0.62
100000028410	0.83	0.87	100000094475	0.80	0.75
3492071	0.60	0.65	**100000022492**	**0.67**	**0.71**
3516627	**0.54**	**0.55**	**3595479**	**0.41**	**0.50**
3564006	**0.43**	**0.42**	100000064783	0.59	0.57
100000004453	0.61	0.54	**3488281**	**0.68**	**0.65**
100000065065	0.44	0.41	**3564063**	**0.59**	**0.59**
3488297	0.78	0.79	3488317	0.86	0.90
100000028363	**0.67**	**0.70**	100000064622	0.88	0.88
3595491	**0.58**	**0.65**	100000065148	0.58	0.60
100000064614	0.80	0.80	**3516327**	**0.44**	**0.49**
100000094472	0.47	0.47	**3564005**	**0.61**	**0.63**
100000028417	0.91	0.87	100000064624	0.81	0.79

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item or Extended Constructed Response (ECR) item.

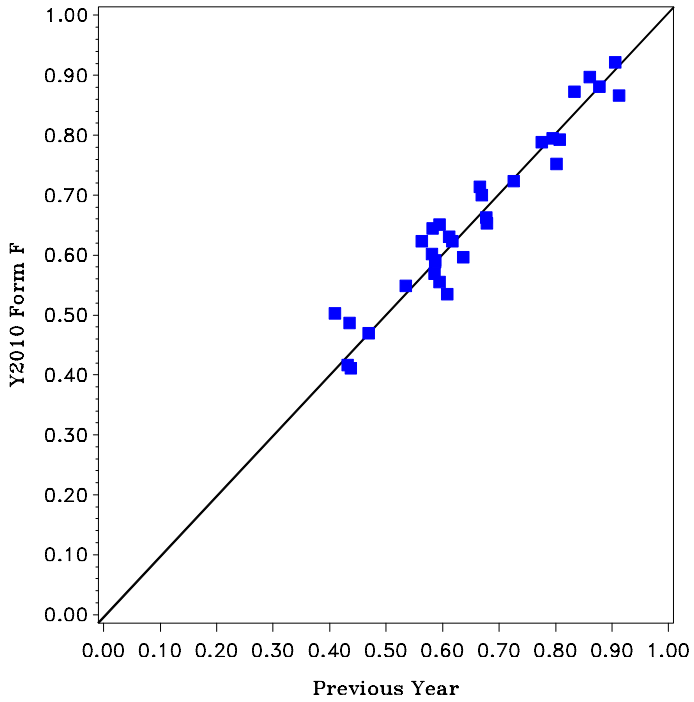

Table 1.60 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 6 Form F

						Score-Point Distribution (%)				
Year	Item CID	Item Type	N	Mean	SD	0	1	2	3	Omit
2008	100000028374	ECR	2,286	0.68	0.47	30.66	67.85			1.49
2008	3595487	ECR	2,286	1.69	1.00	12.82	24.28	36.79	23.75	2.36
2009	3516627	BCR	18,240	0.54	0.50	41.28	53.54			5.18
2009	3564006	BCR	18,240	0.87	0.60	19.67	62.34	12.11		5.88
2008	100000028363	BCR	2,301	0.67	0.47	31.90	67.01			1.09
2008	3595491	BCR	2,301	1.17	0.83	25.73	28.77	43.98		1.52
2008	100000022498	BCR	2,287	0.64	0.48	34.89	63.71			1.40
2008	3595494	BCR	2,287	1.19	0.78	20.77	35.24	41.98		2.01
2008	100000022492	BCR	2,281	0.67	0.47	30.73	66.68			2.59
2008	3595479	BCR	2,281	0.82	0.80	38.93	32.88	24.64		3.55
2007	3488281	BCR	2,125	0.68	0.47	25.79	68.05			6.16
2007	3564063	BCR	2,125	1.18	0.69	10.54	48.94	34.35		6.16
2008	3516327	BCR	30,292	0.44	0.50	54.40	43.67			1.92
2008	3564005	BCR	30,292	1.23	0.74	16.34	39.59	41.47		2.60
2010	100000028374	ECR	29,384	0.66	0.47	32.88	66.31			0.80
2010	3595487	ECR	29,384	1.87	0.99	9.32	25.26	30.85	33.46	1.06
2010	3516627	BCR	29,384	0.55	0.50	40.68	54.92			4.26
2010	3564006	BCR	29,384	0.83	0.57	20.66	64.35	9.55		5.05
2010	100000028363	BCR	29,384	0.70	0.46	28.22	70.12			1.63
2010	3595491	BCR	29,384	1.29	0.80	19.75	27.70	50.69		1.81
2010	100000022498	BCR	29,384	0.60	0.49	37.55	59.70			2.67
2010	3595494	BCR	29,384	1.11	0.79	22.60	36.21	37.54		3.47
2010	100000022492	BCR	29,384	0.71	0.45	25.85	71.47			2.67
2010	3595479	BCR	29,384	1.01	0.86	32.36	26.88	36.97		3.77
2010	3488281	BCR	29,384	0.65	0.48	30.36	65.37			4.21
2010	3564063	BCR	29,384	1.18	0.69	11.69	49.69	34.34		4.21
2010	3516327	BCR	29,384	0.49	0.50	49.70	48.76			1.41
2010	3564005	BCR	29,384	1.26	0.73	14.72	40.03	43.20		1.91

Table 1.61 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 6 Form F

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2009	2	100000065090	SR	-1.8921			
2008	5	100000028429	SR	0.5827			
2008	7	100000028374	ECR_A	-0.0590			
2008	8	3595487	ECR_B	0.5948	-1.2603	-0.2661	1.5263
2008	13	100000028410	SR	-1.0525			
2007	21	3492071	SR	0.1167			
2009	22	3516627	BCR_A	0.6268			
2009	23	3564006	BCR_B	1.3578	-1.9164	1.9164	
2008	24	100000004453	SR	0.5745			
2009	28	100000065065	SR	1.2537			
2008	29	3488297	SR	-0.5874			
2008	31	100000028363	BCR_A	0.0350			
2008	32	3595491	BCR_B	0.5574	-0.3422	0.3422	
2009	33	100000064614	SR	-0.7732			
2009	44	100000094472	SR	1.0702			
2008	46	100000028417	SR	-1.8196			
2008	47	100000022498	BCR_A	0.2274			
2008	48	3595494	BCR_B	0.4522	-0.6967	0.6967	
2007	49	3492094	SR	-0.4771			
2009	52	100000064637	SR	0.3379			
2009	53	100000094475	SR	-0.8688			
2008	59	100000022492	BCR_A	-0.2094			
2008	60	3595479	BCR_B	1.2104	-0.5752	0.5752	
2009	62	100000064783	SR	0.4609			
2007	66	3488281	BCR_A	-0.4248			
2007	67	3564063	BCR_B	-0.0218	-1.4648	1.4648	
2007	69	3488317	SR	-1.4874			
2009	70	100000064622	SR	-1.7223			
2009	71	100000065148	SR	0.5252			
2008	77	3516327	BCR_A	1.1725			
2008	78	3564005	BCR_B	0.2939	-0.8570	0.8570	
2009	79	100000064624	SR	-1.0746			
2010	2	100000065090	SR	-1.9815			
2010	5	100000028429	SR	0.5836			
2010	7	100000028374	ECR_A	0.2351			
2010	8	3595487	ECR_B	0.3640	-1.3201	0.1828	1.1373
2010	13	100000028410	SR	-1.3212			
2010	21	3492071	SR	0.2299			
2010	22	3516627	BCR_A	0.7075			
2010	23	3564006	BCR_B	1.6275	-2.0510	2.0510	
2010	24	100000004453	SR	0.8127			
2010	28	100000065065	SR	1.4833			

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2010	29	3488297	SR	-0.6249			
2010	31	100000028363	BCR_A	-0.0837			
2010	32	3595491	BCR_B	0.3055	-0.3163	0.3163	
2010	33	100000064614	SR	-0.6463			
2010	44	100000094472	SR	1.1803			
2010	46	100000028417	SR	-1.3160			
2010	47	100000022498	BCR_A	0.3903			
2010	48	3595494	BCR_B	0.7504	-0.7203	0.7203	
2010	49	3492094	SR	-0.1972			
2010	52	100000064637	SR	0.3609			
2010	53	100000094475	SR	-0.4276			
2010	59	100000022492	BCR_A	-0.1987			
2010	60	3595479	BCR_B	0.9807	-0.1290	0.1290	
2010	62	100000064783	SR	0.7013			
2010	66	3488281	BCR_A	0.1382			
2010	67	3564063	BCR_B	0.3369	-1.3775	1.3775	
2010	69	3488317	SR	-1.6344			
2010	70	100000064622	SR	-1.4157			
2010	71	100000065148	SR	0.5523			
2010	77	3516327	BCR_A	1.0327			
2010	78	3564005	BCR_B	0.2748	-0.9394	0.9394	
2010	79	100000064624	SR	-0.6630			

Note. These Rasch difficulties were based on a common scale.

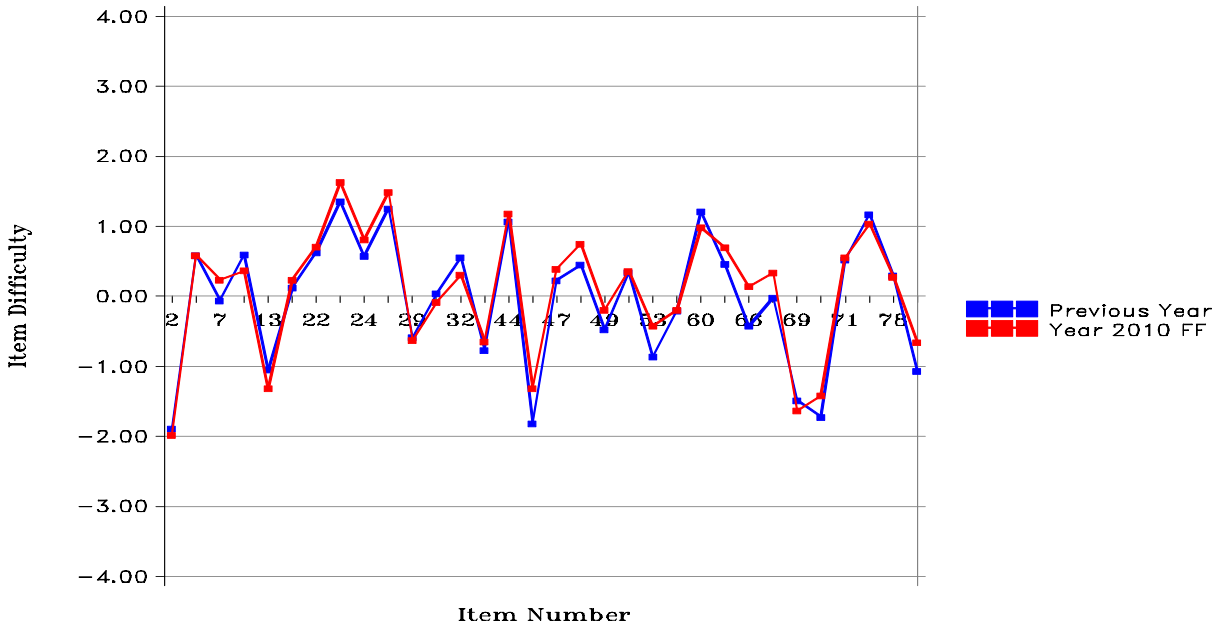

Figure 1.10 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 6 Form F

Table 1.62 P-Value Comparisons of Core Items for Previous Year vs. Year 2010: Grade 7 Form A

Item CID	Previous Year	Y10 FA	Item CID	Previous Year	Y10 FA
100000064084	0.67	0.66	**3517706**	**0.47**	**0.52**
100000026806	**0.50**	**0.51**	**3564025**	**0.28**	**0.32**
3595381	**0.60**	**0.61**	100000026814	0.40	0.43
100000064025	0.71	0.66	100000012782	0.90	0.88
100000026799	0.60	0.58	**100000012779**	**0.57**	**0.50**
100000004159	0.53	0.51	**3595378**	**0.38**	**0.33**
100000063929	0.45	0.33	100000063916	0.41	0.38
100000026808	**0.30**	**0.34**	3547772	0.21	0.29
3595374	**0.33**	**0.37**	**100000004147**	**0.72**	**0.74**
100000004154	0.68	0.63	**3595388**	**0.33**	**0.32**
100000026792	0.65	0.59	100000063988	0.59	0.66
100000063933	0.73	0.71	3492166	0.23	0.25
3517818	**0.33**	**0.30**	100000018133	0.31	0.30
3564023	**0.38**	**0.30**	**3517693**	**0.16**	**0.14**
100000123045	0.62	0.59	**3564028**	**0.45**	**0.45**
3547642	0.72	0.72	3492162	0.71	0.72
100000018124	0.37	0.34	100000070659	0.36	0.39

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item or Extended Constructed Response (ECR) item.

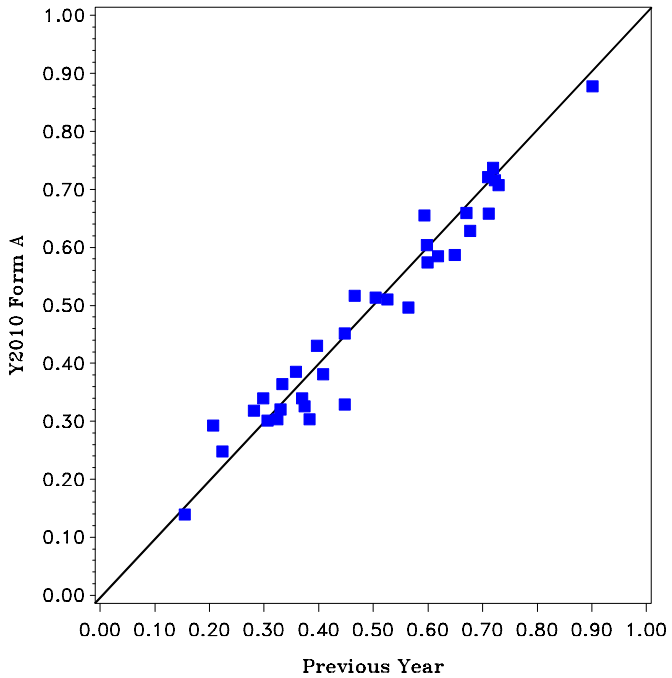

Table 1.63 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 7 Form A

						Score-Point Distribution (%)				
Year	Item CID	Item Type	N	Mean	SD	0	1	2	3	Omit
2008	100000026806	BCR	2,353	0.50	0.50	47.34	50.49			2.17
2008	3595381	BCR	2,353	1.20	0.70	13.64	46.45	36.72		3.19
2008	100000026808	ECR	2,635	0.30	0.46	67.40	29.94			2.66
2008	3595374	ECR	2,635	1.00	1.15	45.96	14.84	19.70	15.41	4.10
2007	3517818	BCR	32,264	0.33	0.47	61.95	32.55			5.50
2007	3564023	BCR	32,264	0.77	0.58	23.33	61.20	7.86		7.61
2007	3517706	BCR	32,000	0.47	0.50	45.48	46.67			7.85
2007	3564025	BCR	32,000	0.57	0.71	48.26	30.38	13.08		8.28
2009	100000012779	ECR	10,425	0.57	0.50	41.53	56.53			1.94
2009	3595378	ECR	10,425	1.13	0.76	17.57	52.59	24.00	0.04	1.83
2008	100000004147	ECR	2,646	0.72	0.45	25.43	72.03			2.53
2008	3595388	ECR	2,646	0.99	0.89	30.61	36.55	23.24	5.40	4.20
2007	3517693	BCR	32,000	0.16	0.36	76.92	15.57			7.51
2007	3564028	BCR	32,000	0.90	0.70	21.54	49.89	19.95		8.62
2010	100000026806	BCR	30,039	0.51	0.50	44.95	51.41			3.62
2010	3595381	BCR	30,039	1.21	0.69	13.02	48.46	36.31		2.10
2010	100000026808	ECR	30,039	0.34	0.47	62.03	34.09			3.82
2010	3595374	ECR	30,039	1.10	1.13	39.73	15.48	24.60	14.95	4.94
2010	3517818	BCR	30,039	0.30	0.46	58.44	30.40			11.15
2010	3564023	BCR	30,039	0.61	0.55	29.77	54.08	3.44		12.69
2010	3517706	BCR	30,039	0.52	0.50	43.43	51.74			4.55
2010	3564025	BCR	30,039	0.64	0.73	45.16	33.66	15.13		5.55
2010	100000012779	ECR	30,039	0.50	0.50	47.19	49.72			3.06
2010	3595378	ECR	30,039	0.98	0.75	23.26	52.56	18.40	2.88	2.75
2010	100000004147	ECR	30,039	0.74	0.44	24.34	73.83			1.75
2010	3595388	ECR	30,039	0.96	0.85	31.03	39.69	22.48	3.92	2.74
2010	3517693	BCR	30,039	0.14	0.35	82.96	13.98			2.93
2010	3564028	BCR	30,039	0.91	0.70	26.64	50.24	20.16		2.82

