

## **1. OVERVIEW OF THE 2008 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. 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 equipercentile method was conducted in that year. Detailed information on scale score transformation can be found in Appendix C, *Year 2006 MSA-Math Recalibration Results from 3-PL IRT to the Rasch Model Using Equipercentile Method* in the *2007 MSA-Math Technical Report*.

In 2007, MSDE decided to drop 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 purpose of year-to-year linking and equating, operational selected-response (SR) items (i.e., multiple choice items) appeared both in 2008 and previous years were exclusively used. It should be noted that Rasch item difficulty parameters of the Maryland-specific items generated by recalibration of 2006 data were kept as fixed parameter during the 2008 linking and equating process. All scale scores of the 2008 assessment were linked back to the 2006 assessment so that all of the scale scores were on the same scale within each content and grade.

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).

### **1.1 Purposes/Uses of the 2008 MSA-Math**

By measuring students’ achievement against the new academic standards, the 2008 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 2008 MSA-Math Blueprint*.

### **1.3 Development and Review of the 2008 MSA-Math Items and Test**

The development of the 2008 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 2008 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.

Table 1.1 identifies responsibilities of each group in developing the 2008 MSA-Math test.

**Table 1.1 The 2008 MSA-Math Responsibility for Test Development**

Development of the 2008 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 Type, 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 2008 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, the 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 2008 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 2008 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 2008 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 conducted 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 1.10, *Form-to-Form Linking Procedures*. More importantly, however, year-to-year linking was conducted with only the core linking items and year-to-year linking procedures on these core linking items can be found in section of chapter 1.10, *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 2008 linking pool, on the other hand, item parameters of these items were recalibrated with the 2008 live, operational data and then reserved in the 2008 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 1.7, *Validation Check with the 2008 Core Items*.

While a few core linking items appeared only on operational form (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 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 2008 scale on the 2006 scale. Because these core linking items carried their operational item parameters on the 2006 scale, they were included in the 2008 year-to-year linking pool. Classical analysis on these items can be found in section of chapter 1.7: *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 1.10, *Linking, Equating, and Scaling Procedures of the 2008 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 items were used for form-to-form calibration and linking. Because these items were not included in the 2008 year-to-year linking pool, new Rasch item and step difficulty parameters were estimated with the 2008 live, operational data set. 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 in terms of classical and Rasch item difficulty can be found in section of chapter 1.7, *Validation Check with the 2008 Core Items*.

**Table 1.3 Item Type of Content Standard for the 2008 MSA-Math: Grades 3 and 4**

Grade	Strand Title	Item Type	No. of Items in Each Form
3	Total CRT	<i>SR, BCR</i>	65
	Algebra	<i>SR, BCR</i>	13
	Geometry	<i>SR, BCR</i>	8
	Measurement	<i>SR, BCR</i>	7
	Statistics	<i>SR, BCR</i>	12
	Probability	<i>SR</i>	2
	Number Computation	<i>SR, BCR</i>	16
	Process	<i>BCR</i>	7
4	Total CRT	<i>SR, BCR</i>	64
	Algebra	<i>SR, BCR</i>	14
	Geometry	<i>SR, BCR</i>	7
	Measurement	<i>SR, BCR</i>	7
	Statistics	<i>SR, BCR</i>	8
	Probability	<i>SR, BCR</i>	7
	Number Computation	<i>SR, BCR</i>	14
	Process	<i>BCR</i>	7

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



**Table 1.4 Item Type of Content Standard for the 2008 MSA-Math: Grades 5 and 6**

Grade	Strand Title	Item Type	No. of Items in Each Form
5	Total CRT	<i>SR, BCR, ECR</i>	65
	Algebra	<i>SR, BCR, ECR</i>	15
	Geometry	<i>SR, BCR</i>	6
	Measurement	<i>SR, BCR</i>	8
	Statistics	<i>SR, BCR</i>	9
	Probability	<i>SR, BCR</i>	4
	Number Computation	<i>SR, BCR</i>	15
	Process	<i>BCR, ECR</i>	8
6	Total CRT	<i>SR, BCR, ECR</i>	62
	Algebra	<i>SR, BCR, ECR</i>	14
	Geometry	<i>SR, BCR</i>	8
	Measurement	<i>SR, BCR</i>	6
	Statistics	<i>SR, BCR</i>	9
	Probability	<i>SR, BCR</i>	4
	Number Computation	<i>SR, BCR</i>	14
	Process	<i>BCR, ECR</i>	7

**Table 1.5 Item Type of Content Standard for the 2008 MSA-Math: Grades 7 and 8**

Grade	Strand Title	Item Type	No. of Items in Each Form
7	Total CRT	<i>SR, SPR, BCR, ECR</i>	62
	Algebra	<i>SR, SPR, BCR, ECR</i>	14
	Geometry	<i>SR, SPR, ECR</i>	7
	Measurement	<i>SR, SPR, BCR</i>	6
	Statistics	<i>SR, SPR, BCR, ECR</i>	9
	Probability	<i>SR, SPR, BCR</i>	5
	Number Computation	<i>SR, SPR</i>	14
	Process	<i>BCR, ECR</i>	7
8	Total CRT	<i>SR, SPR, BCR, ECR</i>	62
	Algebra	<i>SR, SPR, BCR, ECR</i>	14
	Geometry	<i>SR, SPR, ECR</i>	7
	Measurement	<i>SR, SPR, BCR</i>	6
	Statistics	<i>SR, SPR, BCR, ECR</i>	9
	Probability	<i>SR, SPR, BCR</i>	5
	Number Computation	<i>SR, SPR</i>	14
	Process	<i>BCR, ECR</i>	7

**Table 1.6 Item Distribution of Each Content Standard for the 2008 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 2008 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 2008 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 2008 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 2008 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 2008 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 2008 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	<b>65</b>
F	15	6	8	9	4	15	8	<b>65</b>

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 2008 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	<b>74</b>
F	15	14	13	15	17	<b>74</b>

**Table 1.14 Item Type and Score Point Distribution for the 2008 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	<b>65</b>	49	
F	49	7	7	1	1	<b>65</b>	49	7	14	1	3	<b>74</b>

**Table 1.15 Item Distribution of Each Content Standard for the 2008 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	<b>62</b>
F	14	8	6	9	4	14	7	<b>62</b>

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 2008 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	<b>70</b>
F	14	14	13	14	15	<b>70</b>

**Table 1.17 Item Type and Score Point Distribution for the 2008 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	<b>62</b>	48	6	12	1	3	<b>70</b>
F	48	6	6	1	1	<b>62</b>	48	6	12	1	3	<b>70</b>

**Table 1.18 Item Distribution of Each Content Standard for the 2008 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 2008 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 2008 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 2008 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 2008 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 2008 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 Test Administration of the 2008 MSA-Math

The 2008 MSA-Math test was administered to all students in grades 3 through 8. Pearson coordinated test administration procedures with MSDE prior to implementation. This chapter was prepared to provide general information about the 2008 test administration. Detailed information about the 2008 test administration can be obtained from the 2008 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:

- Examiner's Manual- Math

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
- Calculator (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:

- Sign for the door reading "Testing: Do not Disturb"
- Digital clock or a watch, or clock with a second hand
- Copy of the Scoring Service Identification Document (SSID) Header Sheet

Two test-related Examiners Manuals (EMs) were developed for the 2008 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 (April 1-10, 2008, with make-up testing held April 11-16, 2008). However, each Local Education Agency (LEA) set a specific schedule for administration of the MSA within that window for their district. For a given grade and content area, all testing 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 Mathematics must be administered on the same days throughout the LEA. In addition, each content area at each grade was tested on two days during the window.

The MSA-Math testing schedule allowed approximately 2 1/2 hours on each of the two days (including preparation time and breaks).

For the 2008 MSA-Math, the primary testing days were as follows:

- Test materials delivered to schools                      On or Before March 10, 2008  
(Examiner's Manuals, Test/Answer Books,  
and Test Coordinator's Kit)
- Mathematics Primary Testing Window                  April 1 – April 10, 2008
- Make-up Testing Window                                      April 11 – April 16, 2008

Students and parents should be reminded of the importance of students attending school during the administration of the MSA and the importance of student participation in MSA testing. Maryland was held to the 95% participation requirement under NCLB by the US Department of Education, and schools should do all they can to test all students on MSA or Alt-MSA (as applicable).