Table 1.64 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 7 Form A

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2009	4	100000064084	SR	-0.4313			
2008	5	100000026806	BCR_A	0.3784			
2008	6	3595381	BCR_B	-0.2943	-1.3144	1.3144	
2009	9	100000064025	SR	-0.6135			
2008	11	100000026799	SR	0.0661			
2008	17	100000004159	SR	0.4045			
2009	21	100000063929	SPR	0.8469			
2008	22	100000026808	ECR_A	1.4007			
2008	23	3595374	ECR_B	1.0873	-0.1344	-0.6583	0.7928
2008	24	100000004154	SPR	-0.6588			
2008	25	100000026792	SPR	-0.3043			
2009	26	100000063933	SPR	-0.9511			
2007	27	3517818	BCR_A	1.1456			
2007	28	3564023	BCR_B	1.1416	-2.2500	2.2500	
2009	29	100000123045	SR	-0.1139			
2009	38	3547642	SPR	-0.8342			
2008	39	100000018124	SPR	1.3694			
2007	40	3517706	BCR_A	0.2566			
2007	41	3564025	BCR_B	1.4706	-0.8554	0.8554	
2008	42	100000026814	SR	1.1042			
2008	44	100000012782	SR	-2.1659			
2009	45	100000012779	ECR_A	0.3215			
2009	46	3595378	ECR_B	1.5606	-2.5425	0.2773	2.2652
2009	47	100000063916	SPR	1.0818			
2007	48	3547772	SPR	2.0618			
2008	53	100000004147	ECR_A	-0.8842			
2008	54	3595388	ECR_B	1.5355	-1.7269	-0.1094	1.8363
2009	55	100000063988	SPR	-0.1230			
2007	56	3492166	SPR	1.8994			
2009	57	100000018133	SPR	1.6464			
2007	67	3517693	BCR_A	2.4839			
2007	68	3564028	BCR_B	0.4116	-1.5038	1.5038	
2008	71	3492162	SR	-0.6499			
2009	78	100000070659	SPR	1.2870			
2010	4	100000064084	SR	-0.3144			
2010	5	100000026806	BCR_A	0.4422			
2010	6	3595381	BCR_B	-0.1940	-1.3372	1.3372	
2010	9	100000064025	SR	-0.3181			
2010	11	100000026799	SR	0.1445			
2010	17	100000004159	SR	0.5166			
2010	21	100000063929	SPR	1.4038			
2010	22	100000026808	ECR_A	1.4517			

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2010	23	3595374	ECR_B	1.2670	-0.1693	-0.8031	0.9724
2010	24	100000004154	SPR	-0.2708			
2010	25	100000026792	SPR	-0.0337			
2010	26	100000063933	SPR	-0.8874			
2010	27	3517818	BCR_A	1.5333			
2010	28	3564023	BCR_B	2.1002	-2.4025	2.4025	
2010	29	100000123045	SR	-0.1001			
2010	38	3547642	SPR	-0.7284			
2010	39	100000018124	SPR	1.4308			
2010	40	3517706	BCR_A	0.3815			
2010	41	3564025	BCR_B	1.5390	-0.7818	0.7818	
2010	42	100000026814	SR	0.9408			
2010	44	100000012782	SR	-2.0093			
2010	45	100000012779	ECR_A	0.5110			
2010	46	3595378	ECR_B	1.8500	-2.6284	0.3759	2.2526
2010	47	100000063916	SPR	1.1933			
2010	48	3547772	SPR	1.7435			
2010	53	100000004147	ECR_A	-0.8589			
2010	54	3595388	ECR_B	1.8264	-1.9672	-0.1313	2.0984
2010	55	100000063988	SPR	-0.3417			
2010	56	3492166	SPR	2.0183			
2010	57	100000018133	SPR	1.6854			
2010	67	3517693	BCR_A	2.9771			
2010	68	3564028	BCR_B	0.8278	-1.3678	1.3678	
2010	71	3492162	SR	-0.7170			
2010	78	100000070659	SPR	1.1503			

Note. These Rasch difficulties were based on a common scale.

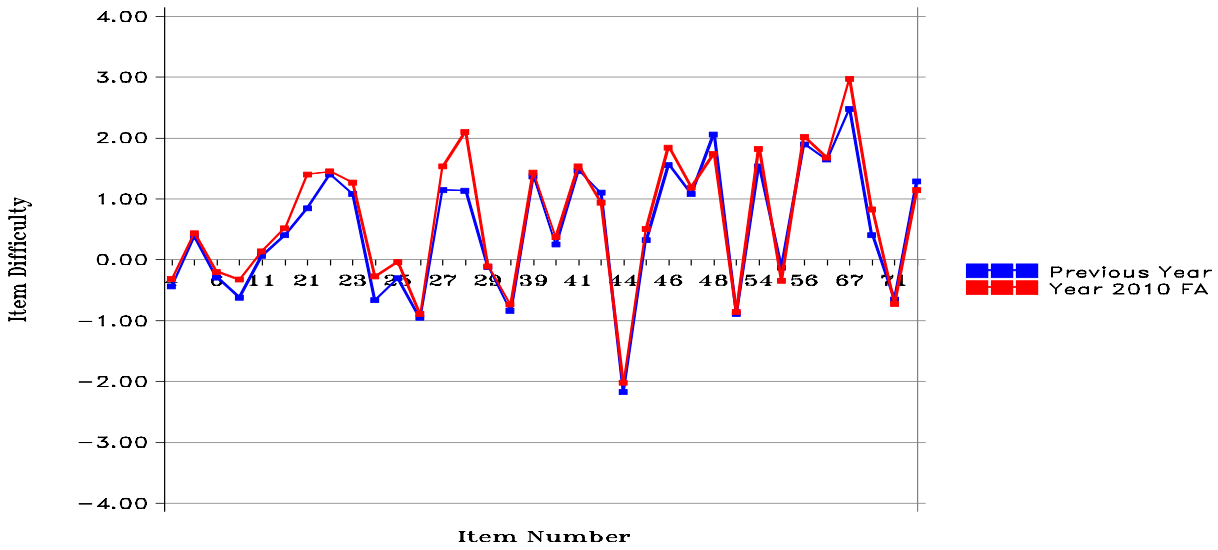

Figure 1.11 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 7 Form A

Table 1.65 P-Value Comparisons of Core Items for Previous Year vs. Year 2010: Grade 7 Form F

Item CID	Previous Year	Y10 FF	Item CID	Previous Year	Y10 FF
3500142	0.52	0.64	**3517706**	**0.47**	**0.53**
100000026820	**0.49**	**0.52**	**3564025**	**0.28**	**0.33**
3595398	**0.49**	**0.55**	100000026814	0.40	0.44
3487535	0.60	0.62	3500139	0.71	0.77
100000026799	0.60	0.61	**100000012779**	**0.57**	**0.52**
3487612	0.75	0.82	**3595378**	**0.38**	**0.35**
100000064095	0.40	0.40	100000064021	0.46	0.45
3491692	**0.47**	**0.46**	3547675	0.59	0.59
3564159	**0.48**	**0.47**	**100000027967**	**0.47**	**0.34**
3487614	0.71	0.72	**3595392**	**0.32**	**0.36**
3492176	0.71	0.66	100000004173	0.44	0.40
100000064098	0.80	0.74	100000070736	0.33	0.37
100000004178	**0.34**	**0.38**	100000064052	0.17	0.19
3595400	**0.41**	**0.40**	**3517693**	**0.16**	**0.13**
100000123045	0.62	0.59	**3564028**	**0.45**	**0.45**
3513630	0.68	0.70	100000012794	0.44	0.43
100000018121	0.29	0.25	3492164	0.35	0.34

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item or Extended Constructed Response (ECR) item.

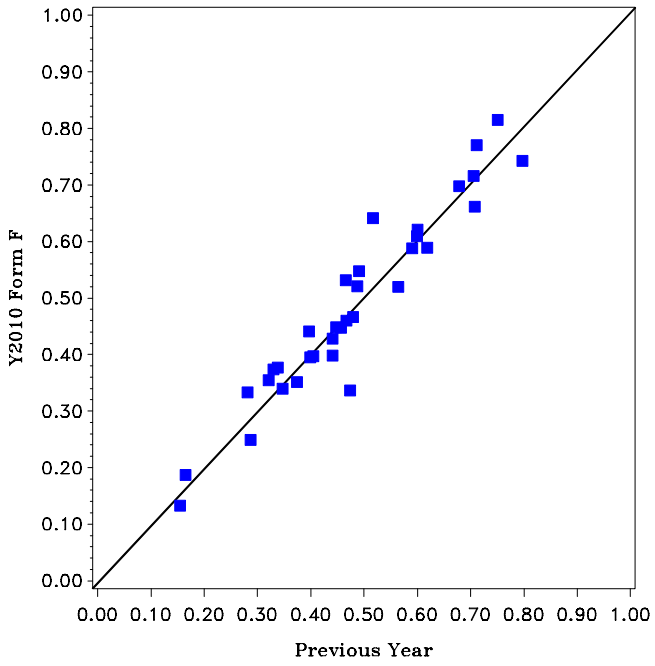

Table 1.66 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 7 Form F

						Score-Point Distribution (%)				
Year	Item CID	Item Type	N	Mean	SD	0	1	2	3	Omit
2008	100000026820	BCR	2,285	0.49	0.50	49.19	48.84			1.97
2008	3595398	BCR	2,285	0.98	0.71	23.24	49.15	24.55		3.06
2009	3491692	ECR	10,238	0.47	0.50	49.92	46.81			3.27
2009	3564159	ECR	10,238	1.44	1.11	19.37	29.20	21.36	24.01	6.06
2008	100000004178	BCR	2,596	0.34	0.47	62.67	33.86			3.47
2008	3595400	BCR	2,596	0.81	0.82	40.72	28.51	26.31		4.47
2007	3517706	BCR	32,000	0.47	0.50	45.48	46.67			7.85
2007	3564025	BCR	32,000	0.57	0.71	48.26	30.38	13.08		8.28
2009	100000012779	ECR	10,425	0.57	0.50	41.53	56.53			1.94
2009	3595378	ECR	10,425	1.13	0.76	17.57	52.59	24.00	0.04	1.83
2008	100000027967	ECR	2,569	0.47	0.50	50.56	47.49			1.95
2008	3595392	ECR	2,569	0.97	0.89	32.89	34.99	24.06	4.52	3.54
2007	3517693	BCR	32,000	0.16	0.36	76.92	15.57			7.51
2007	3564028	BCR	32,000	0.90	0.70	21.54	49.89	19.95		8.62
2010	100000026820	BCR	29,140	0.52	0.50	46.06	52.19			1.72
2010	3595398	BCR	29,140	1.10	0.68	16.05	53.08	28.28		2.50
2010	3491692	ECR	29,140	0.46	0.50	49.46	46.07			4.34
2010	3564159	ECR	29,140	1.40	1.09	19.10	30.55	21.98	21.87	6.25
2010	100000004178	BCR	29,140	0.38	0.48	53.42	37.79			8.79
2010	3595400	BCR	29,140	0.80	0.84	37.70	25.58	27.04		9.68
2010	3517706	BCR	29,140	0.53	0.50	42.42	53.25			4.12
2010	3564025	BCR	29,140	0.67	0.73	43.55	35.21	15.82		5.07
2010	100000012779	ECR	29,140	0.52	0.50	45.40	52.03			2.53
2010	3595378	ECR	29,140	1.06	0.78	21.15	51.20	21.23	3.99	2.34
2010	100000027967	ECR	29,140	0.34	0.47	64.31	33.73			1.90
2010	3595392	ECR	29,140	1.07	0.87	26.14	40.07	25.59	5.19	2.90
2010	3517693	BCR	29,140	0.13	0.34	83.44	13.42			3.04
2010	3564028	BCR	29,140	0.90	0.70	26.95	50.18	19.88		2.89

Table 1.67 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 7 Form F

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2007	4	3500142	SR	0.0659			
2008	5	100000026820	BCR_A	0.4611			
2008	6	3595398	BCR_B	0.4245	-1.3632	1.3632	
2008	9	3487535	SR	-0.1988			
2008	11	100000026799	SR	0.0661			
2007	17	3487612	SR	-1.3047			
2009	21	100000064095	SPR	1.1350			
2009	22	3491692	ECR_A	0.7615			
2009	23	3564159	ECR_B	0.5864	-1.1860	0.4071	0.7788
2007	24	3487614	SPR	-1.1139			
2008	25	3492176	SPR	-0.5940			
2009	26	100000064098	SPR	-1.3514			
2008	27	100000004178	BCR_A	1.3668			
2008	28	3595400	BCR_B	0.9231	-0.4262	0.4262	
2009	29	100000123045	SR	-0.1139			
2008	38	3513630	SPR	-0.5989			
2008	39	100000018121	SPR	1.8207			
2007	40	3517706	BCR_A	0.2566			
2007	41	3564025	BCR_B	1.4706	-0.8554	0.8554	
2008	42	100000026814	SR	1.1042			
2007	44	3500139	SR	-1.1300			
2009	45	100000012779	ECR_A	0.3215			
2009	46	3595378	ECR_B	1.5606	-2.5425	0.2773	2.2652
2009	47	100000064021	SPR	0.7333			
2007	48	3547675	SPR	-0.4714			
2008	53	100000027967	ECR_A	0.5688			
2008	54	3595392	ECR_B	1.6678	-1.7413	-0.2810	2.0223
2008	55	100000004173	SPR	0.6732			
2009	56	100000070736	SPR	1.4414			
2009	57	100000064052	SPR	2.5900			
2007	67	3517693	BCR_A	2.4839			
2007	68	3564028	BCR_B	0.4116	-1.5038	1.5038	
2008	71	100000012794	SR	0.8596			
2007	78	3492164	SPR	0.9591			
2010	4	3500142	SR	-0.2119			
2010	5	100000026820	BCR_A	0.4402			
2010	6	3595398	BCR_B	0.2177	-1.5084	1.5084	
2010	9	3487535	SR	-0.0824			
2010	11	100000026799	SR	0.1445			
2010	17	3487612	SR	-1.4333			
2010	21	100000064095	SPR	1.1139			
2010	22	3491692	ECR_A	0.7610			

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2010	23	3564159	ECR_B	0.6508	-1.3030	0.3303	0.9727
2010	24	3487614	SPR	-0.8015			
2010	25	3492176	SPR	-0.4784			
2010	26	100000064098	SPR	-1.1179			
2010	27	100000004178	BCR_A	1.1288			
2010	28	3595400	BCR_B	0.9665	-0.3186	0.3186	
2010	29	100000123045	SR	-0.1001			
2010	38	3513630	SPR	-0.5803			
2010	39	100000018121	SPR	2.0178			
2010	40	3517706	BCR_A	0.3815			
2010	41	3564025	BCR_B	1.5390	-0.7818	0.7818	
2010	42	100000026814	SR	0.9408			
2010	44	3500139	SR	-1.0226			
2010	45	100000012779	ECR_A	0.5110			
2010	46	3595378	ECR_B	1.8500	-2.6284	0.3759	2.2526
2010	47	100000064021	SPR	0.8403			
2010	48	3547675	SPR	0.0400			
2010	53	100000027967	ECR_A	1.5638			
2010	54	3595392	ECR_B	1.6174	-2.0279	-0.0882	2.1160
2010	55	100000004173	SPR	1.1111			
2010	56	100000070736	SPR	1.2585			
2010	57	100000064052	SPR	2.4514			
2010	67	3517693	BCR_A	2.9771			
2010	68	3564028	BCR_B	0.8278	-1.3678	1.3678	
2010	71	100000012794	SR	0.9999			
2010	78	3492164	SPR	1.4419			

Note. These Rasch difficulties were based on a common scale.

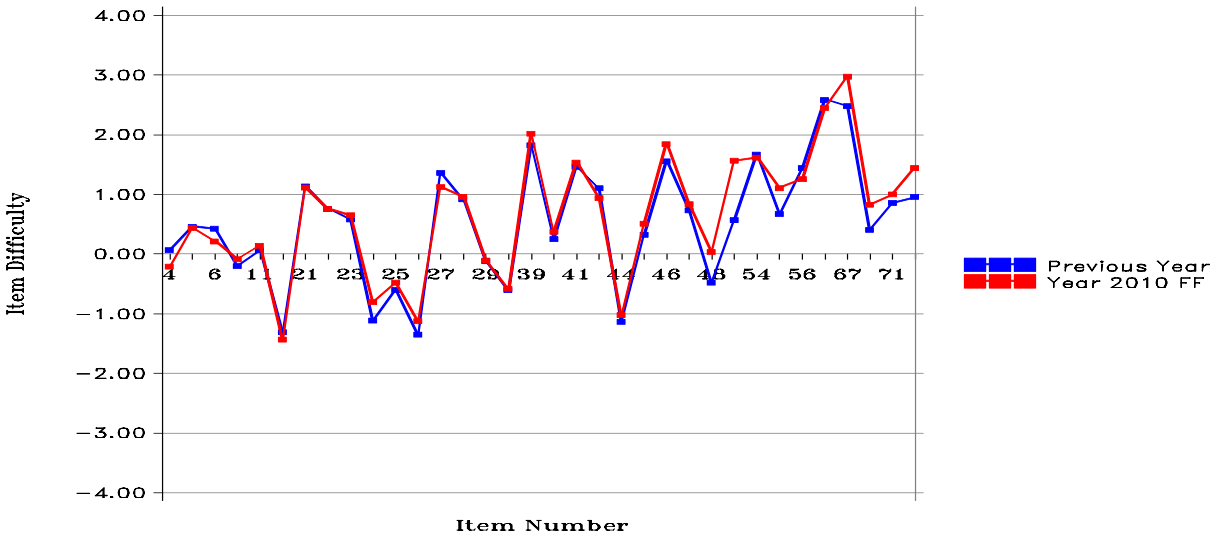

Figure 1.12 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 7 Form F

Table 1.68 P-Value Comparisons of Core Items for Previous Year vs. Year 2010: Grade 8 Form A

Item CID	Previous Year	Y10 FA	Item CID	Previous Year	Y10 FA
100000043324	**0.42**	**0.46**	**3487626**	**0.25**	**0.24**
3595434	**0.40**	**0.44**	**3595433**	**0.49**	**0.46**
100000064981	0.65	0.62	3487724	0.38	0.38
3514161	0.23	0.24	**3514117**	**0.42**	**0.40**
100000012774	**0.28**	**0.26**	**3564111**	**0.48**	**0.47**
3595421	**0.52**	**0.48**	3487537	0.72	0.79
100000026779	0.67	0.68	**3487933**	**0.27**	**0.33**
100000018170	0.77	0.74	**3595416**	**0.29**	**0.29**
100000065006	0.78	0.75	100000064900	0.38	0.39
100000012737	0.39	0.39	100000071596	0.28	0.31
3487718	0.59	0.64	100000016549	0.83	0.84
3487637	**0.30**	**0.28**	100000018146	0.78	0.77
3595418	**0.56**	**0.57**	3500160	0.29	0.30
3514279	0.27	0.27	3500164	0.47	0.42
3487636	**0.21**	**0.21**	**100000012745**	**0.29**	**0.28**
3595428	**0.33**	**0.33**	**3595427**	**0.45**	**0.40**
3492058	0.64	0.46	3492031	0.34	0.35
3514131	0.39	0.43			

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item or Extended Constructed Response (ECR) item.

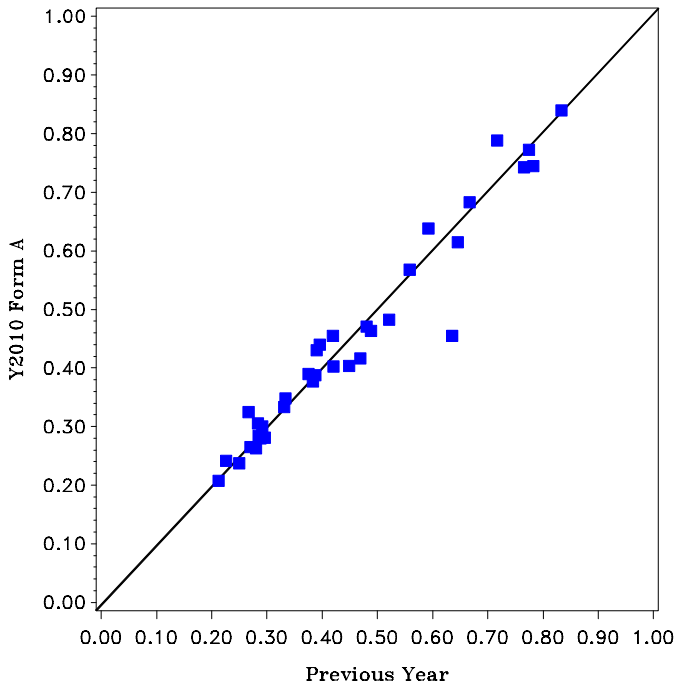

Table 1.69 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 8 Form A

						Score-Point Distribution (%)				
Year	Item CID	Item Type	N	Mean	SD	0	1	2	3	Omit
2008	100000043324	BCR	2,254	0.42	0.49	53.50	42.01			4.48
2008	3595434	BCR	2,254	0.79	0.91	46.45	13.53	32.92		7.10
2008	100000012774	ECR	2,055	0.28	0.45	69.29	28.13			2.58
2008	3595421	ECR	2,055	1.57	0.93	9.25	38.64	29.98	19.37	2.77
2008	3487637	BCR	2,279	0.30	0.46	68.54	29.66			1.80
2008	3595418	BCR	2,279	1.12	0.80	23.34	34.97	38.48		3.20
2008	3487636	BCR	2,314	0.21	0.41	76.92	21.31			1.77
2008	3595428	BCR	2,314	0.66	0.82	53.03	21.95	22.21		2.81
2008	3487626	ECR	2,254	0.25	0.43	73.29	25.07			1.64
2008	3595433	ECR	2,254	1.47	1.03	15.53	39.57	19.52	22.72	2.66
2009	3514117	BCR	11,185	0.42	0.49	52.91	42.16			4.93
2009	3564111	BCR	11,185	0.96	0.74	21.62	45.05	25.58		7.75
2008	3487933	ECR	2,260	0.27	0.44	70.00	26.81			3.19
2008	3595416	ECR	2,260	0.86	0.91	35.58	33.58	17.35	5.84	7.65
2008	100000012745	BCR	2,215	0.29	0.45	61.04	28.98			9.98
2008	3595427	BCR	2,215	0.90	0.90	35.49	18.69	35.62		10.20
2010	100000043324	BCR	30,331	0.46	0.50	50.64	45.57			3.70
2010	3595434	BCR	30,331	0.88	0.92	43.86	13.70	37.22		5.04
2010	100000012774	ECR	30,331	0.26	0.44	68.58	26.43			4.64
2010	3595421	ECR	30,331	1.45	1.00	14.49	30.57	31.62	17.10	5.74
2010	3487637	BCR	30,331	0.28	0.45	69.24	28.17			2.50
2010	3595418	BCR	30,331	1.14	0.80	22.85	33.87	39.95		3.14
2010	3487636	BCR	30,331	0.21	0.41	76.54	20.87			2.48
2010	3595428	BCR	30,331	0.67	0.81	51.19	23.84	21.47		3.31
2010	3487626	ECR	30,331	0.24	0.43	73.01	23.81			3.06
2010	3595433	ECR	30,331	1.39	1.12	22.46	29.43	20.31	23.13	4.28
2010	3514117	BCR	30,331	0.40	0.49	52.91	40.37			6.46
2010	3564111	BCR	30,331	0.94	0.79	25.32	36.65	28.89		8.61
2010	3487933	ECR	30,331	0.33	0.47	64.61	32.58			2.76
2010	3595416	ECR	30,331	0.86	0.89	36.93	31.84	20.59	4.22	6.06
2010	100000012745	BCR	30,331	0.28	0.45	59.54	28.06			12.23
2010	3595427	BCR	30,331	0.81	0.87	36.25	19.96	30.51		12.34

Table 1.70 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 8 Form A

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	3	100000043324	BCR_A	0.5470			
2008	4	3595434	BCR_B	0.5498	0.5281	-0.5281	
2009	6	100000064981	SR	-0.4984			
2009	15	3514161	SPR	1.7775			
2008	16	100000012774	ECR_A	1.4093			
2008	17	3595421	ECR_B	-0.1128	-2.0744	0.4235	1.6508
2008	18	100000026779	SPR	-0.8103			
2008	19	100000018170	SPR	-1.4299			
2009	20	100000065006	SPR	-1.4885			
2008	21	100000012737	SPR	0.6593			
2008	23	3487718	SR	-0.4267			
2008	24	3487637	BCR_A	1.2982			
2008	25	3595418	BCR_B	-0.2110	-0.6488	0.6488	
2009	26	3514279	SPR	1.5561			
2008	34	3487636	BCR_A	1.9471			
2008	35	3595428	BCR_B	1.0128	-0.0934	0.0934	
2007	36	3492058	SPR	-0.6456			
2007	37	3514131	SPR	0.5831			
2008	39	3487626	ECR_A	1.6956			
2008	40	3595433	ECR_B	0.1148	-1.7322	0.7638	0.9683
2007	43	3487724	SR	0.6740			
2009	44	3514117	BCR_A	0.7346			
2009	45	3564111	BCR_B	0.2987	-1.0752	1.0752	
2007	49	3487537	SR	-1.3157			
2008	54	3487933	ECR_A	1.4354			
2008	55	3595416	ECR_B	1.3557	-1.4592	0.0694	1.3898
2009	56	100000064900	SPR	0.9334			
2009	57	100000071596	SPR	1.3398			
2008	63	100000016549	SR	-1.8904			
2008	64	100000018146	SR	-1.4333			
2009	74	3500160	SPR	1.5068			
2009	75	3500164	SPR	0.5224			
2008	76	100000012745	BCR_A	1.3209			
2008	77	3595427	BCR_B	0.2848	0.1446	-0.1446	
2007	80	3492031	SR	0.8468			
2010	3	100000043324	BCR_A	0.3845			
2010	4	3595434	BCR_B	0.3994	0.6339	-0.6339	
2010	6	100000064981	SR	-0.4269			
2010	15	3514161	SPR	1.7253			
2010	16	100000012774	ECR_A	1.5604			
2010	17	3595421	ECR_B	0.1961	-1.5626	-0.0169	1.5796
2010	18	100000026779	SPR	-0.8747			

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2010	19	100000018170	SPR	-1.2527			
2010	20	100000065006	SPR	-1.2763			
2010	21	100000012737	SPR	0.7805			
2010	23	3487718	SR	-0.5452			
2010	24	3487637	BCR_A	1.4578			
2010	25	3595418	BCR_B	-0.1928	-0.6025	0.6025	
2010	26	3514279	SPR	1.5409			
2010	34	3487636	BCR_A	1.9670			
2010	35	3595428	BCR_B	1.0288	-0.1750	0.1750	
2010	36	3492058	SPR	0.4102			
2010	37	3514131	SPR	0.5542			
2010	39	3487626	ECR_A	1.7702			
2010	40	3595433	ECR_B	0.2890	-1.0401	0.3739	0.6662
2010	43	3487724	SR	0.8736			
2010	44	3514117	BCR_A	0.6586			
2010	45	3564111	BCR_B	0.2436	-0.8169	0.8169	
2010	49	3487537	SR	-1.5304			
2010	54	3487933	ECR_A	1.2142			
2010	55	3595416	ECR_B	1.5956	-1.5055	-0.3011	1.8066
2010	56	100000064900	SPR	0.8337			
2010	57	100000071596	SPR	1.2240			
2010	63	100000016549	SR	-1.9345			
2010	64	100000018146	SR	-1.3641			
2010	74	3500160	SPR	1.2752			
2010	75	3500164	SPR	0.5753			
2010	76	100000012745	BCR_A	1.3321			
2010	77	3595427	BCR_B	0.4893	0.0455	-0.0455	
2010	80	3492031	SR	0.9789			

Note. These Rasch difficulties were based on a common scale.