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 2008 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

All students in grades 3 through 8 had to participate in the 2008 MSA-Math. 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. The criteria that students should need to be tested in the Alt-MSA program instead of the MSA-Math can be viewed in section 2, Appendix C of the TACM.

On May 9, 2007, the U.S. Department of Education issued guidance for the development of Alternative Assessment based on Modified Academic Achievement Standards (also known as AA-MAAS or “Modified Assessments”). Maryland was in the process of developing the Modified Maryland School Assessment (Mod-MSA), but the assessment was not completed in time for the 2008 administration window. Students, however, might have been identified through the Individualized Education Program (IEP) process in the current school year as takers of the Mod-MSA. For 2008, these students were assessed using the regular MSA-Math.

## 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 2008 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 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 2008 TACM in section 4.

### **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 2008 MSA-Math. The accommodation was provided by a live reader or through technology. Section 1 of the 2008 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 (Harcourt, 2008):

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 or MSDE disciplinary action. (p. 13)

The Test/Answer Books for the 2008 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 (Harcourt, 2008):

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. 13)

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

## 1.6 Scoring Procedures of the 2008 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.

Once received by Pearson, Test/Answer Books were scanned into an electronic imaging system so that the information necessary to score responses was captured and converted into an electronic format. Students' identification and demographic information, school information, and answers to *SR* items were converted to alphanumeric format; hand-written responses were captured in digital image format.

### Machine-Scored Items

After students' responses to *SR* and *SPR* items were converted to text format, the scoring key was applied to the captured item responses. Correct answers were assigned a score of one point. Incorrect answers, blank responses (omits), and responses with multiple marks were also assigned a score of zero.

### Hand-Scored Items

Test/Answer Books were scanned into the electronic imaging system, allowing scorers to score these responses online at all scoring sites while maintaining the live documents at the contractor's facility. The imaging system randomly distributed responses, ensuring no one scorer scored a disproportionate number of responses from any one school. This online scoring system maintained a database of actual student responses and the scores associated with those responses. An off-site backup of all images and scores was maintained as well to guard against potential loss of data and images due to system failure. The system also provided continuous, up-to-date monitoring of all scoring activities. Detailed information on MSA scoring specification can be obtained in the document *Performance Assessment Scoring Center: Spring 2008 Scoring Specification for MSA-Reading and Math*, which is available from either MSDE or Pearson.

### Scoring Staff

The MSDE had one Room Director (RD) dedicated to each grade level, domain (Math), and site. The RD worked closely with the PASC Training Supervisor and the PASC Math Specialists. The PASC Training Supervisor, Math Specialist, and RDs participated in the anchor-pulling sessions in Maryland. (Detailed information about anchor-pulling procedures can be found in the following portion of this section: *Development Procedures for Anchor Pulling*.) The Room Director/Training Team Leader 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 RD responsible for training the Team Leaders (TLs) and Readers for the respective Maryland prompts. For MSDE scoring projects, PASC had qualified alternate RDs available at the beginning of the project to ensure a timely start of training in the event that the primary RD was unavailable to start as scheduled. The alternate RD acted as a TL unless the RD couldn't fulfill his/her duties.

#### 1) Reader/Scorer

A graduate of a four-year accredited college or university who had successfully passed the PASC new reader exam and new reader training. The Readers were eligible to score custom programs for which they had been trained and successfully qualified.

**2) Team Leader (TL)**

An experienced reader who directly monitored the scoring of a team of Readers and retrained as needed. The reader had successfully completed the PASC TL training program.

**3) Room Director (RD)**

A knowledgeable team leader who had been selected to work with team leaders and the training supervisor to oversee the scoring of several teams. An RD's main duty was to rule on validity of questionable papers and to maintain consistency in scoring decisions. RDs also served as trainers.

**4) Reader's Aide (RA)**

PASC storeroom personnel whose main responsibilities during scoring were to do copying and printing for the PASC materials center. During anchor pulling, RA responsibility might include duplicating student papers. They might also be assigned a variety of clerical duties.

**5) Developers**

An experienced PASC reader that was responsible for selecting a wide variety of student responses for such activities as benchmarking, anchor pulling, range finding, and training materials. Selected papers were then submitted to MSDE for comment and approval. Developers remained on the project as anchor-pulling participants and trainers whenever possible.

**6) Trainers**

Experienced personnel who were TLs or RDs and selected by the Training Supervisor to train and qualify readers for Maryland. Additionally these experienced personnel might also train new Readers and do domain-specific training.

**Reader Recruitment and Qualifications**

All Readers for MSDE had to provide Pearson's staffing vendor their résumé and documentation of a four-year college degree. As part of the initial screening process for recruiting Readers into Pearson's general pool, applicants had to respond to an open-ended prompt. This writing sample ensured that all applicants were able to perform the kinds of tasks they would assess. The writing sample was intended to screen out those who couldn't write standard, idiomatically correct English or who couldn't organize their thoughts clearly. The writing prompt was scored by a qualified PASC staff member. If successful on the preliminary screening, applicants then participated in a one-day general introductory training workshop presented by a PASC staff member. These workshops allowed Pearson to eliminate potential Readers who might seem qualified according to their educational and professional experience but who couldn't learn to score to a scale consistently or who were otherwise unsuitable for assignment to large-scale scoring projects. The PASC staff member who presented the workshop evaluated each potential Reader and submitted these evaluations to the Training Supervisor/Site Supervisor with his/her recommendations. Those who successfully completed the workshop were added to Pearson's general pool of Readers who were potential scorers of Math assessments. This addition to the general pool did not necessarily qualify these Readers for scoring the MSDE program.

### **Team Leader Selection and Qualification**

The training for new TLs consisted of a two-day course focusing on the duties and responsibilities necessary to successfully manage a team of Readers. The workshop was led by two PASC Training Supervisors. The instruction included a review of PASC policies and procedures, sessions on use of the Reader monitoring reports to track a Reader's speed and accuracy, practice annotating anchors and simulated training of the annotated papers, role playing activities which explored various situations that could occur with Readers during the scoring of a project, and Reader counseling and retraining guidelines. Hands-on training on the various TL computer applications was also provided in the workshop. Upon completion of the workshop, the two PASC Training Supervisors reviewed each participant's performance, making sure that each had a complete understanding of the TL role and its responsibilities. Any participant they found who did not perform to their satisfaction was not added to the qualified TL list.

### **Team Leader Project Training**

Project-specific TL training for MSDE was conducted in the days immediately preceding scoring and Reader training. This training began with the RD reading the rubrics aloud and answering any questions the TL or assistant RD might have regarding the rubric. The RD then read each anchor paper aloud to the TLs. Each response in the anchor set was thoroughly explained, including the notes and comments of the anchor-pulling committee. Training set A was reviewed next. The TLs scored the training set individually, recorded the scores on the answer sheet, and then waited for all TLs to complete the scoring. When everyone had completed scoring the training set, the RD discussed the answers one by one, focusing on why it was that score and not another. The RD reviewed with the group the reason for assigning each score point and discussed each paper in its entirety. The TLs were then ready to score Training set B. Training set B was scored and reviewed exactly as Training set A.

Having thoroughly discussed both training sets with the group, the RD explained that in order for a participant to qualify as a TL, it was required that the TL should score at least an 80% perfect match on both of the qualifying sets (Qualification Rules, Attachment M). The TLs scored the first qualifying set individually and recorded their scores on the appropriate answer sheet. As each TL finished scoring, he/she brought the answer sheet to the RD for grading. Each answer was reviewed and any questions the TL had were addressed before the TL attempted the next qualifying set. The TL followed the same procedure with Qualifying set 2. Upon completing the second qualifying set, the TL submitted the answer sheet to the RD for grading. TLs had to pass both sets 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.

After the qualification process, the RD continued the training process with the decision set. This set was read aloud and each paper thoroughly explained and discussed. By following these procedures, the RD ensured that the anchor-pulling committees' notes and comments were completely understood.

### **Team Leader Duties**

TLs were responsible for monitoring the training and qualifying of the Readers assigned to their team. The TLs assisted the RD, if requested, during the training of the Readers. The TL was responsible for grading the Readers' qualifying sets and discussing the results with the Readers so everyone received the same direction. The TL certified to the RD and Training Supervisor that the Reader was qualified and recorded the scores under Qualification scores on the Reader evaluation



form. The TL was also responsible for monitoring each Reader's assignment of scores to the responses. Additionally, the TL reviewed the daily Reader statistical reports with each individual on the team. The TL consulted the RD regarding variations by the team members from the acceptable standards (95% for Math Step A, and 85% for Math Step B). The TL had the initial responsibility to see that the Reader maintained the set standards through individual retraining. The RD monitored the TL by reviewing team statistics and working one-on-one with the TL.