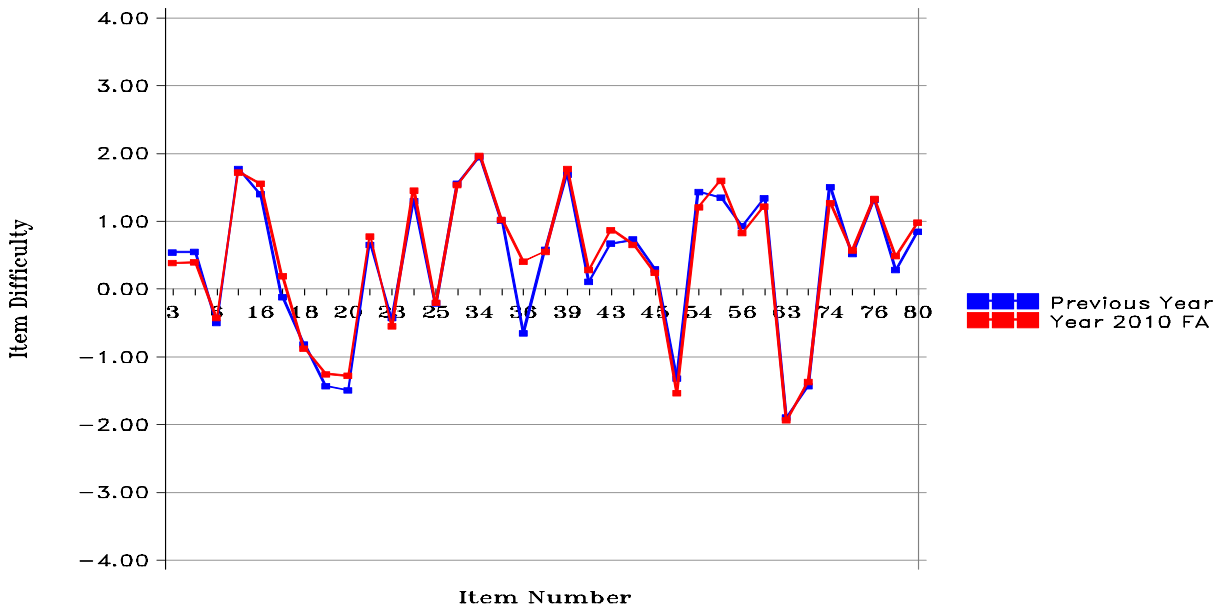

Figure 1.13 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 8 Form A

Table 1.71 P-Value Comparison of Core Items for Previous Year vs. Year 2010: Grade 8 Form F

Item CID	Previous Year	Y10 FF	Item CID	Previous Year	Y10 FF
3487624	**0.23**	**0.28**	**3514607**	**0.27**	**0.30**
3595415	**0.27**	**0.30**	**3564112**	**0.28**	**0.32**
100000064858	0.60	0.62	100000004089	0.83	0.82
100000064903	0.35	0.33	**3514117**	**0.42**	**0.41**
3487680	**0.34**	**0.31**	**3564111**	**0.48**	**0.48**
3564133	**0.37**	**0.35**	3487566	0.45	0.53
100000004110	0.70	0.69	**100000016550**	**0.66**	**0.77**
100000065028	0.75	0.70	**3595411**	**0.35**	**0.46**
100000064919	0.39	0.38	3513638	0.36	0.35
100000012737	0.39	0.39	100000064832	0.68	0.73
100000026755	0.65	0.65	100000065017	0.78	0.81
3514147	**0.38**	**0.45**	3487904	0.55	0.65
3564115	**0.33**	**0.38**	100000018176	0.32	0.26
100000026745	0.20	0.19	3492050	0.73	0.72
3514266	**0.33**	**0.35**	**100000012744**	**0.22**	**0.20**
3564120	**0.50**	**0.53**	**3595424**	**0.30**	**0.27**
3487907	0.45	0.41	3487719	0.82	0.78
3514114	0.41	0.45			

Note. Bold-faced number indicates that it is Brief Constructed Response (BCR) item or Extended Constructed Response (ECR) item.

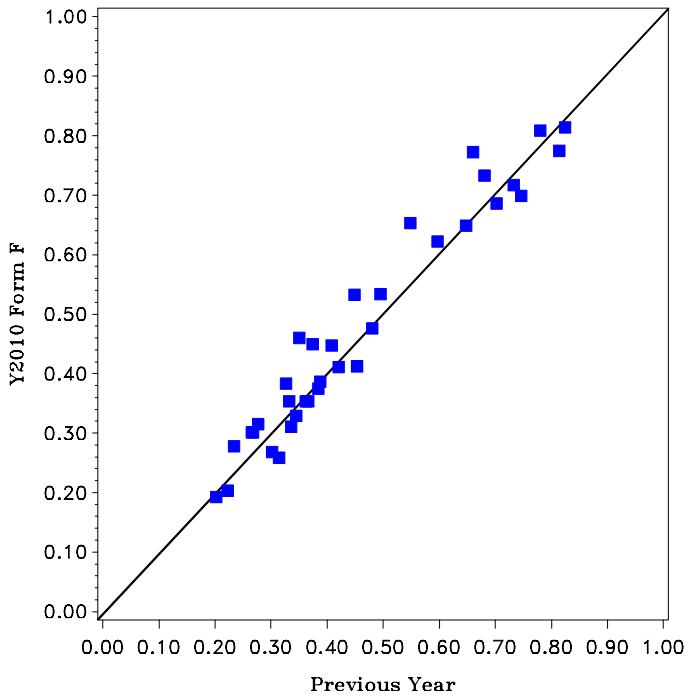

Table 1.72 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2010: Grade 8 Form F

						Score-Point Distribution (%)				
Year	Item CID	Item Type	N	Mean	SD	0	1	2	3	Omit
2008	3487624	BCR	2,260	0.23	0.42	69.47	23.45			7.08
2008	3595415	BCR	2,260	0.53	0.69	47.26	30.71	11.33		10.71
2009	3487680	BCR	11,185	0.34	0.47	62.32	33.71			3.97
2009	3564133	ECR	11,185	1.10	1.24	43.28	17.08	10.84	23.81	5.00
2007	3514147	ECR	32,480	0.38	0.48	56.35	37.55			6.09
2007	3564115	BCR	32,480	0.66	0.86	50.84	14.99	25.33		8.84
2008	3514266	BCR	31,743	0.33	0.47	60.28	33.36			6.36
2008	3564120	BCR	31,743	0.99	0.73	18.71	46.43	26.39		8.47
2008	3514607	ECR	32,318	0.27	0.44	64.36	26.92			8.72
2008	3564112	ECR	32,318	0.83	1.09	45.22	13.36	17.36	11.71	12.35
2009	3514117	BCR	11,185	0.42	0.49	52.91	42.16			4.93
2009	3564111	BCR	11,185	0.96	0.74	21.62	45.05	25.58		7.75
2008	100000016550	ECR	2,324	0.66	0.47	31.41	66.18			2.41
2008	3595411	ECR	2,324	1.05	0.95	29.22	32.10	25.09	7.70	5.90
2008	100000012744	BCR	2,055	0.22	0.42	71.39	22.34			6.28
2008	3595424	BCR	2,055	0.60	0.53	34.79	56.11	2.19		6.91
2010	3487624	BCR	29,748	0.28	0.45	66.94	27.88			5.08
2010	3595415	BCR	29,748	0.60	0.72	44.96	32.33	14.08		8.20
2010	3487680	ECR	29,748	0.31	0.46	62.89	31.17			5.75
2010	3564133	ECR	29,748	1.06	1.28	44.70	13.63	7.14	26.19	7.74
2010	3514147	BCR	29,748	0.45	0.50	50.66	45.04			4.23
2010	3564115	BCR	29,748	0.77	0.86	43.99	21.11	27.84		6.80
2010	3514266	BCR	29,748	0.35	0.48	60.89	35.47			3.50
2010	3564120	BCR	29,748	1.07	0.73	18.27	45.97	30.46		4.91
2010	3514607	ECR	29,748	0.30	0.46	61.86	30.22			7.75
2010	3564112	ECR	29,748	0.95	1.11	41.53	15.76	19.65	13.25	8.91
2010	3514117	BCR	29,748	0.41	0.49	52.61	41.24			5.90
2010	3564111	BCR	29,748	0.95	0.80	25.51	36.56	29.41		7.98
2010	100000016550	ECR	29,748	0.77	0.42	21.41	77.39			1.18
2010	3595411	ECR	29,748	1.38	0.93	18.02	30.98	37.69	10.68	2.55
2010	100000012744	BCR	29,748	0.20	0.40	68.58	20.38			10.47
2010	3595424	BCR	29,748	0.54	0.53	35.18	51.07	1.40		11.41

Table 1.73 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 8 Form F

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	3	3487624	BCR_A	1.6041			
2008	4	3595415	BCR_B	1.3736	-0.8303	0.8303	
2009	6	100000064858	SR	-0.2975			
2009	15	100000064903	SPR	1.0549			
2009	16	3487680	ECR_A	1.2610			
2009	17	3564133	ECR_B	0.8697	-0.1110	0.3415	-0.2305
2008	18	100000004110	SPR	-1.0594			
2009	19	100000065028	SPR	-1.1766			
2009	20	100000064919	SPR	0.8188			
2008	21	100000012737	SPR	0.6593			
2008	23	100000026755	SR	-0.6151			
2007	24	3514147	BCR_A	0.6730			
2007	25	3564115	BCR_B	0.7108	0.4113	-0.4113	
2008	26	100000026745	SPR	1.9488			
2008	34	3514266	BCR_A	1.1230			
2008	35	3564120	BCR_B	0.0332	-1.3167	1.3167	
2008	36	3487907	SPR	0.3647			
2008	37	3514114	SPR	0.5733			
2008	39	3514607	ECR_A	1.3428			
2008	40	3564112	ECR_B	1.1814	0.0189	-0.8420	0.8231
2008	43	100000004089	SR	-1.7778			
2009	44	3514117	BCR_A	0.7346			
2009	45	3564111	BCR_B	0.2987	-1.0752	1.0752	
2007	49	3487566	SR	0.3428			
2008	54	100000016550	ECR_A	-0.9409			
2008	55	3595411	ECR_B	0.9205	-1.5093	-0.2570	1.7663
2008	56	3513638	SPR	0.8139			
2009	57	100000064832	SPR	-0.9109			
2009	63	100000065017	SR	-1.4481			
2007	64	3487904	SR	-0.2926			
2008	74	100000018176	SPR	1.1230			
2007	75	3492050	SPR	-1.3521			
2008	76	100000012744	BCR_A	1.7840			
2008	77	3595424	BCR_B	2.1127	-2.5577	2.5577	
2007	80	3487719	SR	-1.9117			
2010	3	3487624	BCR_A	1.5161			
2010	4	3595415	BCR_B	1.3253	-0.8164	0.8164	
2010	6	100000064858	SR	-0.4014			
2010	15	100000064903	SPR	1.1908			
2010	16	3487680	ECR_A	1.2585			
2010	17	3564133	ECR_B	0.7715	0.2687	0.5204	-0.7890
2010	18	100000004110	SPR	-0.8713			

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2010	19	100000065028	SPR	-0.9745			
2010	20	100000064919	SPR	0.8664			
2010	21	100000012737	SPR	0.7805			
2010	23	100000026755	SR	-0.5724			
2010	24	3514147	BCR_A	0.4865			
2010	25	3564115	BCR_B	0.7613	0.0351	-0.0351	
2010	26	100000026745	SPR	2.1143			
2010	34	3514266	BCR_A	1.0607			
2010	35	3564120	BCR_B	-0.0377	-1.2136	1.2136	
2010	36	3487907	SPR	0.6857			
2010	37	3514114	SPR	0.4998			
2010	39	3514607	ECR_A	1.2960			
2010	40	3564112	ECR_B	1.1473	-0.2529	-0.6027	0.8555
2010	43	100000004089	SR	-1.6932			
2010	44	3514117	BCR_A	0.6586			
2010	45	3564111	BCR_B	0.2436	-0.8169	0.8169	
2010	49	3487566	SR	0.0785			
2010	54	100000016550	ECR_A	-1.3817			
2010	55	3595411	ECR_B	0.5644	-1.6158	-0.4646	2.0804
2010	56	3513638	SPR	1.0684			
2010	57	100000064832	SPR	-1.1628			
2010	63	100000065017	SR	-1.6215			
2010	64	3487904	SR	-0.5972			
2010	74	100000018176	SPR	1.6711			
2010	75	3492050	SPR	-1.1053			
2010	76	100000012744	BCR_A	2.0170			
2010	77	3595424	BCR_B	2.5056	-2.6114	2.6114	
2010	80	3487719	SR	-1.6419			

Note. These Rasch difficulties were based on a common scale.

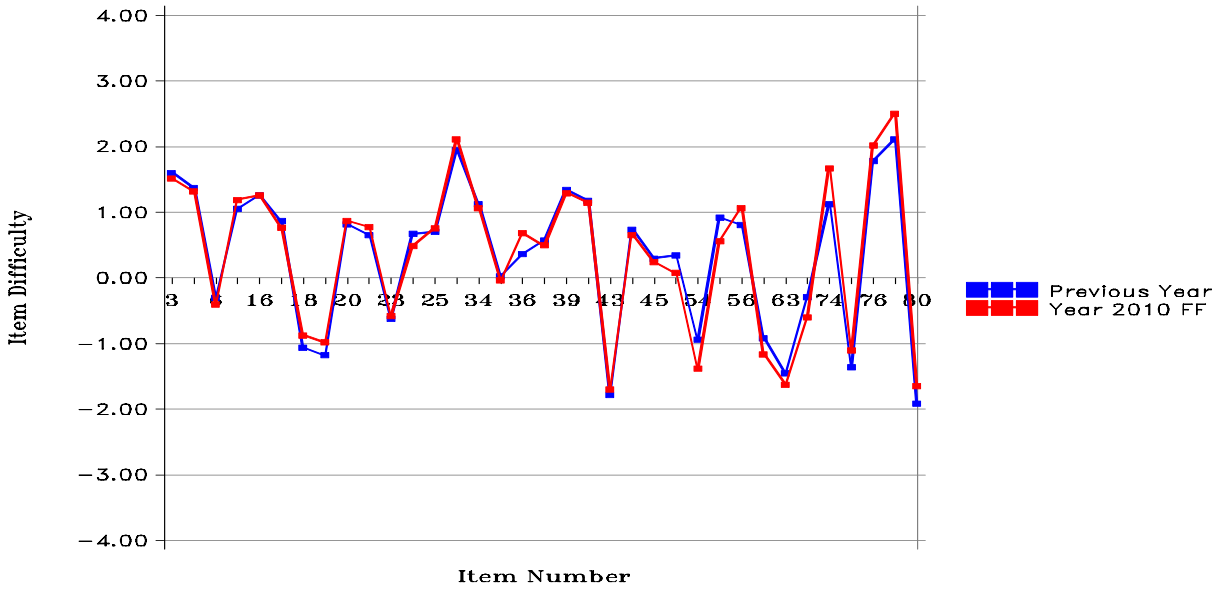

Figure 1.14 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2010: Grade 8 Form F

1.9 Linking, Equating, and Scaling Procedures of the 2010 MSA-Math

For the purpose of year-to-year linking and equating, we constructed a 2010 linking pool: only operational selected-response (SR) items (i.e., multiple-choice items) were included in the linking pool. It should be noted that these SR items appeared both in current and previous years' assessments and were used as either core or core link item in previous years' assessments (i.e., in any assessment before 2010). After setting up the linking pool, we conducted a stability check of linking items and decided which items should be excluded from or which items should remain in the linking pool. During the calibration and equating process, we kept and fixed the original operational Rasch item difficulty parameters of any linking items that remained through the stability check to put the 2010 assessment on a common scale. Accordingly, scale scores of the 2010 assessment were linked back to the 2006 assessment and all the scale scores of different years were comparable within each content and grade. It should be noted that Rasch recalibration was conducted using the 2006 MSA-Math data in 2007 due to the IRT model transition (i.e., from 3-PL to the Rasch). Detailed information on the 2006 Rasch recalibration and results can be obtained in the 2007 MSA-Math technical report.

Stratified Random Sampling Procedures

To select equating samples of grades 3, 4, and 5, a stratified random sampling method was applied to the 2010 state examinee population. To verify that the sample was representative of the statewide examinee population, the distributions of LEA, gender, and ethnicity of the 2010 sample were compared with those of the 2010 population. Appendix A, *The 2010 MSA-Math Stratified Random Sampling*, provides the results of 2010 sampling. The results indicated that the equating samples were well representative of the statewide examinee population in terms of LEA, gender, and ethnicity.

Robust Z Procedures

After selecting equating samples, each operational form was independently calibrated to estimate Rasch item difficulty of each item. Then Robust z values of all anchor items were calculated using the following calculations (South Carolina Department of Education, 2001):

- The mean and standard deviation of the linking pool's item difficulties for each operational form
- The ratio of the standard deviations between operational form A and form F
- The correlation between operational form A and F item difficulties
- The difference between operational form A and F for each item in the linking pool
- The mean of the differences calculated above
- The median of the differences calculated above
- The interquartile range of the differences calculated above
- The robust z is defined as (the difference between the test form1 and other test form item difficulty minus the median of the differences) / (interquartile range multiplied by 0.74).

Guidelines for Selecting Form-to-Form or Year-to-Year Linking Items

Once the above calculations were made, the following guidelines were followed in determining form-to-form or year-to-year common items used for Rasch linking and equating:

- Conform to the following "Protocol Criteria:" A correlation greater than 0.95 and a standard deviation ratio between 0.9 and 1.1. For example, use all the possible linking items as anchors if an original set of linking items meets these two criteria.
- Try not to include items with an absolute value of robust z exceeding 1.645.
- If one item difficulty on one form of the current year is eliminated from the linking pool, other item difficulties of the other forms should not be included.
- Should not eliminate more than 20 percent of the linking pool items.

Figure 1.15 depicts how we applied the anchor stability guidelines into the 2010 MSA-Math equating.

Form-to-Form Linking Procedures

The stability of the common items appearing on both operational forms was verified at each grade level:

- Calibrate the two operational test forms separately
- Calculate robust z values of Rasch item difficulties for forms A and F
- Correlate Rasch item difficulties between form A and form F
- Calculate standard deviation ratio between two forms

After examining the robust z values, correlation coefficient, and standard deviation ratio between the two separate calibrations, it was determined that the common item difficulties were consistent across the two forms for all items and could be included as form-to-form linking items in the fixed calibration of the two forms.

Year-to-Year Linking Procedures

The two 2010 operational forms included a set of year-to-year linking common items that appeared on both current and previous operational forms. We utilized the Rasch item fixed equating method for all of the operational items to be placed on a common scale within each grade.

The stability of the linking common items was evaluated using robust z values, correlation coefficients, and standard deviation ratios.

Tables 1.74 through 1.79 include Rasch item difficulties used for calculating robust z values, correlation coefficients, and standard deviations. Figures 1.16 through 1.27 depict item difficulty plots between current and previous years. It should be noted that the item difficulties of the 2010 operational forms were obtained from independent calibration, and those of previous assessments were on a common scale (i.e., linked to the 2006 assessment).

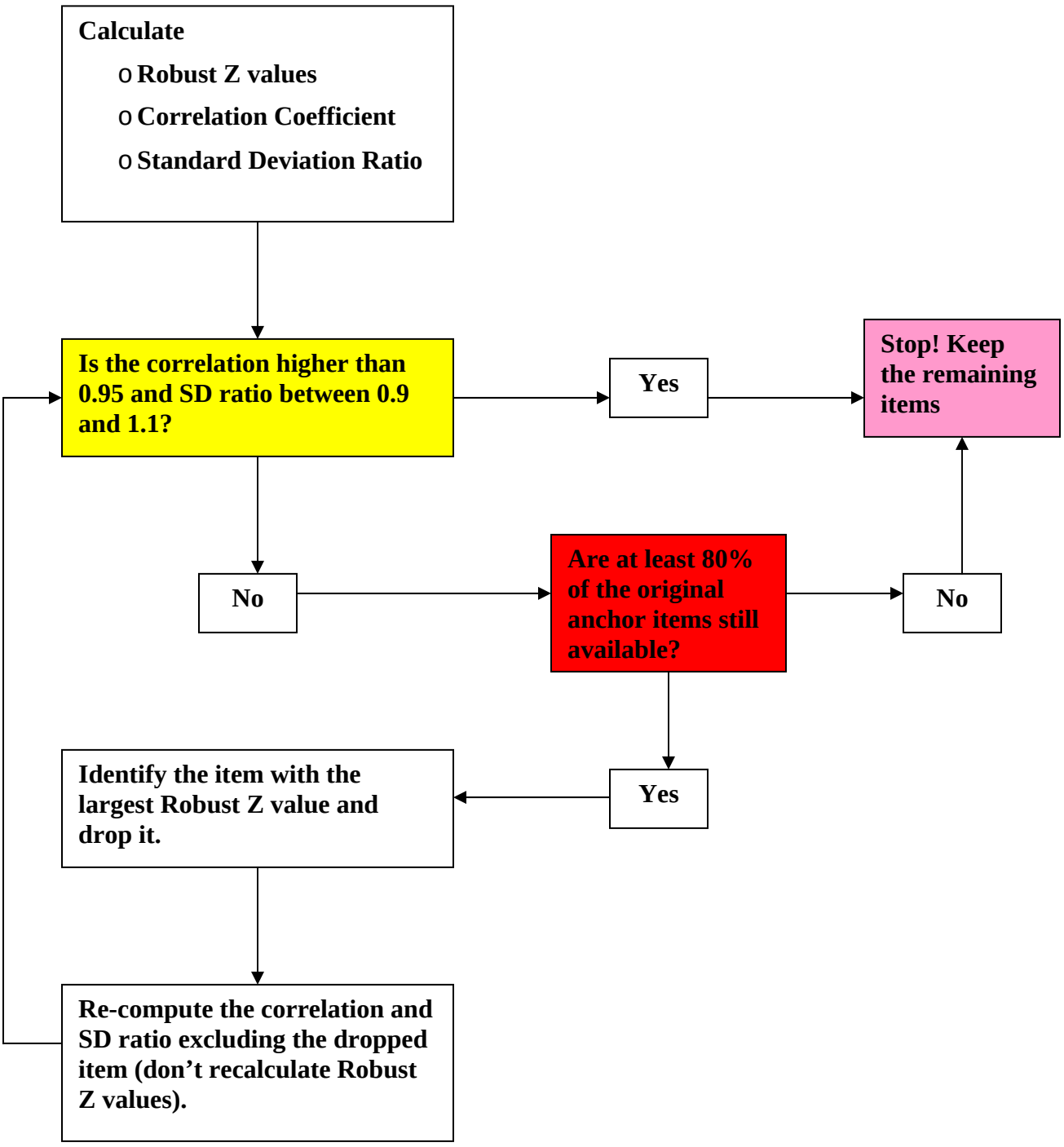

Figure 1.15 Anchor Evaluation Steps Chart for MSA-Math

Table 1.74 Core Linking Item Difficulties of Previous Year vs. Year 2010: Grade 3

Item Seq No.	Previous Form A	Y2010 Form A	Robust Z	Item Seq No.	Previous Form F	Y2010 Form F	Robust Z
1	0.9627	0.615	.6608	1	0.9627	0.6056	.7427
2	0.6288	0.0871	-.1583	2	0.6288	0.0413	-.3787
5	0.069	-0.4352	.0000	5	0.069	-0.4852	-.2166
6	1.5712	1.1152	.2035	6	1.5712	1.2043	.6950
7	0.5502	-0.2219	-1.1312	7	0.5502	-0.2565	-1.4455
13	1.3184	0.6694	-.6114	13	1.3184	0.7252	-.4064
14	-1.1315	-1.7321	-.4070	14	-1.1315	-1.9408	-1.4582
15	2.4187	1.6623	-1.0649	15	2.4187	1.6942	-1.0455
16	0.3981	-0.2813	-.7397	16	0.3981	-0.2002	-.4312
17	-0.3864	-1.182	-1.2304	17	-0.3864	-1.2561	-1.7522
18	-0.2038	-0.5165	.8086	18	-0.2038	-0.569	.7033
19	2.0077	1.3059	-.8343	19	2.0077	1.4999	.0092
20	-1.147	-1.5022	.6291	20	-1.147	-1.4288	1.1092
23	0.4123	-0.1425	-.2136	23	0.4123	-0.0974	.0000
31	-0.2784	-1.1928	-1.7320	31	-0.2784	-1.2082	-2.0447
32	0.5005	0.2982	1.2747	32	0.5005	0.2222	1.1262
33	1.3765	1.3319	1.9406	33	1.3765	1.215	1.6947
41	0.4861	-0.11	-.3880	41	0.4861	-0.0255	-.0092
48	-2.1822	-2.2978	1.6408	48	-2.1822	-2.0723	3.0157
49	-1.3667	-1.554	1.3381	49	-1.3667	-1.4347	2.1498
51	0.2953	-0.2284	-.0823	51	0.2953	-0.3038	-.4351
52	-0.6165	-0.9388	.7680	52	-0.6165	-1.0101	.5651
56	-0.6059	-1.1702	-.2538	56	-0.6059	-1.2018	-.4195
62	0.9229	0.3245	-.3977	62	0.9229	0.3555	-.2808
63	-0.2691	-0.5236	1.0543	63	-0.2691	-0.4813	1.4480
64	-0.0075	-0.4547	.2407	64	-0.0075	-0.5978	-.3923
65	-0.6248	-0.476	2.7572	65	-0.6248	-0.4547	3.3087
66	-0.5397	-0.6905	1.4922	66	-0.5397	-0.7741	1.3399
67	1.5719	0.7668	-1.2705	67	1.5719	0.8961	-.8084
68	0.0473	-0.2495	.8757	68	0.0473	-0.3774	.4137
69	0.0444	0.1536	2.5900	69	0.0444	0.1316	2.9052
71	0.4822	0.119	.5953	71	0.4822	-0.1106	-.4045
72	0.8475	0.0584	-1.2029	72	0.8475	0.1138	-1.0902
81	1.2021	0.7839	.3631	81	1.2021	0.6974	.0243
82	1.8021	1.2539	-.1858	82	1.8021	1.3478	.2696

Form Statistics

Form Statistics	Previous Year Form A	Y10 Form A	Previous Year Form F	Y10 Form F
Mean	.302	-.153	.302	-.158
SD	1.019	.946	1.019	.960

*Note: mean and standard deviation of Year 10 is calculated with freely calibrated estimates.

Correlation and Standard Deviation Ratio

Correlation Coefficient	.967	.965
SD Ratio	93%	94%

Values Used for Robust Z Statistics

Mean Diff	-.455	-.460
Median Diff	-.504	-.510
IQR Diff	.320	.278

Based on correlation coefficients and SD ratios, none of the linking common items were dropped from the linking pool.

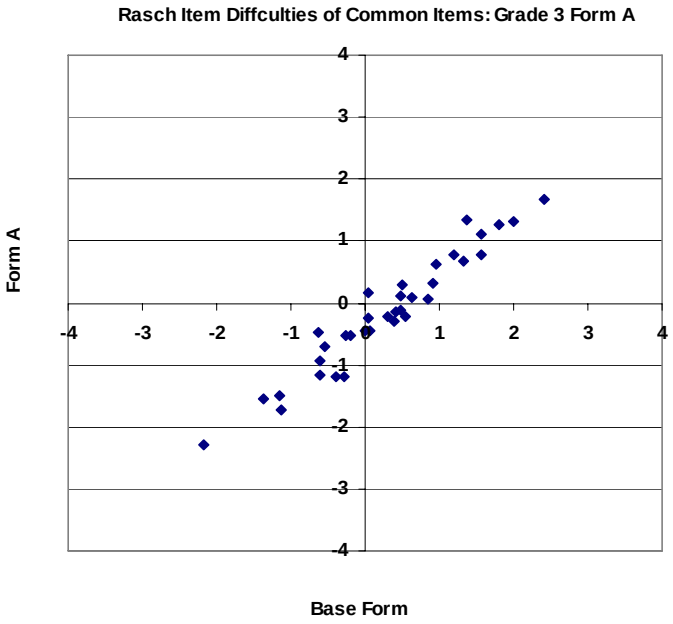

Figure 1.16 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 3 Form A

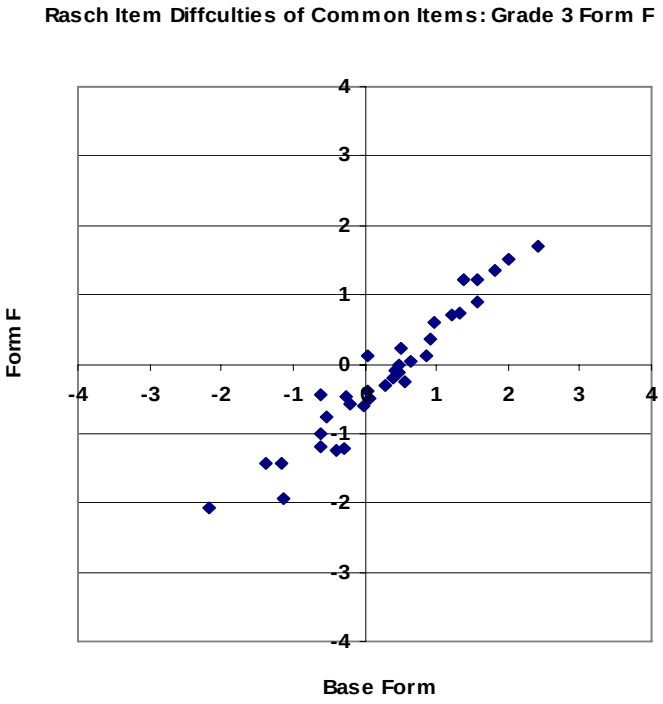

Figure 1.17 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 3 Form F

Table 1.75 Core Linking Item Difficulties of Previous Year vs. Year 2010: Grade 4

Item Seq No.	Previous Form A	Y2010 Form A	Robust Z	Item Seq No.	Previous Form F	Y2010 Form F	Robust Z
2	-0.799	-1.0164	-.6356	2	-0.799	-1.035	-.3795
3	1.757	1.5966	-.4333	3	1.757	1.4745	-.5226
6	0.1763	-0.4163	-1.9671	6	0.1763	-0.4848	-1.6877
7	-0.8522	-0.5925	1.0575	7	-0.8522	-0.7422	.6854
8	-1.055	-1.1633	-.2484	8	-1.055	-1.2443	-.2357
10	0.9009	0.7907	-.2552	10	0.9009	0.6152	-.5324
18	-0.4437	-0.3399	.5043	18	-0.4437	-0.5096	.1440
19	1.4979	1.2501	-.7435	19	1.4979	1.1047	-.8633
22	0.394	0.3557	.0000	22	0.394	0.1343	-.4524
23	-0.7461	-0.4319	1.2509	23	-0.7461	-0.5537	.9390
24	0.9747	1.4967	1.9883	24	0.9747	1.3698	1.5628
25	0.07969	0.0343	-.0252	25	0.07969	-0.0106	.0690
26	1.1949	1.2353	.2793	26	1.1949	1.024	-.1791
30	-0.2781	-0.2879	.1011	30	-0.2781	-0.4341	-.1333
32	-0.5617	-1.0275	-1.5171	32	-0.5617	-1.0884	-1.2741
33	-2.7781	-2.621	.6934	33	-2.7781	-2.5045	1.1889
47	-0.1077	-1.1799	-3.6690	47	-0.1077	-1.3004	-3.3238
49	-0.9767	-0.8156	.7076	49	-0.9767	-1.0137	.2330
50	0.9291	1.1867	1.0501	50	0.9291	1.0767	.8011
54	0.6901	0.8583	.7328	54	0.6901	0.8933	.9722
55	-0.4674	-0.2177	1.0220	55	-0.4674	-0.3308	.7672
62	1.0327	1.6145	2.2006	62	1.0327	1.586	2.0496
63	-0.2435	-0.2898	-.0284	63	-0.2435	-0.546	-.5841
64	-0.106	-0.5417	-1.4103	64	-0.106	-0.6489	-1.3240
66	**0.6281**	**1.6572**	**3.7879**	**66**	**0.6281**	**1.3408**	**2.5402**
67	-0.3619	-0.4311	-.1097	67	-0.3619	-0.4746	.0000
68	-2.3	-2.41	-.2544	68	-2.3	-2.8712	-1.4111
69	-0.9014	-0.787	.5419	69	-0.9014	-0.905	.3358
70	-2.8436	-3.0562	-.6185	70	-2.8436	-2.7692	.5758
71	-0.8156	-1.0433	-.6721	71	-0.8156	-1.1365	-.6407
78	-1.2169	-1.5857	-1.1729	78	-1.2169	-1.7108	-1.1732
80	-0.0118	0.2757	1.1562	80	-0.0118	0.1542	.8577
81	-0.1831	-0.1003	.4297	81	-0.1831	-0.2842	.0357

Form Statistics

Form Statistics	Previous Year Form A	Y10 Form A	Previous Year Form F	Y10 Form F
Mean	-.236	-.243	-.236	-.358
SD	1.101	1.219	1.101	1.188

*Note: mean and standard deviation of Year 10 is calculated with freely calibrated estimates.