### **Room Director Selection and Qualification**

The candidates for RD had been recommended by the PASC Managers or Training Supervisors. The recommendations were based upon the evaluations the candidates received as Readers and TLs and were part of their personnel file. The Training Supervisors met as a group to discuss who might be considered for the position of RD. The Training Supervisor group reviewed the evaluations and the duties that the potential RDs had performed. The candidates generally had been TLs on large-scale projects for multiple teams, and/or they had served as TLs on small-scale projects where TLs trained their individual teams. They had been evaluated on their ability to train Readers as well as their ability to monitor the scoring accuracy and consistency of Readers. These evaluations were submitted in writing at the end of each scoring project by the Readers and RDs that had observed the work of the RD candidates.

### **Room Director Project Training**

The RDs familiarized themselves with the rubric. Any questions regarding the rubric were addressed by the PASC Math Specialists or MSDE. The next step was for the RD/TTL to prepare the anchors by annotating each response to all score points in the Anchor Set utilizing the notes from the anchor-pulling session. The MSDE approved the anchor-pulling notes and the Training Supervisor confirmed that the RD had accurately added the anchor-pulling notes to the training materials. The RD continued the process by annotating the training sets and decision sets with all notes and comments from the anchor-pulling session. Additionally, the RDs became familiar with the wording of all of the other prompts for the administration to which they were assigned.

### **Room Director Duties**

The RD's job was to conduct the training of the TLs and Readers, oversee the actual scoring of the papers, monitor the work of the TL, and act as the decision maker for situations or questions that may arise during the scoring process. For example, all invalid (foreign language, off-topic, off-mode, etc.) responses were reviewed by the RD, who had to confirm any such decision and ensure consistency of decisions. (Blanks were confirmed at the TL level and did not require RD confirmation.) Additionally the RD and TL (after approval of Training Supervisor) conducted all resolution readings. Responses for which scores were non-matching or non-adjacent were automatically routed to the RD for an independent resolution scoring. The resolution score became the reported score.

The RD was familiar with all prompts and trained the TLs and Readers to recognize these alternate prompts. Thus, should the student have written his/her answer in the wrong place, the answer was recognized by the RD, who could electronically move the response to the appropriate space for scoring by a Reader qualified on the appropriate prompt. The RD also reviewed any potential questionable content responses and forwarded those to the Training Supervisor to consult with the MSDE before processing.

The RD was also responsible for daily statistical review and analysis of all monitoring reports to ensure the quality of the scoring within the room. Review of the data allowed the RD not only to

monitor the Reader but also to provide the TL with additional input. Available data included 1) individual Reader agreement rates between two independent scorings; 2) score point distributions by Reader and trend review; 3) prompt statistics for agreement rates and score point distributions; 4) Resolution data.

### **Project Scoring Parameters**

MSDE had a long-standing history of implementing assessments that were composed of multiple item types: selected response (SR), brief constructed response (BCR), extended constructed response (ECR), and gridded or student-produced response (SPR). The MSA-Math contained all such item types for operational scoring, and each of the 10 forms per grade also contained field test items of each of these types. Open-ended items were scored using a generic rubric as follows:

- Mathematics BCR items: Step A 0-1 scale, Step B 0-2 scale
- ECR items Step A 0-1 scale, Step B 0-3 scale

All MSA-Math response documents were image-scanned at Pearson's scoring center in San Antonio, Texas. The image scanner captured document identification (ID), demographic information, SR responses, and created a bi-tonal image of the entire document, allowing images of the BCR and ECR responses to be distributed to Readers for human scoring while images of the SR, SPR and all other data were made available to Scoring Editing for human review.

All constructed responses were scored by Pearson's Performance Assessment Scoring Center (PASC). The PASC mission was to provide accurate, reliable, on-time scores for all student responses entrusted to our care. PASC maintained large pools of qualified, trained, professional Readers who were well-experienced in scoring a wide range of writing assessments and open-ended assessments in reading, mathematics, science, social science, and other subjects, at each of our scoring sites.

### **Reader Project Training**

Reader training was lead by the RD/TL and was conducted utilizing our central scoring model. There was one RD responsible for each site, grade, and Domain (Math). After all student responses were scored for the first item, the RD 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 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 Readers' 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 Readers had all their questions answered and the discussion of the anchor set was finished, the Readers began to score the first training set. Each Reader independently read and scored the responses in the training set. The trainer scored and recorded each reader's responses on a training record form. The correct scores were then read to the group when everyone had completed the scoring. In addition, each training paper was discussed as to reasons for applying each given score. At this point, Readers interacted with the RD in discussing the characteristics of each response that earned the assigned score point. The same format was followed for each training set. During this process, the job of the Reader was to internalize the scoring scale and adjust his or her

individual scoring to conform to that scale. Once all training papers had been scored and fully discussed, Readers began the qualifying process.

For MSDE, there were three qualifying sets. MSDE informed PASC in writing for each specific administration how many qualifying sets were approved and were available to the Readers. Readers had to score at least an 80% on at least one of two qualifying sets for Math.

### **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 MSDE scoring. These reports were compiled by prompt, listed the entire prompt's Readers, and provided the results of their scoring for each day. Information on these reports included the number of responses read by the Readers during the period, the number and percent of invalid 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 Reader was a point lower than the second Reader; the number and percent of responses on which the first Reader was a point higher than the second Reader (Adjacent); and the number and percent of responses differing by more than one score point (Non-Adjacent/Non-Perfect). The Training Supervisor 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*.

### **Reader Retraining**

When a Reader's performance fell below acceptable parameters for a project, the Reader was retrained. Retraining was the process by which the RD or TL 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 Reader back on track with the guidelines provided by a specific program. Group retraining was conducted by the RD every Monday (or following any extended break) during the scoring project. In addition, daily retraining occurred as deemed necessary by the MSDE representative and Training Supervisor.

### **Read Behinds**

Pearson's system allowed TLs and/or RDs to conduct read behinds as an additional monitoring method. When conducting read behinds, the TL or RD received images of student responses and the scores assigned by the Reader. Responses selected for read behinds might be randomly selected or might be targeted read behinds (e.g., responses receiving specific scores, etc.). These read behinds were very useful in tracking specific areas of confusion for a given Reader or group of Readers and assisted the TL and RD in knowing just how to direct retraining activities for individual Readers or teams. The initial read behind percentage was set at 50%. This percentage might be adjusted either higher or lower by the TL based upon the performance of the Reader.

### **Retraining Readers with < 80% Agreement rates**

It was the responsibility of the Team Leader (TL) to not only address questions and provide guidance to the Readers, but to also monitor and manage performance; this included Calibrations, Read Behinds, Agreement rates, and Resolution rates. At times, TLs could become easily side-tracked and spend more time acting as a resource for Readers than managing performance. PASC had identified this issue and planned to allocate additional TLs whose primary job responsibility

was to manage/monitor performance. This level of staffing allowed us to monitor each Reader daily and provide retraining when the level of acceptable performance had not been met.

### **Pre-“Live” training on Field Test prompts**

For 2008, PASC used scored student responses from the appropriate field test administration. This allowed the Readers to build familiarity with the program prior to live scoring.

### **Trainers Earlier and Longer**

In addition to increasing the number of TLs dedicated to the program, PASC also felt it more effective to expedite and extend the time the Trainers were onsite. PASC trained a qualified individual at each site to act as the remote Trainer once the primary left. This individual was responsible for retraining Readers as needed.