Correlation and Standard Deviation Ratio

Correlation Coefficient	**.953**	**.949**
SD Ratio	**111%**	**108%**

Values Used for Robust Z Statistics

Mean Diff	-.006	-.122
Median Diff	-.038	-.113
IQR Diff	.381	.439

Based on correlation coefficients, SD ratios, robust z, and item difficulty plot, item number 66 appearing on both forms was dropped from the linking pool.

The following correlation coefficients and SD ratios were calculated after dropping that item:

Correlation with Base	**.961**	**.954**
SD Ratio	**107%**	**105%**

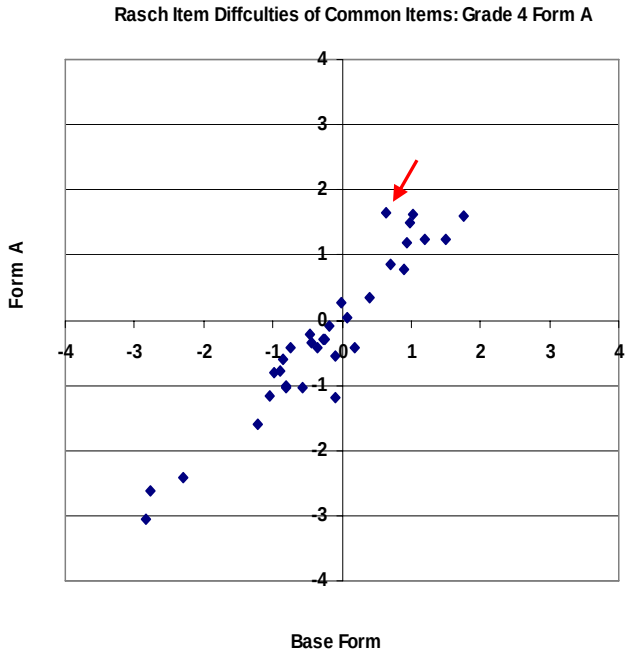

Figure 1.18 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 4 Form A

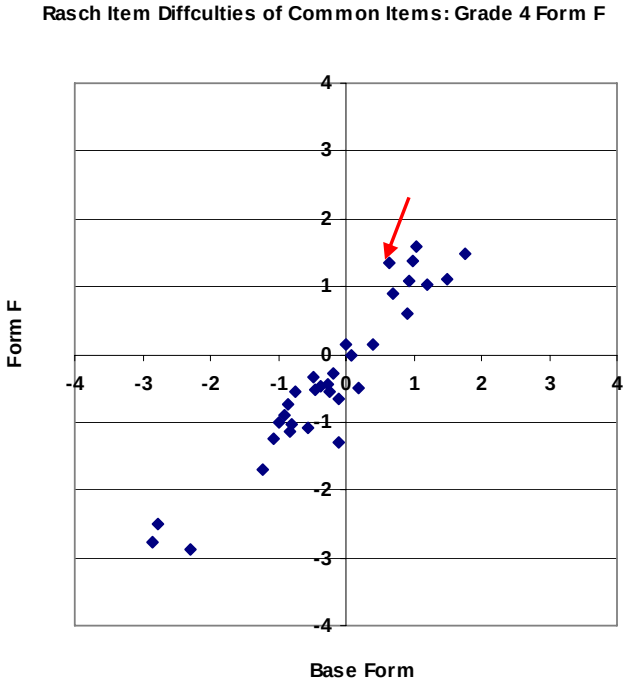

Figure 1.19 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 4 Form F

Table 1.76 Core Linking Item Difficulties of Previous Year vs. Year 2010: Grade 5

Item Seq No.	Previous Form A	Y2010 Form A	Robust Z	Item Seq No.	Previous Form F	Y2010 Form F	Robust Z
2	-1.0845	-1.3821	-.7145	2	-1.0845	-1.4565	-.7202
8	-1.3086	-2.1453	-2.9398	8	-1.3086	-2.3202	-3.4521
16	0.6094	0.7552	1.1158	16	0.6094	0.5947	.8060
18	0.179	-0.1068	-.6658	18	0.179	-0.1356	-.4750
19	-0.9093	-1.1127	-.3257	19	-0.9093	-1.0842	.1217
20	0.4459	0.4978	.7282	20	0.4459	0.5577	1.3463
21	0.3045	0.3457	.6840	21	0.3045	0.2593	.6757
23	0.335	-0.0155	-.9329	23	0.335	-0.0994	-.9867
26	-2.4397	-2.5478	.0677	26	-2.4397	-2.4309	.9064
27	0.2606	-0.0775	-.8817	27	0.2606	-0.1267	-.7855
28	-1.1318	-1.3703	-.4706	28	-1.1318	-1.6072	-1.1618
37	-0.8175	-1.2613	-1.3180	37	-0.8175	-1.4027	-1.6308
38	-0.331	-0.6018	-.6039	38	-0.331	-0.8486	-1.3421
39	1.6277	1.282	-.9131	39	1.6277	1.3021	-.5220
42	0.1548	0.2376	.8557	42	0.1548	0.0222	.3024
43	-1.1293	-1.0669	.7715	43	-1.1293	-1.1913	.6040
47	-1.2591	-1.3778	.0239	47	-1.2591	-1.5475	-.3631
48	-0.213	-0.2786	.2431	48	-0.213	-0.4732	-.2426
49	-1.4589	-1.6677	-.3480	49	-1.4589	-1.6675	-.0222
55	-0.6828	-0.4841	1.3341	55	-0.6828	-0.4565	1.8354
56	-0.5908	-0.3694	1.4278	56	-0.5908	-0.7461	.2055
58	-0.9439	-1.0742	-.0239	58	-0.9439	-1.3396	-.8214
60	-1.6307	-1.6103	.5981	60	-1.6307	-1.7131	.5168
61	-0.5025	-0.3871	.9903	61	-0.5025	-0.4849	.9440
64	-0.926	-1.2572	-.8532	64	-0.926	-1.3888	-1.1080
70	0.2463	0.1832	.2534	70	0.2463	0.0481	.0222
71	0.5581	0.3934	-.1659	71	0.5581	0.3603	.0239
72	1.0014	0.7609	-.4788	72	1.0014	0.7687	-.1252
82	-0.0717	-0.1606	.1470	82	-0.0717	-0.158	.5002
83	-0.6839	-0.6372	.7067	83	-0.6839	-0.6476	1.0238

Form Statistics

Form Statistics	Previous Year Form A	Y10 Form A	Previous Year Form F	Y10 Form F
Mean	-.413	-.551	-.413	-.647
SD	.881	.920	.881	.936

*Note: mean and standard deviation of Year 10 is calculated with freely calibrated estimates.

Correlation and Standard Deviation Ratio

Correlation Coefficient	**.970**	**.965**
SD Ratio	**104%**	**106%**

Values Used for Robust Z Statistics

Mean Diff	-.138	-.234
Median Diff	-.125	-.203
IQR Diff	.327	.316

Based on correlation coefficients and SD ratios, none of the linking common items were dropped from the linking pool.

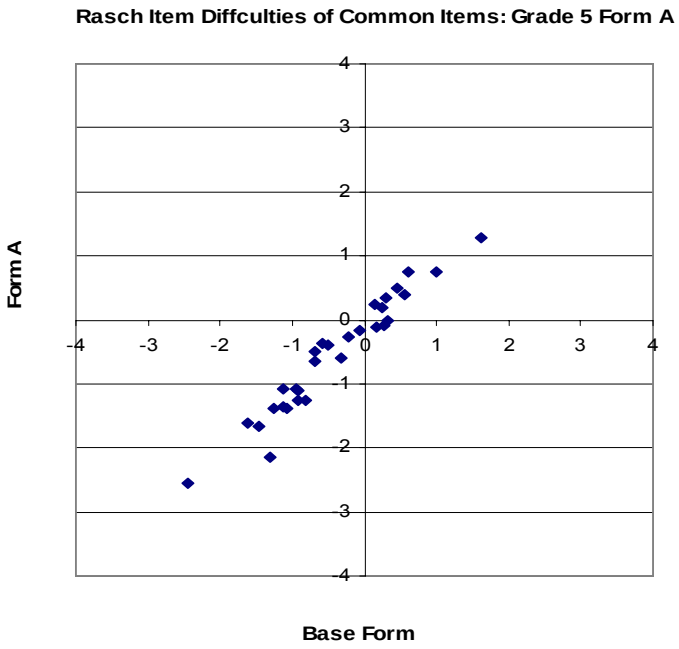

Figure 1.20 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 5 Form A

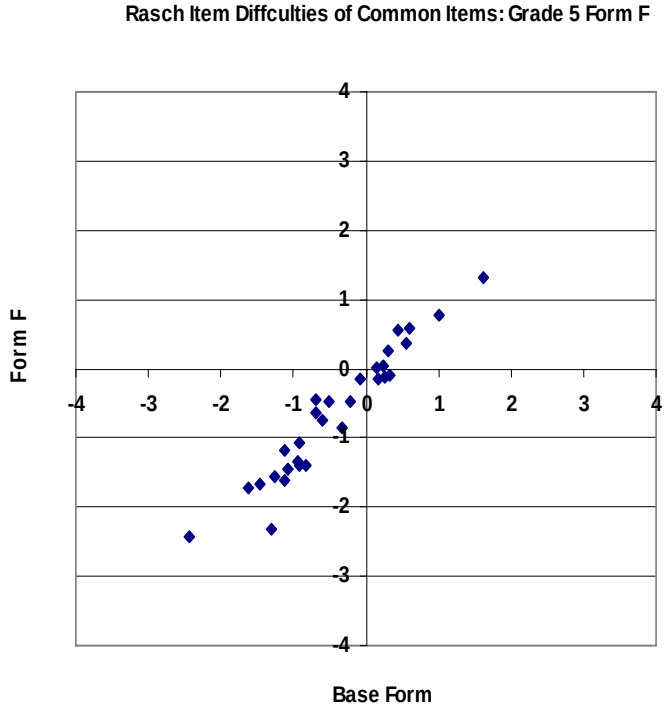

Figure 1.21 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 5 Form F

Table 1.77 Core Linking Item Difficulties of Previous Year vs. Year 2010: Grade 6

Item Seq No.	Previous Form A	Y2010 Form A	Robust Z	Item Seq No.	Previous Form F	Y2010 Form F	Robust Z
1	-1.2053	-1.789	-2.2240	1	-1.2053	-1.5836	-1.3112
3	0.6406	0.7206	1.9898	3	0.6406	0.8133	1.0468
4	0.6958	0.4597	-.0171	4	0.6958	0.6312	.0312
6	-0.2844	-0.564	-.2933	6	-0.2844	-0.4085	-.2234
9	**-0.319**	**-1.029**	**-3.0309**	**9**	**-0.319**	**-0.852**	**-1.9758**
10	-1.4432	-2.0445	-2.3357	10	-1.4432	-1.9057	-1.6716
11	-0.4703	-0.7208	-.1086	11	-0.4703	-0.6209	-.3368
12	0.3254	0.0591	-.2089	12	0.3254	0.182	-.3060
19	0.2409	-0.202	-1.3301	19	0.2409	-0.1597	-1.4067
20	0.4042	0.2433	.4603	20	0.4042	0.4074	.3214
25	**0.4777**	**0.7239**	**3.0449**	**25**	**0.4777**	**0.8175**	**1.7619**
26	-0.1396	-0.3928	-.1257	26	-0.1396	-0.2466	-.1502
27	0.2101	0.0023	.1625	27	0.2101	0.1028	-.1515
30	-0.6658	-0.7279	1.0876	30	-0.6658	-0.5676	.7279
34	1.1378	0.7221	-1.1574	34	1.1378	0.8195	-1.0545
35	-1.4702	-1.4094	1.8678	35	-1.4702	-1.3111	.9886
36	0.3674	0.1283	-.0362	36	0.3674	0.1884	-.4583
37	0.5144	0.2152	-.4178	37	0.5144	0.0648	-1.6164
38	-0.7259	-0.585	2.3764	38	-0.7259	-0.4978	1.2838
45	-0.9261	-1.1568	.0171	45	-0.9261	-0.732	1.1383
50	0.8786	0.7197	.4730	50	0.8786	0.9569	.6428
51	-0.0834	-0.1537	1.0355	51	-0.0834	-0.0194	.5816
54	0.2864	0.2075	.9809	54	0.2864	0.4976	1.2115
55	0.5885	0.7131	2.2729	55	0.5885	0.8091	1.2517
56	-0.2182	-0.7704	-2.0240	56	-0.2182	-0.7872	-2.1273
57	0.8821	0.7034	.3473	57	0.8821	0.8029	-.0312
58	-0.2639	-0.5121	-.0940	58	-0.2639	-0.3935	-.2469
61	0.283	0.0138	-.2273	61	0.283	0.2883	.3304
68	0.4071	0.5878	2.6291	68	0.4071	0.6594	1.3874

Form Statistics

Form Statistics	Previous Year Form A	Y10 Form A	Previous Year Form F	Y10 Form F
Mean	.026	-.175	.026	-.043
SD	.692	.784	.692	.771

*Note: mean and standard deviation of Year 10 is calculated with freely calibrated estimates.

Correlation and Standard Deviation Ratio

Correlation Coefficient	**.956**	**.945**
SD Ratio	**113%**	**111%**

Values Used for Robust Z Statistics

Mean Diff	-.201	-.069
Median Diff	-.233	-.072
IQR Diff	.213	.316

Based on correlation coefficients, SD ratio, robust z, and item difficulty plot, item numbers 9 and 25 appearing on both forms were dropped from the linking pool.

The following correlation coefficients and SD ratios were calculated after dropping those items:

Correlation with Base	**.965**	**.952**
SD Ratio	**109%**	**108%**

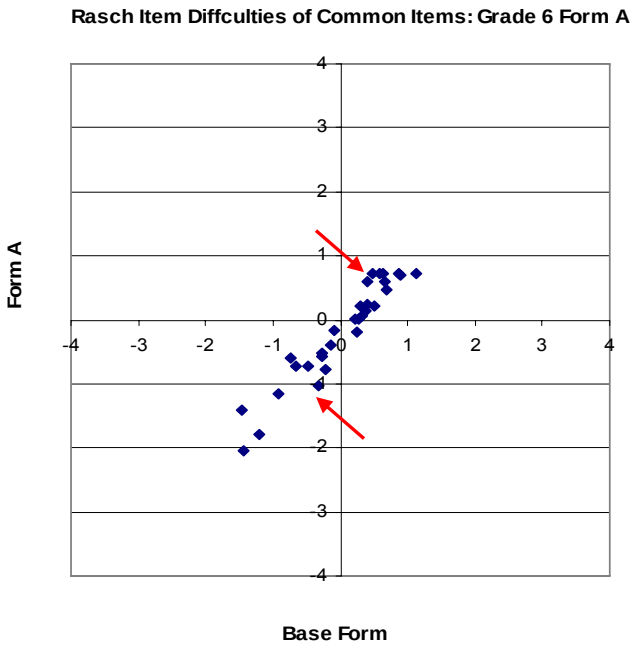

Figure 1.22 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 6 Form A

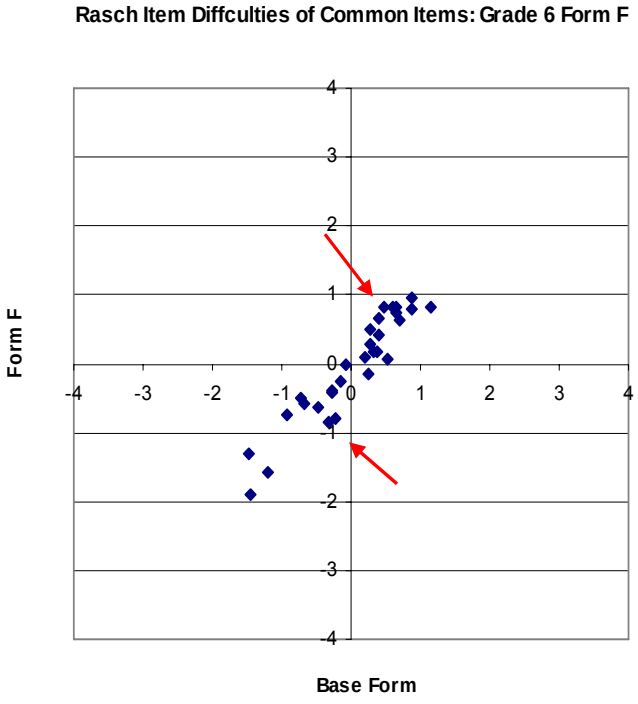

Figure 1.23 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 6 Form F

Table 1.78 Core Linking Item Difficulties of Previous Year vs. Year 2010: Grade 7

Item Seq No.	Previous Form A	Y2010 Form A	Robust Z	Item Seq No.	Previous Form F	Y2010 Form F	Robust Z
1	**1.0539**	**1.1766**	**3.0259**	**1**	**1.0539**	**1.2863**	**2.7079**
2	0.4455	0.0834	-.4445	2	0.4455	0.0905	-.3726
3	-2.682	-3.3599	-2.7052	3	-2.682	-3.5242	-2.9277
7	-0.1398	-0.5762	-.9764	7	-0.1398	-0.6867	-1.3790
8	-0.4706	-0.7918	-.1518	8	-0.4706	-0.781	-.1387
10	-0.3998	-0.6213	.5619	10	-0.3998	-0.6227	.3202
12	-1.4762	-1.8462	-.5011	12	-1.4762	-1.8908	-.6852
18	-0.6359	-0.8743	.4410	18	-0.6359	-1.0779	-.8289
19	-1.1243	-1.7982	-2.6766	19	-1.1243	-1.934	-2.7572
20	-1.9183	-2.118	.7180	20	-1.9183	-2.1155	.4550
30	1.4497	1.1756	.1854	30	1.4497	1.3115	.7644
31	0.9745	0.8174	1.0230	31	0.9745	0.9319	1.2657
32	0.0227	-0.5428	-1.9006	32	0.0227	-0.4267	-.8677
43	-0.4094	-0.6368	.5197	43	-0.4094	-0.6497	.2289
49	0.2502	0.2981	2.4905	49	0.2502	0.3978	2.2632
50	0.7735	0.2688	-1.4654	50	0.7735	0.2209	-1.4089
51	-0.0583	-0.1583	1.4317	51	-0.0583	-0.1941	.7770
52	-1.4991	-1.8934	-.6750	52	-1.4991	-1.8071	-.1261
63	0.5663	0.2041	-.4453	63	0.5663	0.34	.3023
64	0.0092	-0.2927	-.0136	64	0.0092	-0.2511	.1240
65	1.0861	0.4948	-2.0853	65	1.0861	0.5971	-1.0754
66	0.0819	-0.2508	-.2341	66	0.0819	-0.2261	-.1261
69	0.5231	0.2358	.0909	69	0.5231	0.1887	-.2646
70	-0.2784	-0.7218	-1.0265	70	-0.2784	-0.586	-.1240
72	0.6673	0.4131	.3279	72	0.6673	0.495	.5855
79	-1.4603	-1.4237	2.4096	79	-1.4603	-1.4479	1.5542
80	-0.0202	-0.3183	.0136	80	-0.0202	-0.2388	.3427
81	-0.0385	0.026	2.6093	81	-0.0385	0.0241	1.8174

Form Statistics

Form Statistics	Previous Year Form A	Y10 Form A	Previous Year Form F	Y10 Form F
Mean	-.168	-.465	-.168	-.449
SD	.985	1.049	.985	1.104

*Note: mean and standard deviation of Year 10 is calculated with freely calibrated estimates.

Correlation and Standard Deviation Ratio

Correlation Coefficient	**.981**	**.978**
SD Ratio	**106%**	**112%**

Values Used for Robust Z Statistics

Mean Diff	-.297	-.281
Median Diff	-.300	-.284
IQR Diff	.189	.258

Based on correlation coefficients, SD ratio, robust z, and item difficulty plot, item number 1 appearing on both forms were dropped from the linking pool.

The following correlation coefficients and SD ratios were calculated after dropping item 1:

Correlation with Base	**.982**	**.979**
SD Ratio	**104%**	**110%**

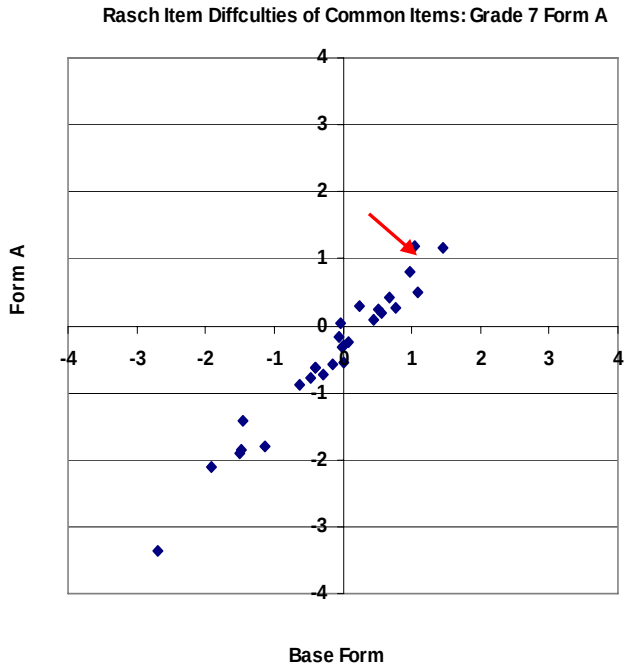

Figure 1.24 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 7 Form A

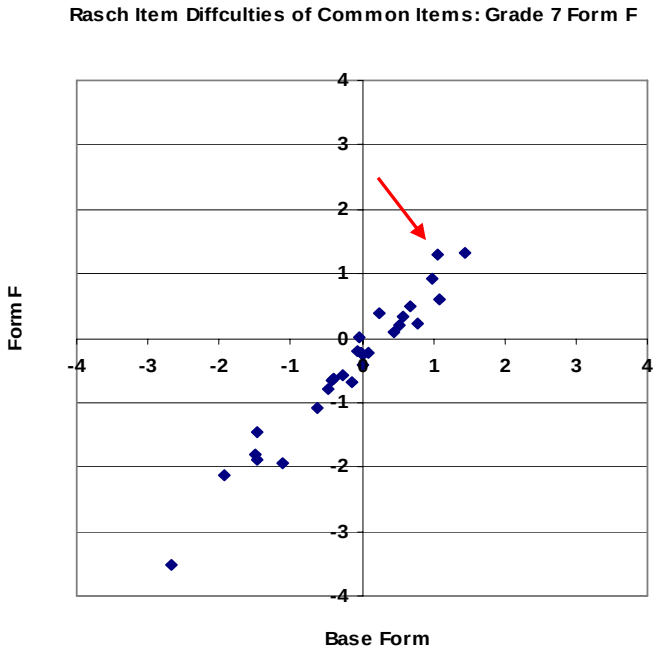

Figure 1.25 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 7 Form F

Table 1.79 Core Linking Item Difficulties of Previous Year vs. Year 2010: Grade 8

Item Seq No.	Previous Form A	Y2010 Form A	Robust Z	Item Seq No.	Previous Form F	Y2010 Form F	Robust Z
1	1.4965	1.4781	.7787	1	1.4965	1.5703	1.4430
2	-0.2177	-0.2389	.7613	2	-0.2177	-0.2083	.9065
5	-0.755	-0.9984	-.6129	5	-0.755	-0.9224	-.5665
7	-1.2003	-1.5009	-.9667	7	-1.2003	-1.4712	-1.4288
8	0.3621	0.2785	.3754	8	0.3621	0.2627	.0000
14	0.3158	0.0413	-.8052	14	0.3158	0.2347	.1525
22	-0.854	-1.103	-.6475	22	-0.854	-1.08	-1.0548
27	0.2581	0.2049	.5634	27	0.2581	0.4826	2.6985
32	-1.16	-1.3528	-.3000	32	-1.16	-1.2193	.3341
33	0.5139	0.3996	.1855	33	0.5139	0.3824	-.2674
38	-0.5927	-0.7885	-.3185	38	-0.5927	-0.6572	.2908
41	0.5661	0.1416	-1.7329	41	0.5661	0.0959	-3.0893
42	-1.4001	-1.8792	-2.0706	42	-1.4001	-1.6195	-.9998
46	-0.2581	-0.3998	.0161	46	-0.2581	-0.4006	-.3591
47	-2.1555	-2.3357	-.2220	47	-2.1555	-2.324	-.5757
48	-1.4852	-1.812	-1.1287	48	-1.4852	-2.0039	-3.4934
50	0.0551	-0.0892	.0000	50	0.0551	-0.0647	-.1700
51	-0.515	-0.5414	.7292	51	-0.515	-0.4395	1.4572
52	0.3257	0.4715	1.7942	52	0.3257	0.459	1.9387
53	-0.6275	-0.7689	.0179	53	-0.6275	-0.7257	.0100
62	1.2102	1.0535	-.0767	62	1.2102	1.0184	-.7698
65	-1.8764	-2.2085	-1.1615	65	-1.8764	-2.224	-2.0679
66	0.0346	-0.3301	-1.3631	66	0.0346	-0.0603	.0375
72	-0.3805	-0.3389	1.1497	72	-0.3805	-0.4258	.4507
73	**1.0119**	**1.3888**	**3.2235**	**73**	**1.0119**	**1.3826**	**3.9166**
78	-0.0934	-0.2145	.1435	78	-0.0934	-0.2272	-.2866
79	-0.0487	-0.082	.6865	79	-0.0487	0.0721	1.8346

Form Statistics

Form Statistics	Previous Year Form A	Y10 Form A	Previous Year Form F	Y10 Form F
Mean	-.277	-.427	-.277	-.375
SD	.906	1.005	.906	1.012

*Note: mean and standard deviation of Year 10 is calculated with freely calibrated estimates.

Correlation and Standard Deviation Ratio

Correlation Coefficient	**.988**	**.986**
SD Ratio	**111%**	**112%**

Values Used for Robust Z Statistics

Mean Diff	-.150	-.098
Median Diff	-.144	-.099
IQR Diff	.219	.162

Based on correlation coefficients, SD ratio, robust z, and item difficulty plot, item number 73 appearing on both forms were dropped from the linking pool.

The following correlation coefficients and SD ratios were calculated after dropping item 73:

Correlation with Base	**.998**	**.986**
SD Ratio	**108%**	**108%**

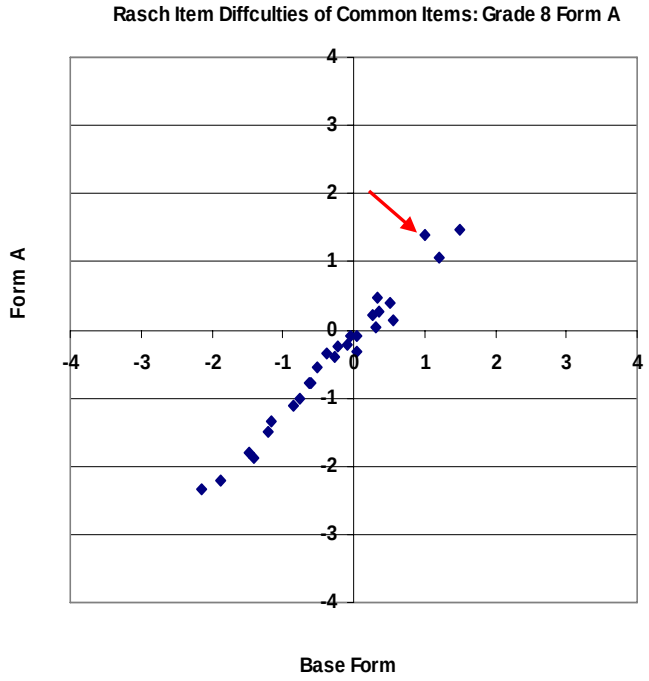

Figure 1.26 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 8 Form A

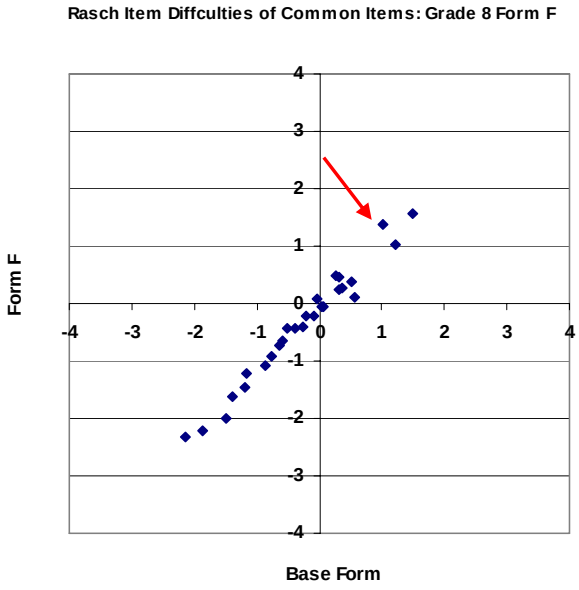

Figure 1.27 Item Difficulty Plot of Previous Year Form (Base Form) vs. Current Year (2010) Form: Grade 8 Form F

Reporting Scale Scores

In order to facilitate the use and interpretation of the results of the 2010 MSA-Math, the following formula was used to convert each student's ability or theta to the reporting scale score:

$$ReportingAbilityScaleScore = 32.8398 \cdot theta + 380.2954$$

$$ReportingSE = 32.8398 \cdot SE$$

where

theta = the Rasch (i.e., 1-PL *IRT*) ability estimate, and

SE = the conditional standard error of the ability estimate.

The following table contains information about the slopes and intercepts used to generate the 2010 scale scores. First of all, it should be noted that the slopes and intercepts were obtained during the 2006 recalibration. The same slopes and intercepts have been used since the 2006 assessment.