### **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 Reader’s scores provided they were adjacent.
- For example:

1 <sup>st</sup> Reader	2 <sup>nd</sup> Reader	Final Score
1	2	2
2	3	3

- A resolution reader was used if two non-adjacent initial scores were received.
- The resolution reader’s score was used in place of both the 1st and 2nd Readers’ scores.
- For example:

1 <sup>st</sup> Reader	2 <sup>nd</sup> Reader	Resolution Reader	Final Score
0	2	1	1
0	3	2	2
1	3	3	3
2	0	1	1
3	0	2	2

## Development Procedures for Anchor Pulling

A Developer is a PASC Reader who was selected by the PASC Training Supervisor to prepare sets of papers for client approval. These experienced Readers were judged by the Training Supervisor for their ability to recognize and assemble a wide variety of responses. A Material Development Evaluation was completed by the Math Specialists for review by the Training Supervisor. This evaluation was part of the Developer's personnel file. The Developer also participated with the clients as a facilitator during the anchor-pulling session in order to make notes and be prepared to assemble the finished sets to the client's specifications. In the case of the MSDE, the Developer was also the RD. For a given math prompt, the PASC Developers 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.
  - Identified and flagged problem papers—off-topic, off-task, verbatim copying, strange, potential teacher interference, etc.
  - Marked the flag with score range or the nature of the problem and paper ID.
- 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 anchor pulling
  - Decided which particular papers from the sorted piles should go into which set for anchor pulling. Each paper selected went into only one set.
  - Used the following guidelines in deciding for which set a paper was most appropriate.
    - A. **Anchor set:** At least three examples of each score point, depending upon the score scale (no invalids). These had to be clean papers but needed to illustrate different types of the same score point, if there were such clear differences. Once completed, this set was submitted to the Training Supervisor and to MSDE for review and approval.

B. **Decision set:** This had to be a set of whatever size necessary to illustrate the various kinds of problems that might arise with this prompt or item. If the number of such responses was small, these might be incorporated into the first training set instead of being grouped into a separate additional set.

C. **Training sets:** These were at least two sets of up to 20 papers each (again, this varied according to the score point scale). They had to contain a range of responses including clean papers, line papers, and problem papers. The responses had to be in random order of quality and unmarked.

D. **Qualifying sets:** There were three sets of these. Generally there were 10 responses per set, but there could have been fewer, depending upon the score scale. These had to consist heavily of clean papers but not exclusively so. One of the sets might include an example of an invalid response, but it had to be clearly so.

E. **Calibration sets (validity sets):** These were composed of five responses of mixed quality, arranged in random order. Pearson created as many different sets as there were expected to be scoring days on a single prompt or group of items—minus one or two for the training day and the initial scoring day.

Comprehensive notes concerning the specific problems presented in these papers (and the solutions as decided by the committee during the anchor-pulling session) were to be recorded by the Pearson representatives (Developers and Training Specialists) and were to be discussed with the Readers during training. Any subsequent notes or communication from MSDE were incorporated into the training material as well.

### **Anchor Pulling Procedures**

The objective of anchor-pulling 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 PASC Math Specialists, Managers, Training Supervisors, and the Developers, 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, Decision Sets, Training 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 anchor pulling, the procedure for assigning scores to the papers in each set was as follows:

- Papers were read aloud and discussed by the anchor-pulling 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.

- Each subsequent set was read and scored by each panel member, 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.

The responses in which score points were not in perfect agreement were discussed, starting with the lowest, but least controversial, score point. The papers that had the widest discrepancies of assigned scores around this lowest score point were discussed next before moving on to the papers whose assigned scores were in the next higher range. There might be frequent reference to previous sets to make sure that decisions on score points were consistent.

This iterative process of reading, charting, and discussing successive sets 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 papers in earlier sets became firm.





## 1.7 The 2008 MSA-Math Operational Item Analyses

### Classical Analysis with Form-to-Form Common Items

As mentioned in chapter 1.4, two operational forms were linked using common items appearing on both forms (i.e., operational forms A and F) and randomly distributed to students. 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 1.24.

**Table 1.24 Descriptive Statistics of Form-to-Form Common Items**

Grade	Form	No. of Items	$N$	$M$	$SD$
3	A	44	29,364	33.93	8.28
	F	44	29,253	34.46	8.05
4	A	36	30,101	26.77	7.76
	F	36	29,933	27.05	7.55
5	A	51	30,537	37.29	11.88
	F	51	30,289	37.83	11.65
6	A	44	31,060	29.97	10.56
	F	44	30,292	30.79	10.23
7	A	30	31,804	18.56	7.22
	F	30	31,048	18.73	7.06
8	A	35	32,318	21.94	9.60
	F	35	31,743	22.38	9.46

*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 year's assessment was linked using core linking items. This section was prepared to provide information about how much p-values (i.e., classical item difficulty) of the 2008 core linking items varied from previous years.

First of all, it should be noted that detailed information about Rasch analysis on these core linking items can be found in chapter 1.10, *Calibration, Equating, Scaling*. Second, only SR items were used for the purpose of year-to-year linking. Third, classical analysis (e.g., p-value) on these items was conducted with a statewide population, and item sequence number on the tables was assigned based on the 2008 assessment. As seen from Tables 1.25 through 1.36, we could concluded that most of the 2008 p-values were almost the same or slightly increased compared to those of previous years across all grades.

**Table 1.25 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 3 Form A**

Item Seq. No.	Item CID	Previous Year	2008	Item Seq. No.	Item CID	Previous Year	2008
1	3509931	0.65	0.69	48	3510065	0.96	0.94
2	3548059	0.71	0.75	49	3510063	0.78	0.78
5	3510009	0.79	0.84	50	100000044158	0.77	0.86
6	3509974	0.66	0.65	51	3510018	0.77	0.78
7	3548057	0.73	0.80	52	3510035	0.87	0.88
8	3509955	0.57	0.61	55	3510055	0.62	0.62
14	3509959	0.70	0.70	56	3510027	0.87	0.87
16	3509960	0.76	0.78	62	3510347	0.68	0.74
17	3509964	0.74	0.79	63	3510053	0.84	0.84
21	3510068	0.81	0.84	64	3510058	0.86	0.88
22	3510022	0.47	0.51	65	3510051	0.54	0.57
23	3509927	0.78	0.80	66	3509929	0.54	0.53
24	3510006	0.61	0.59	67	3510329	0.55	0.55
32	3509935	0.67	0.61	68	3510033	0.79	0.82
33	3510066	0.80	0.80	69	3510043	0.76	0.77
41	3510125	0.52	0.56	70	3510012	0.78	0.80
44	100000044163	0.85	0.76	72	3509962	0.88	0.90
45	3509926	0.36	0.39	82	3510036	0.85	0.85
47	3509961	0.92	0.91				

*Note.* Analysis was conducted with a statewide population.

*Note.* Item sequence numbers were assigned based on the 2008 assessment.

### Descriptive Statistics of Year-to-Year Core Linking Items: Grade 3 Form A

Form	Year	No. of Items	<i>M</i>	<i>SD</i>
A	Previous Year	37	0.73	0.14
	Year 2008	37	0.74	0.13

**Table 1.26 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 3 Form F**

Item Seq. No.	Item CID	Previous FA	Y08 FA	Item Seq. No.	Item CID	Previous FA	Y08 FA
1	3509931	0.65	0.70	48	3510065	0.96	0.94
2	3548059	0.71	0.76	49	3510063	0.78	0.78
5	3510009	0.79	0.84	50	100000044158	0.77	0.85
6	3509974	0.66	0.66	51	3510018	0.77	0.79
7	3548057	0.73	0.81	52	3510035	0.87	0.89
8	3509955	0.57	0.63	55	3510055	0.62	0.62
14	3509959	0.70	0.72	56	3510027	0.87	0.87
16	3509960	0.76	0.81	62	3510347	0.68	0.76
17	3509964	0.74	0.82	63	3510053	0.84	0.85
18	3509956	0.64	0.65	64	3510058	0.86	0.88
21	3510068	0.81	0.87	65	3510051	0.54	0.58
22	3510022	0.47	0.52	66	3509929	0.54	0.55
23	3509927	0.78	0.80	67	3510329	0.55	0.56
24	3510006	0.61	0.59	68	3510033	0.79	0.84
29	3510126	0.78	0.76	69	3510043	0.76	0.79
31	100000044154	0.81	0.87	70	3510012	0.78	0.80
32	3509935	0.67	0.63	72	3509962	0.88	0.91
33	3510066	0.80	0.81	76	3510020	0.82	0.84
45	3509926	0.36	0.47	82	3510036	0.85	0.85
47	3509961	0.92	0.92				

*Note.* Analysis was conducted with a statewide population.