Table 1.80 The 2010 MSA-Mathematic Slope and Intercept: Grades 3 through 8

Grade	Slope	Intercept
3	32.6935	352.2959
4	32.8398	380.2954
5	30.7057	390.2866
6	29.6236	398.5595
7	28.1690	405.9549
8	28.3634	418.4843

1.10 Score Interpretation

To help provide appropriate interpretation of the 2010 MSA-Math test scores, two types of scores were created: 240-650 scale scores, and performance levels and descriptions.

240-650 Scale Scores

As explained in section 1.9, *Linking, Equating, and Scaling Procedures*, the 2010 scale scores were placed on a common scale (i.e., 2006 assessment) within the same grade and ranged from 240 to 260. As a result, these scale scores have the same meaning and are comparable across different years' assessments. However, it should be noted that they are not comparable across grade levels.

For scale scores, a higher score simply means a higher performance on math tests. Thus, performance levels and descriptions can give a specific interpretation other than a simple interpretation because they were developed to bring meaning to those scale scores.

Performance Level Descriptors

As previously explained, performance level descriptors provide specific information about students' performance levels and help interpret the 2010 MSA-Math scale scores. They describe what students at a particular level generally know and can be applicable to all students within each grade level.

Maryland standards are divided into three levels of achievement (*www.marylandpublicshools.org*):

- Advanced is a highly challenging and exemplary level of achievement indicating outstanding accomplishment in meeting the needs of students.
- Proficient is a realistic and rigorous level of achievement indicating proficiency in meeting the needs of students.
- Basic is a level of achievement indicating that more work is needed to attain proficiency in meeting the needs of students.

Table 2.1 shows a range of scale scores at each performance level; for example, grade 4 math scale scores from 374 to 432 indicate the level of *Proficient*. Students in this level passed the MSA-Math standard. This level is considered a realistic and rigorous level of achievement. Further information about the 2010 MSA-Math score interpretation can be obtained from MSDE.

1.11 Test Validity of the 2010 MSA-Math

As noted in the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999), "validity is the most important consideration in test evaluation."

Messick (1989) defined validity as follows:

> Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment. (p.5)

This definition implies that test validation is the process of accumulating evidence to support intended use of test scores. Consequently, test validation is a series of ongoing and independent processes that are essential investigations of the appropriate use or interpretation of test scores from a particular measurement procedure (Suen, 1990).

In addition, test validation embraces all of the experimental, statistical, and philosophical means by which hypotheses and scientific theories can be evaluated. This is the reason that validity is now recognized as a unitary concept (Messick, 1989).

To investigate the validity evidence of the 2010 MSA-Math, content-related evidence, item development procedures, differential item functioning (DIF) analysis on gender and ethnicity, and evidence from internal structure were collected.

Content-Related Evidence

Content validity is frequently defined in terms of the sampling adequacy of test items. That is, content validity is the extent to which the items in a test adequately represent the domain of items or the construct of interest (Suen, 1990). Consequently, content validity provides judgmental evidence in support of the domain relevance and representativeness of the content in the test (Messick, 1989).

The 2010 MSA-Math blueprints provide extensive evidence regarding the alignment between the content in the 2010 MSA-Math and the *VSC*. It should be noted that the 2010 MSA-Math operational test forms were built exclusively using a Maryland item bank program which contained both content and statistical information about both operational and field-tested items. Information on the item composition of the operational test forms can be obtained from section 1.4, *Test Form Design, Specifications, Item Type, and Item Roles*. In addition, the 2010 MSA-Math blueprints are presented in Appendix D.

Item Development

Test development for MSA-Math is ongoing and continuous. Content specialists, teachers from across Maryland, Pearson, and MSDE were greatly involved in developing and reviewing items. Committees such as content review, bias review, and vision review reviewed all of the items, which were finally stored in a Maryland item bank. Specifically, an internal review by MSDE and Pearson staff for content alignment and quality required a great deal of time and energy. More specific information on item (test) development and review can be obtained in section 1.3, *Development and Review of the 2010 MSA-Math Items and Test*.

Field test items were embedded and administered in one of ten test forms. Once these items were scored, MSDE and Pearson conducted additional item analysis and content review. Any field test items that exhibited statistical results that suggested potential problems were carefully reviewed by both MSDE and Pearson content specialists. A determination was then made as to whether an item should be eliminated, revised, or field-tested again. Information on statistical analyses for field test items can be obtained in section 1.13, *Field Test Analyses and Item Bank Construction*.

Differential Item Functioning **(DIF)**

1) Bias Review of Items

A separate Bias Review Committee examined each math item, with looking for indications of bias that could impact the performance of an identifiable group of students. They discussed or rejected items biased on gender, ethnic, religious, or geographical bias.

2) *DIF* Statistics

For DIF analyses, subgroups were first identified according to either reference or focal groups. For the 2010 MSA-Math, males and whites were assigned to the reference group and females and African-Americans were assigned to the focal group.

While the Mantel-Haenszel procedure was used for SR and SPR items, the standardized mean difference (SMD) and the standard deviation (SD), along with the Mantel statistic, were calculated for BCR and ECR items. All of the items were classified based on Educational Testing Service (ETS) guidelines. All *DIF* results were kept in the 2010 Maryland item bank. More information on *DIF* analyses can be obtained in section 3.7, *Differential Item Functioning*.

Evidence from Internal Structure

The 2010 MSA-Math has five reporting math standards: *Algebra, Geometry and Measurement, Statistics and Probability, Numbers and Computations,* and *Process*. Tables 4.3 through 4.8 show the correlations among the math standards.

1.12 Unidimensionality Analyses of the 2010 MSA-Math

Measurement implies order and magnitude along a single dimension (Andrich, 1989). Consequently, in the case of scholastic achievement, one-dimensional scale is required to reflect this idea of measurement (Andrich, 1988, 1989). However, unidimensionality cannot be strictly met in a real testing situation because students' cognitive, personality, and test-taking factors usually have a unique influence on their test performance to some level (Andrich, 1988; Hambleton, Swaminathan, & Rogers, 1991). Consequently, what is required for unidimensionality to be met is an investigation of the presence of a dominant factor that influences test performance. This dominant factor is considered as the ability measured by the test (Andrich, 1988; Hambleton et al., 1991; Ryan, 1983).

To check the unidimensionality of the 2010 MSA-Math, we examined the relative sizes of the eigenvalues associated with a principal component analysis of the item set. First, polychoric correlation coefficients were computed with *LISREL 8.5* (Jöreskog & Sörbom, 1993) because they were polytomously scored on math items. Principal component analysis was then applied to produce eigenvalues. The first and the second principal component eigenvalues were compared *without rotation*. Table 1.81 summarizes the results of the first and second principal component eigenvalues of the 2010 MSA-Math.

A general rule of thumb in exploratory factor analysis suggests that a set of items may represent as many factors as there are eigenvalues greater than 1 in this analysis because there is one unit of information per item and the eigenvalues sum to the total number of items. However, a set of items may have multiple eigenvalues greater than 1 and still be sufficiently unidimensional for analysis with IRT (Loehlin, 1987; Orlando, 2004). As seen from the following table, the first component extracted substantially larger eigenvalues across all grades: the size of the eigenvalue of the first component was over ten times greater than the second eigenvalue for each form at each grade. As a result, we could conclude that the assumption of unidimensionality for the 2010 MSA-Math was met.

Table 1.81 The 2010 MSA-Math Eigenvalues between the First and Second Components

Grade	Form	Number of Items	First Eigenvalue	Second Eigenvalue
3	A	65	22.38	1.84
	F	65	21.95	1.99
4	A	64	22.97	1.88
	F	64	21.80	1.68
5	A	65	21.32	2.04
	F	65	20.74	1.97
6	A	62	21.54	1.99
	F	62	20.24	1.83
7	A	62	24.63	2.17
	F	62	24.96	2.09
8	A	62	24.34	2.05
	F	62	24.82	1.98

Note. Form A designates the operational portion of Forms A, B, C, D, and E, which is identical. Form F designates the operational portion of Forms F, G, H, J, and K, which is identical.

Note. Analysis was conducted with a statewide population.

1.13 Field Test Analyses and Item Bank Construction

All field test items embedded in operational forms were subjected to rigorous statistical analyses for their properties in order to provide information about which items may be included as operational items in the future. All statistical results concerning field test items were preserved in the 2010 Maryland item bank. The following field test analyses were conducted:

- Classical item analyses for *SR*, *SPR*, *BCR*, and *ECR* items
- *Differential item functioning* (*DIF*) analyses
- *IRT* analyses

Classical Item Analyses for *SR, SPR, BCR, and ECR* items

Classical item analyses for *SR*, *SPR*, *BCR*, and *ECR* items were conducted within each field test form.

SR items were flagged for further scrutiny if:

- An item distractor was not selected by any students (i.e., nonfunctional distractor)
- An item was selected by a high proportion of high-ability students while being selected by a low proportion of low-ability students (i.e., ambiguous distractor)
- An item *p*-value was less than .20 or greater than .90.
- An item point-biserial was less than .10 (i.e., poorly discriminating). If an item point-biserial was close to zero or negative, the item was checked for a miskeyed answer.

SPR items were flagged for further scrutiny if:

- An item *p*-value was less than .20 or greater than .90.
- An item point-biserial was less than .10 (i.e., poorly discriminating). If an item point-biserial was close to zero or negative, the item was checked for a miskeyed answer.

BCR and *ECR* items were flagged for further scrutiny if:

- An item did not elicit the full range of rubric scores.
- The ratio of mean item score to maximum score (i.e.., adjusted p-value) was less than .20 or greater than .90.
- An item-total correlation was less than .10.

All items required a careful decision. For example, an item that was flagged as being difficult (*p*-value less than .20) and poorly discriminating (point-biserial less than .10) was considered for being dropped as a possible operational item. However, if the item represented important content that had not been extensively taught, a justification could have been made for including it in an operational test form.

Differential Item Functioning Analyses

Analyses of *Differential item functioning* (*DIF*) are intended to compare the performance of different subgroups of the population on specific items, when the group have been statistically matched on their tested proficiency.

In present analyses, the gender reference group was males, and the ethnic reference group was Caucasians. The gender focal group was females and the ethic focal group was African-Americans. For each operational form, the student's total score was used as the matching variable.

Any *SR*, *SPR*, *BCR*, and *ECR* items that were flagged as showing *DIF* were subjected to further examination. For each of these items, for example, math experts judged whether the differential difficulty of the item was unfairly related to group membership using the following criteria:

- If the differential difficulty of the item is related to group membership, and the difference is deemed unfair, then the item should not be used at all.
- If the differential difficulty of the item is related to group membership, but the difference is not deemed unfair, then the item should only be used if there is no other item matching the test blueprint.

It should be noted that DIF analysis results for all the field test items were archived in the 2010 Maryland item bank. Detailed information about the *DIF* procedures can be found in section 3.7, *Differential Item Functioning*.

Item Response Theory (IRT) Analyses

To put the 2010 field test items on a common scale (i.e., the 2006 scale), each field test item was freely calibrated after fixing Rasch item and step difficulty parameters of the 2010 operational items that had been already placed on the base scale during the 2010 operational calibration and equating. For example, each unique field test item appearing on one of five math test forms (i.e., A, B, C, D, and E) was independently calibrated after fixing the same operational items appearing across the field test forms with the same Rasch item and step difficulties because these unique field test forms all correspond to the same operational form (i.e., operational form A).

It should be noted that all the Rasch item difficulties, step difficulties, and fit statistics (i.e., Rasch Infit and Outfit indices) of the field test items were archived in the 2010 Maryland item bank. These field test items are eligible to be used as operational items in subsequent years.

Item Bank Construction

The number of test forms constructed each year and the need to replace items that are released to the public necessitates the availability of a large pool of items. The 2010 MSA-Math item bank continues to be maintained by Pearson in the form of computer files and paper copies. This enables the test items to be readily available to both Pearson and MSDE staff for reference, test construction, test book design, and printing.

1.14 Quality Control Procedures

A standard quality procedure at Pearson was to create a test deck for MSA programs. The test deck began when Quality Assurance entered mock data into the enrollment system, which was transferred to the materials requisition system; the order was packaged by our Distribution Center, and shipped to the Quality Assurance Department. We then reviewed the packing list against the data entered, the materials algorithms applied, the materials packaged against the packing list, and the actual packaging of the documents. These documents were then used to create a test deck of mock data, along with advance copies of documents that were received from the printer. Advance printer copies were inclusive of documents throughout the print run to assure we were randomly testing printed documents. The Maryland test deck was a comprehensive set of all documents that:

- Verified all scan positions for item responses and demographics to verify scanning setup and scan densities
- Verified all constructed response score points, zoning of image, reader scoring, reader resolution, and reader check scores
- Verified the handling of blank documents through the system
- Test all demographic and item edits
- Verified pre-id bar code read, match and no-match
- Verified attemptedness rules applied by subtest
- Verified duplicate student handling (same test duplicate, different test duplicate)
- Verified duplicate student with different demographics ruedles applied
- Verified the document counts to the enrollment, pre-id and actual document receipt
- Verified pre-id matching and application to student record
- Verified various raw score points and access to dummy and live scoring tables
- Verified cut scores applied
- Verified valid score on one subtest and invalid score on other subtest
- Verified scoring applied to Braille and Large Print
- Verified valid multiple choice and invalid constructed response
- Verified valid constructed response and invalid multiple choice
- Verified all special scoring rules
- Verified all summary programs for rounding
- Verified summary inclusion and exclusion (Braille, standard and non-standard student summarization)
- Verified each scoring level for group reporting
- Verified all reporting programs for accuracy in all text and data presented
- Verified class, school, district, and state summary data on home reports
- Verified all data file programs to assure valid information in every field

- Verified data descriptions for accuracy against data file
- Created compare programs to allow for update of files

The Maryland test deck was the first order processed through the Maryland system to verify all aspects of the materials packaging, scanning, editing, scoring, summary, and reporting. Pre-determined conditions were included in the test deck to assure the programs were processing all data to meet the requirements of the program with zero defects. Processing of live orders could not proceed until each phase of the test deck had been approved by our Quality Assurance Department. An Issues Log with sign-off approvals was utilized to assure we were addressing any issues that arose in the review of the test deck data across all functional groups at Pearson.

Prior to release of any order for reporting we received a preliminary file from Scoring Operations to run a key check TRIAN to assure that all scoring keys had been determined and applied accurately. Any item that was not performing as expected was flagged and reviewed by our content specialist and psychometrician. Upon completion of the key check, we proceeded to run the pilot level reports.

We ran the pilot district utilizing live data. The pilot district included multiple buildings, all grades, and any unique accommodations. A formal pilot review process was conducted with Pearson staff experts prior to release of the information to MSDE.

Upon completion of the processing of all district-level data, Pearson Scoring Operations provided the Quality Assurance Department with one or more state-level data files, along with state data for review and approval. Pearson Quality Assurance programmers duplicated all data independently to ensure accurate interpretation of the expected results. A series of SAS programs were run on these files to ensure 100% accuracy. These included but were not limited to:

- Statewide Duplicate Student
- Statewide FD of Demographic Variables
- District/Building/N-Count
- Statewide RS/SS/Cut Score tables
- Proc Means to verify summary statistics
- Item Response listing to verify all constructed responses were scored and within the valid range
- Normative data check for all raw scores
- Reader Resolution report to verify all readings and resolution combinations

Upon complete review and approval by Quality Assurance, we posted the statewide student files to a secure FTP site for review by MSDE.