*Note.* Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 3 Form F**

Form	Year	No. of Items	<i>M</i>	<i>SD</i>
F	Previous Year	39	0.73	0.13
	Year 2008	39	0.76	0.12

**Table 1.27 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 4 Form A**

Item Seq. No.	Item CID	Previous FA	Y08 FA	Item Seq. No.	Item CID	Previous FA	Y08 FA
1	3515406	0.60	0.64	49	3515471	0.86	0.86
2	3515407	0.85	0.86	50	3515630	0.50	0.52
3	100000044146	0.89	0.91	53	3515787	0.51	0.54
6	3515408	0.68	0.76	54	3515533	0.85	0.84
7	3515641	0.83	0.79	55	3515631	0.77	0.78
8	3515410	0.81	0.87	56	3515486	0.57	0.59
10	3515605	0.53	0.61	57	3515484	0.92	0.91
19	3515447	0.45	0.52	63	3515543	0.79	0.80
22	3515604	0.64	0.69	64	3515853	0.71	0.80
24	3515576	0.61	0.65	66	3548078	0.50	0.49
25	3515470	0.69	0.73	67	3515933	0.76	0.76
26	3515643	0.38	0.42	68	3515519	0.82	0.86
27	3515645	0.71	0.72	69	3515795	0.60	0.65
30	3515559	0.72	0.69	70	3515545	0.86	0.87
31	3515426	0.44	0.48	71	3548086	0.76	0.81
32	3515571	0.85	0.80	78	3515506	0.89	0.90
34	3515421	0.82	0.85	79	3515887	0.89	0.86
35	3515574	0.85	0.86	80	3515632	0.71	0.69
47	3515575	0.77	0.88	81	3548088	0.74	0.75
48	3515705	0.75	0.81				

*Note.* Analysis was conducted with a statewide population.

*Note.* Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 4 Form A**

Form	Year	<i>N</i>	<i>M</i>	<i>SD</i>
A	Previous Year	39	0.71	0.15
	Year 2008	39	0.74	0.14

**Table 1.28 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 4 Form F**

Item Seq. No.	Item CID	Previous FF	Y08 FF	Item Seq. No.	Item CID	Previous FF	Y08 FF
1	3515406	0.60	0.65	50	3515630	0.50	0.54
2	3515407	0.85	0.86	55	3515631	0.77	0.79
6	3515408	0.68	0.77	56	3515486	0.57	0.58
7	3515641	0.83	0.81	57	1000000441	0.53	0.75
8	3515410	0.81	0.87	64	3515853	0.71	0.79
10	3515605	0.53	0.62	65	3515836	0.58	0.58
19	3515447	0.45	0.53	66	3548078	0.50	0.51
22	3515604	0.64	0.68	67	3515933	0.76	0.77
24	3515576	0.61	0.65	68	3515635	0.60	0.54
25	3515470	0.69	0.73	69	3515795	0.60	0.63
26	3515643	0.38	0.44	70	3515545	0.86	0.88
27	3515645	0.71	0.74	71	3548086	0.76	0.80
32	3515571	0.85	0.81	77	3548079	0.94	0.95
33	100000044145	0.86	0.95	78	3515506	0.89	0.92
34	3515421	0.82	0.85	79	3515887	0.89	0.90
47	3515575	0.77	0.87	80	3515632	0.71	0.71
49	3515471	0.86	0.86	81	3548088	0.74	0.76

Note. Analysis was conducted with a statewide population.

Note. Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 4 Form F**

Form	Year	<i>N</i>	<i>M</i>	<i>SD</i>
F	Previous Year	34	0.70	0.14
	Year 2008	34	0.74	0.14

**Table 1.29 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 5 Form A**

Item Seq. No.	Item CID	Previous FA	Y08 FA	Item Seq. No.	Item CID	Previous FA	Y08 FA
1	3511312	0.39	0.42	40	3511479	0.51	0.61
2	3511269	0.81	0.88	41	3511504	0.90	0.88
8	3511203	0.87	0.91	43	3511513	0.85	0.85
10	3512535	0.46	0.55	47	3511266	0.71	0.70
16	3511196	0.55	0.58	49	3511470	0.81	0.86
17	3511307	0.41	0.42	50	3511499	0.63	0.63
19	3511467	0.85	0.82	51	3511330	0.63	0.61
20	3512529	0.56	0.58	55	3512595	0.79	0.80
21	3511339	0.62	0.66	56	3511521	0.67	0.62
23	100000043853	0.57	0.67	59	3511376	0.81	0.88
26	3511216	0.67	0.71	60	3511396	0.84	0.88
27	3512638	0.64	0.74	61	3511429	0.75	0.77
28	3512691	0.52	0.60	69	3512625	0.88	0.90
34	3512702	0.54	0.50	70	3511631	0.76	0.78
37	3511566	0.66	0.66	72	3511439	0.79	0.77
38	3511246	0.78	0.76	79	3511442	0.61	0.62
39	3511458	0.92	0.87	83	3511448	0.76	0.77

Note. Analysis was conducted with a statewide population.

Note. Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 5 Form A**

Form	Year	<i>N</i>	<i>M</i>	<i>SD</i>
A	Previous Year	34	0.69	0.15
	Year 2008	34	0.71	0.14

**Table 1.30 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 5 Form F**

Item Seq. No.	Item CID	Previous FF	Y08 FF	Item Seq. No.	Item CID	Previous FF	Y08 FF
1	3511312	0.39	0.42	44	3512632	0.39	0.42
2	3511269	0.81	0.89	47	3511266	0.71	0.70
8	3511203	0.87	0.92	49	3511470	0.81	0.87
10	3512535	0.46	0.56	50	3511499	0.63	0.64
16	3511196	0.55	0.59	51	3511330	0.63	0.62
17	3511307	0.41	0.42	55	3512595	0.79	0.80
19	3511467	0.85	0.82	56	3511521	0.67	0.63
21	3511339	0.62	0.67	59	3511376	0.81	0.88
23	100000043853	0.57	0.68	60	3511396	0.84	0.88
26	3511216	0.67	0.70	61	3511429	0.75	0.77
38	3511246	0.78	0.77	71	3512628	0.77	0.82
39	3511458	0.92	0.90	72	3511439	0.79	0.79
40	3511479	0.51	0.63	79	3511442	0.61	0.63
41	3511504	0.90	0.89	82	100000043851	0.64	0.66
43	3511513	0.85	0.87	83	3511448	0.76	0.79

*Note.* Analysis was conducted with a statewide population.

*Note.* Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 5 Form F**

Form	Year	<i>N</i>	<i>M</i>	<i>SD</i>
F	Previous Year	30	0.69	0.15
	Year 2008	30	0.72	0.15

**Table 1.31 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 6 Form A**

Item Seq. No.	Item CID	Previous FA	Y08 FA	Item Seq. No.	Item CID	Previous FA	Y08 FA
1	3516257	0.83	0.88	37	3516329	0.62	0.60
3	3516291	0.47	0.53	38	3516355	0.66	0.70
5	3516295	0.65	0.70	44	3516351	0.51	0.52
6	3516243	0.69	0.72	46	3516249	0.67	0.67
9	3516248	0.75	0.83	49	3516573	0.67	0.75
10	3516559	0.84	0.91	51	3516242	0.38	0.47
11	3516255	0.70	0.77	52	3516281	0.44	0.50
12	3516258	0.54	0.61	53	3516354	0.72	0.70
13	3516298	0.29	0.36	55	3516332	0.51	0.52
19	3516240	0.56	0.64	56	3516256	0.60	0.61
21	3516283	0.43	0.48	57	3516302	0.69	0.69
25	3516285	0.54	0.58	62	3517000	0.51	0.58
26	3516290	0.64	0.75	68	3516613	0.55	0.54
33	3516453	0.76	0.78	69	3516313	0.83	0.79
34	3516331	0.41	0.49	70	3516318	0.88	0.87
35	3516241	0.84	0.84	79	3516323	0.67	0.69
36	3516247	0.55	0.60	80	3516303	0.53	0.55

Note. Analysis was conducted with a statewide population.

Note. Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 6 Form A**

Form	Year	<i>N</i>	<i>M</i>	<i>SD</i>
A	Previous Year	34	0.61	0.15
	Year 2008	34	0.65	0.14



**Table 1.32 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 6 Form F**

Item Seq. No.	Item CID	Previous FF	Y08 FF	Item Seq. No.	Item CID	Previous FF	Y08 FF
3	3516291	0.47	0.54	38	3516355	0.66	0.72
4	3516625	0.84	0.91	44	3516351	0.51	0.56
5	3516295	0.65	0.70	45	3516565	0.44	0.54
6	3516243	0.69	0.73	46	3516249	0.67	0.70
9	3516248	0.75	0.85	49	3516573	0.67	0.78
10	3516559	0.84	0.92	51	3516242	0.38	0.47
11	3516255	0.70	0.77	52	1000000438	0.72	0.69
19	3516240	0.56	0.65	53	3516354	0.72	0.67
21	3516283	0.43	0.50	55	3516332	0.51	0.54
25	3516285	0.54	0.59	56	3516256	0.60	0.63
26	3516290	0.64	0.75	57	3516302	0.69	0.70
33	3516453	0.76	0.85	62	3517000	0.51	0.57
34	3516331	0.41	0.50	68	3516613	0.55	0.52
35	3516241	0.84	0.85	69	3516313	0.83	0.81
36	3516247	0.55	0.62	70	3516318	0.88	0.89
37	3516329	0.62	0.65	80	3516303	0.53	0.58

Note. Analysis was conducted with a statewide population.

Note. Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 6 Form F**

Form	Year	<i>N</i>	<i>M</i>	<i>SD</i>
F	Previous Year	32	0.63	0.14
	Year 2008	32	0.68	0.13

**Table 1.33 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 7 Form A**

Item Seq. No.	Item CID	Previous FA	Y08 FA	Item Seq. No.	Item CID	Previous FA	Y08 FA
1	3517604	0.32	0.34	31	3517678	0.88	0.92
2	3517601	0.45	0.51	32	3517742	0.50	0.59
3	3517609	0.50	0.58	42	3517710	0.61	0.69
4	3517613	0.62	0.69	43	3517656	0.63	0.65
7	3517616	0.55	0.63	49	3547535	0.76	0.81
8	3517634	0.63	0.67	51	3517687	0.56	0.57
9	3517642	0.42	0.48	52	3517692	0.79	0.83
10	3517638	0.69	0.77	64	3517714	0.54	0.55
12	3517650	0.60	0.66	65	3517716	0.61	0.68
18	3517652	0.66	0.69	66	3517718	0.61	0.70
19	3547473	0.77	0.80	69	3517721	0.42	0.52
20	3517663	0.27	0.32	71	3517709	0.64	0.68
27	3517665	0.35	0.37	79	3555859	0.74	0.74
30	3517667	0.57	0.53	80	3517752	0.62	0.64

*Note.* Analysis was conducted with a statewide population.

*Note.* Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 7 Form A**

Form	Year	<i>N</i>	<i>M</i>	<i>SD</i>
A	Previous Year	31	0.57	0.15
	Year 2008	31	0.61	0.14

**Table 1.34 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 7 Form F**

Item Seq. No.	Item CID	Previous FF	Y08 FF	Item Seq. No.	Item CID	Previous FF	Y08 FF
1	3517604	0.32	0.34	31	3517678	0.88	0.92
2	3517601	0.45	0.52	32	3517742	0.50	0.59
3	3517609	0.50	0.60	42	3517710	0.61	0.75
4	3517613	0.62	0.70	43	3517656	0.63	0.66
7	3517616	0.55	0.64	49	3547535	0.76	0.81
8	3517634	0.63	0.68	51	3517687	0.56	0.58
9	3517642	0.42	0.48	52	3517692	0.79	0.82
10	3517638	0.69	0.78	64	3517714	0.54	0.59
12	3517650	0.60	0.58	65	3517716	0.61	0.68
18	3517652	0.66	0.73	66	3517718	0.61	0.70
19	3547473	0.77	0.81	69	3517721	0.42	0.52
20	3517663	0.27	0.32	71	3517709	0.64	0.68
27	3517665	0.35	0.38	79	3555859	0.74	0.75
30	3517667	0.57	0.50	80	3517752	0.62	0.65

*Note.* Analysis was conducted with a statewide population.

*Note.* Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 7 Form F**

Form	Year	<i>N</i>	<i>M</i>	<i>SD</i>
F	Previous Year	28	0.58	0.14
	Year 2008	28	0.63	0.15

**Table 1.35 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 8 Form A**

Item Seq. No.	Item CID	Previous FA	Y08 FA	Item Seq. No.	Item CID	Previous FA	Y08 FA
1	3514015	0.23	0.28	46	3514055	0.57	0.56
2	3514014	0.56	0.57	47	3514052	0.50	0.53
5	3514016	0.75	0.78	52	3514074	0.42	0.42
7	3514053	0.71	0.73	53	3514075	0.63	0.65
22	3514059	0.63	0.64	58	3514092	0.42	0.43
32	3514058	0.30	0.33	64	3514095	0.31	0.31
33	3514062	0.41	0.43	66	100000043309	0.13	0.18
38	3514291	0.73	0.75	67	3514103	0.60	0.68
41	100000043323	0.36	0.49	79	3514710	0.53	0.54
42	3514057	0.65	0.68	80	3514139	0.73	0.68
43	3514121	0.69	0.71				

*Note.* Analysis was conducted with a statewide population.

*Note.* Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 8 Form A**

Form	Year	<i>N</i>	<i>M</i>	<i>SD</i>
A	Previous Year	21	0.52	0.18
	Year 2008	21	0.54	0.17

**Table 1.36 P-Value Comparisons of Core Linking Items for Previous Year vs. Year 2008: Grade 8 Form F**

Item Seq. No.	Item CID	Previous FF	Y08 FF	Item Seq. No.	Item CID	Previous FF	Y08 FF
1	3514015	0.23	0.28	46	3514055	0.57	0.59
2	3514014	0.56	0.57	47	3514052	0.50	0.53
5	3514016	0.75	0.79	50	3514056	0.79	0.77
7	3514053	0.71	0.75	52	3514074	0.42	0.46
22	3514059	0.63	0.67	53	3514075	0.63	0.66
32	3514058	0.30	0.35	58	3514092	0.42	0.44
33	3514062	0.41	0.44	64	3514095	0.31	0.31
38	3514291	0.73	0.77	65	3514174	0.58	0.52
41	100000043323	0.36	0.50	67	3514103	0.60	0.70
42	3514057	0.65	0.69	79	3514710	0.53	0.54
43	3514121	0.69	0.71	80	3514139	0.73	0.69

*Note.* Analysis was conducted with a statewide population.

*Note.* Item sequence numbers were assigned based on the 2008 assessment.

**Descriptive Statistics of Year-to-Year Core Linking Items: Grade 8 Form F**

Form	Year	<i>N</i>	<i>M</i>	<i>SD</i>
F	Previous Year	22	0.55	0.16
	Year 2008	22	0.58	0.15



### **Validation Check with the 2008 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 2008 year-to-year linking pool, Rasch item and step difficulty parameters of the core items were reestimated with the 2008 random samples during calibration and equating. (Please see section 1.10 and appendix A for 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 2008 core and core linking items can be found in section 1.4, *Test Form Design, Specifications, Item Type, and Item Roles*.

As previously mentioned, 2008 Forms A, B, C, D, and E (Operational Form A) are the same, and Year 2008 Forms F, G, H, J, and K (Operational Form F) are the same except for field test items. A smaller number of cases (approximately 2,500) were available for conducting field test analyses. Both BCR and ECR item p-values 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 Year 2008 p-values were almost the same or somewhat increased compared to those in previous years across all grades.

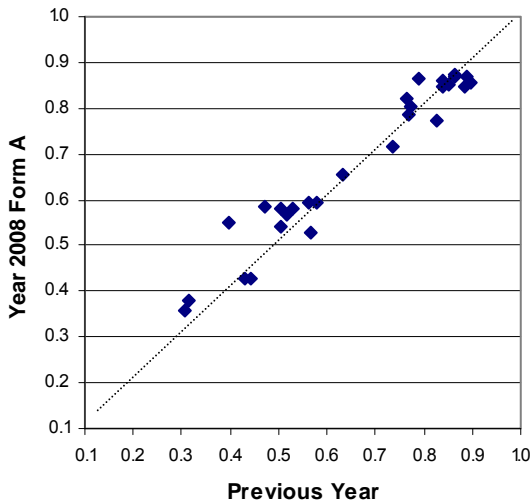
With respect to Rasch item calibration, it should be at first noted that we coded “Omit” of each item as “missing” before we ran the data with the Rasch model. In general, most of the 2008 items were almost the same or somewhat easier compared to those in previous years across all grades. It should be 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 most of the items became easier.

**Table 1.37 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 3 Form A**

Item CID	Previous Year	Year 08 Form A	Item CID	Previous Year	Year 08 Form A
<b>3509918</b>	<b>0.76</b>	<b>0.82</b>	<b>3510073</b>	<b>0.77</b>	<b>0.79</b>
<b>3595500</b>	<b>0.50</b>	<b>0.54</b>	<b>3595503</b>	<b>0.47</b>	<b>0.58</b>
100000044161	0.63	0.66	<b>3510072</b>	<b>0.85</b>	<b>0.85</b>
3488196	0.90	0.85	<b>3595504</b>	<b>0.58</b>	<b>0.60</b>
3488126	0.89	0.87	100000044152	0.79	0.86
<b>3509941</b>	<b>0.50</b>	<b>0.58</b>	<b>3510060</b>	<b>0.84</b>	<b>0.86</b>
<b>3595501</b>	<b>0.40</b>	<b>0.55</b>	<b>3595505</b>	<b>0.53</b>	<b>0.58</b>
3496696	0.83	0.77	3487779	0.84	0.85
<b>3509957</b>	<b>0.77</b>	<b>0.80</b>	<b>3510034</b>	<b>0.30</b>	<b>0.36</b>
<b>3595502</b>	<b>0.43</b>	<b>0.43</b>	<b>3595506</b>	<b>0.32</b>	<b>0.38</b>
3488123	0.56	0.60	3488178	0.57	0.53
3548507	0.88	0.85	3496700	0.86	0.87
100000044159	0.52	0.57	3509950	0.73	0.72
3488038	0.44	0.43	3490570	0.86	0.87

Note. Bold-faced item indicates a BCR item.





**Table 1.38 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 3 Form A**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2007	3509918	BCR_A	29,897	0.76	0.43	22.75	76.32	N/A	N/A	0.93
2007	3595500	BCR_B	29,897	1.01	0.66	19.76	55.87	22.43	N/A	1.94
2006	3509941	BCR_A	2,845	0.50	0.50	48.61	50.40	N/A	N/A	0.98
2006	3595501	BCR_B	2,845	0.79	0.44	29.67	56.63	11.35	N/A	2.36
2007	3509957	BCR_A	29,897	0.77	0.42	21.66	77.18	N/A	N/A	1.16
2007	3595502	BCR_B	29,897	0.86	0.61	24.80	60.79	12.82	N/A	1.59
2006	3510073	BCR_A	2,860	0.77	0.42	22.66	76.92	N/A	N/A	0.42
2006	3595503	BCR_B	2,860	0.94	0.39	17.73	68.64	12.87	N/A	0.77
2007	3510072	BCR_A	29,897	0.85	0.36	13.93	85.13	N/A	N/A	0.94
2007	3595504	BCR_B	29,897	1.16	0.79	22.53	35.47	40.07	N/A	1.93
2007	3510060	BCR_A	29,897	0.84	0.37	13.69	83.95	N/A	N/A	2.36
2007	3595505	BCR_B	29,897	1.06	0.61	13.95	62.34	21.71	N/A	1.99
2007	3510034	BCR_A	29,897	0.30	0.46	68.51	30.47	N/A	N/A	1.02
2007	3595506	BCR_B	29,897	0.63	0.62	42.80	47.44	7.89	N/A	1.87
2008	3509918	BCR_A	29,364	0.82	0.38	17.44	82.14	N/A	N/A	0.41
2008	3595500	BCR_B	29,364	1.08	0.59	12.73	64.81	21.51	N/A	0.95
2008	3509941	BCR_A	29,364	0.58	0.49	41.23	57.86	N/A	N/A	0.91
2008	3595501	BCR_B	29,364	1.10	0.61	12.19	62.16	23.74	N/A	1.91
2008	3509957	BCR_A	29,364	0.80	0.40	18.50	80.16	N/A	N/A	1.33
2008	3595502	BCR_B	29,364	0.85	0.62	25.91	59.61	12.79	N/A	1.68
2008	3510073	BCR_A	29,364	0.79	0.41	20.72	78.64	N/A	N/A	0.65
2008	3595503	BCR_B	29,364	1.17	0.57	8.16	64.41	26.25	N/A	1.17
2008	3510072	BCR_A	29,364	0.85	0.36	14.45	84.99	N/A	N/A	0.56
2008	3595504	BCR_B	29,364	1.19	0.79	22.14	33.94	42.60	N/A	1.32
2008	3510060	BCR_A	29,364	0.86	0.35	12.57	86.14	N/A	N/A	1.29
2008	3595505	BCR_B	29,364	1.16	0.61	11.21	60.07	27.80	N/A	0.91
2008	3510034	BCR_A	29,364	0.36	0.48	63.43	35.62	N/A	N/A	0.95
2008	3595506	BCR_B	29,364	0.76	0.64	33.30	54.01	11.12	N/A	1.57

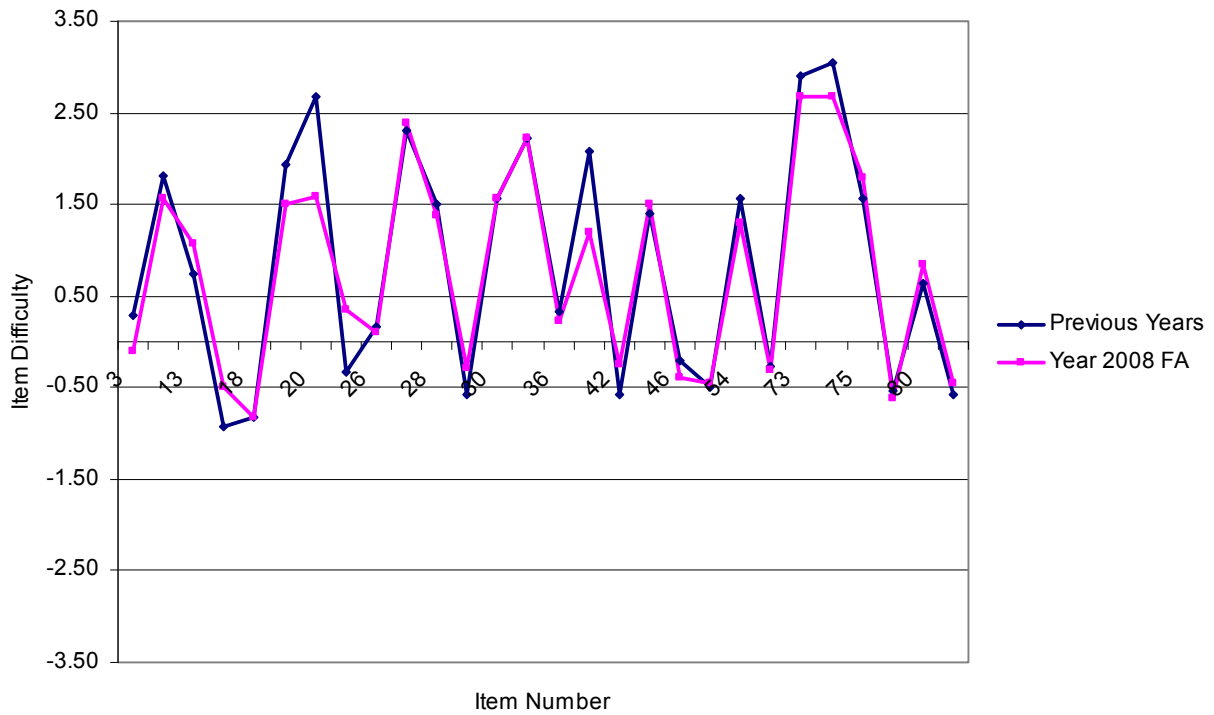
**Table 1.39 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 3 Form A**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2007	3	3509918	BCR_A	0.2848		
2007	4	3595500	BCR_B	1.8054	-1.5584	1.5584
2004	13	100000044161	SR	0.7424		
2007	15	3488196	SR	-0.9320		
2007	18	3488126	SR	-0.8244		
2006	19	3509941	BCR_A	1.9297		
2006	20	3595501	BCR_B	2.6804	-1.6811	1.6811
2007	25	3496696	SR	-0.3314		
2007	26	3509957	BCR_A	0.1695		
2007	27	3595502	BCR_B	2.3042	-1.8051	1.8051
2007	28	3488123	SR	1.5067		
2007	29	3548507	SR	-0.5758		
2005	30	100000044159	SR	1.5728		
2007	31	3488038	SR	2.2301		
2006	36	3510073	BCR_A	0.3226		
2006	37	3595503	BCR_B	2.0893	-2.0585	2.0585
2007	42	3510072	BCR_A	-0.5702		
2007	43	3595504	BCR_B	1.3990	-0.6985	0.6985
2004	46	100000044152	SR	-0.2072		
2007	53	3510060	BCR_A	-0.4888		
2007	54	3595505	BCR_B	1.5699	-1.8116	1.8116
2007	71	3487779	SR	-0.2575		
2007	73	3510034	BCR_A	2.8934		
2007	74	3595506	BCR_B	3.0491	-1.5541	1.5541
2007	75	3488178	SR	1.5629		
2007	76	3496700	SR	-0.5360		
2006	80	3509950	SR	0.6292		
2007	81	3490570	SR	-0.5833		
2008	3	3509918	BCR_A	-0.1008		
2008	4	3595500	BCR_B	1.5690	-1.8761	1.8761
2008	13	100000044161	SR	1.0707		
2008	15	3488196	SR	-0.4959		
2008	18	3488126	SR	-0.8251		
2008	19	3509941	BCR_A	1.5122		
2008	20	3595501	BCR_B	1.5891	-1.8002	1.8002
2008	25	3496696	SR	0.3479		
2008	26	3509957	BCR_A	0.1115		
2008	27	3595502	BCR_B	2.3975	-1.7280	1.7280
2008	28	3488123	SR	1.3765		
2008	29	3548507	SR	-0.2784		
2008	30	100000044159	SR	1.5712		
2008	31	3488038	SR	2.2208		

**Table 1.39 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2008	36	3510073	BCR_A	0.2315		
2008	37	3595503	BCR_B	1.1948	-1.9764	1.9764
2008	42	3510072	BCR_A	-0.2447		
2008	43	3595504	BCR_B	1.5000	-0.5243	0.5243
2008	46	10000044152	SR	-0.3864		
2008	53	3510060	BCR_A	-0.4451		
2008	54	3595505	BCR_B	1.3069	-1.7023	1.7023
2008	71	3487779	SR	-0.2992		
2008	73	3510034	BCR_A	2.6680		
2008	74	3595506	BCR_B	2.6710	-1.4737	1.4737
2008	75	3488178	SR	1.7968		
2008	76	3496700	SR	-0.6248		
2008	80	3509950	SR	0.8475		
2008	81	3490570	SR	-0.4471		

Note. Rasch item and step difficulties are on a common scale.

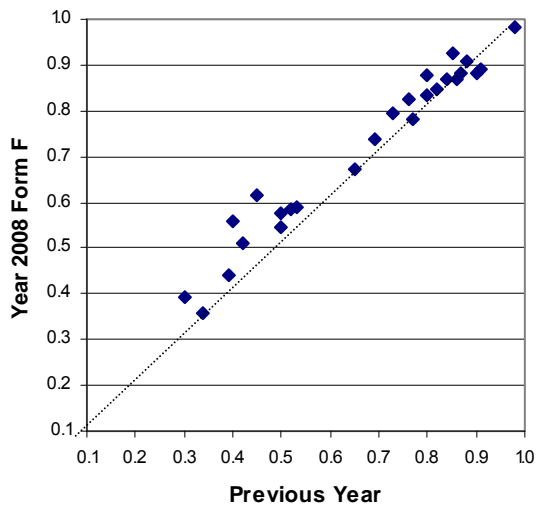


**Figure 1.1 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 3 Form A**

**Table 1.40 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 3 Form F**

Item CID	Previous Year	Year 08 Form F	Item CID	Previous Year	Year 08 Form F
<b>3509918</b>	<b>0.76</b>	<b>0.83</b>	<b>3509924</b>	<b>0.45</b>	<b>0.62</b>
<b>3595500</b>	<b>0.50</b>	<b>0.55</b>	<b>3595509</b>	<b>0.30</b>	<b>0.39</b>
10000044160	0.85	0.93	3488171	0.69	0.74
3488196	0.90	0.88	3488127	0.77	0.78
<b>3509941</b>	<b>0.50</b>	<b>0.58</b>	<b>3510060</b>	<b>0.84</b>	<b>0.87</b>
<b>3595501</b>	<b>0.40</b>	<b>0.56</b>	<b>3595505</b>	<b>0.53</b>	<b>0.59</b>
3487972	0.42	0.51	3488033	0.87	0.88
<b>3509922</b>	<b>0.65</b>	<b>0.67</b>	<b>3509932</b>	<b>0.98</b>	<b>0.98</b>
<b>3595507</b>	<b>0.34</b>	<b>0.36</b>	<b>3595510</b>	<b>0.39</b>	<b>0.44</b>
10000044153	0.80	0.88	3490561	0.88	0.91
10000044159	0.52	0.59	10000044162	0.80	0.83
<b>3510067</b>	<b>0.82</b>	<b>0.85</b>	3490570	0.86	0.87
<b>3595508</b>	<b>0.73</b>	<b>0.79</b>			
3488069	0.91	0.89			

Note. Bold-faced item indicates a BCR item.



**Table 1.41 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 3 Form F**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2007	3509918	BCR_A	29,897	0.76	0.43	22.75	76.32	N/A	N/A	0.93
2007	3595500	BCR_B	29,897	1.01	0.66	19.76	55.87	22.43	N/A	1.94
2006	3509941	BCR_A	2,845	0.50	0.50	48.61	50.40	N/A	N/A	0.98
2006	3595501	BCR_B	2,845	0.79	0.44	29.67	56.63	11.35	N/A	2.36
2007	3509922	BCR_A	29,858	0.65	0.48	32.58	65.40	N/A	N/A	2.02
2007	3595507	BCR_B	29,858	0.68	0.55	34.46	59.79	3.93	N/A	1.82
2007	3510067	BCR_A	29,858	0.82	0.38	16.14	82.47	N/A	N/A	1.40
2007	3595508	BCR_B	29,858	1.47	0.66	7.53	33.92	56.44	N/A	2.11
2006	3509924	BCR_A	2,818	0.45	0.50	53.94	45.28	N/A	N/A	0.78
2006	3595509	BCR_B	2,818	0.60	0.42	43.79	48.83	5.68	N/A	1.70
2007	3510060	BCR_A	29,897	0.84	0.37	13.69	83.95	N/A	N/A	2.36
2007	3595505	BCR_B	29,897	1.06	0.61	13.95	62.34	21.71	N/A	1.99
2007	3509932	BCR_A	29,858	0.98	0.15	1.94	97.58	N/A	N/A	0.48
2007	3595510	BCR_B	29,858	0.78	0.63	31.48	56.08	11.15	N/A	1.28
2008	3509918	BCR_A	29,253	0.83	0.38	17.02	82.58	N/A	N/A	0.40
2008	3595500	BCR_B	29,253	1.09	0.57	11.30	66.53	21.32	N/A	0.84
2008	3509941	BCR_A	29,253	0.58	0.49	41.40	57.72	N/A	N/A	0.88
2008	3595501	BCR_B	29,253	1.12	0.60	10.94	63.07	24.25	N/A	1.74
2008	3509922	BCR_A	29,253	0.67	0.47	30.34	67.24	N/A	N/A	2.42
2008	3595507	BCR_B	29,253	0.71	0.56	32.05	60.86	5.26	N/A	1.83
2008	3510067	BCR_A	29,253	0.85	0.36	14.75	84.62	N/A	N/A	0.62
2008	3595508	BCR_B	29,253	1.59	0.61	5.26	28.70	65.01	N/A	1.03
2008	3509924	BCR_A	29,253	0.62	0.49	37.80	61.59	N/A	N/A	0.60
2008	3595509	BCR_B	29,253	0.78	0.55	27.11	64.61	6.83	N/A	1.46
2008	3510060	BCR_A	29,253	0.87	0.34	12.00	87.08	N/A	N/A	0.92
2008	3595505	BCR_B	29,253	1.18	0.59	9.21	62.34	27.60	N/A	0.85
2008	3509932	BCR_A	29,253	0.98	0.14	1.31	98.07	N/A	N/A	0.63
2008	3595510	BCR_B	29,253	0.88	0.62	24.86	59.60	14.23	N/A	1.32

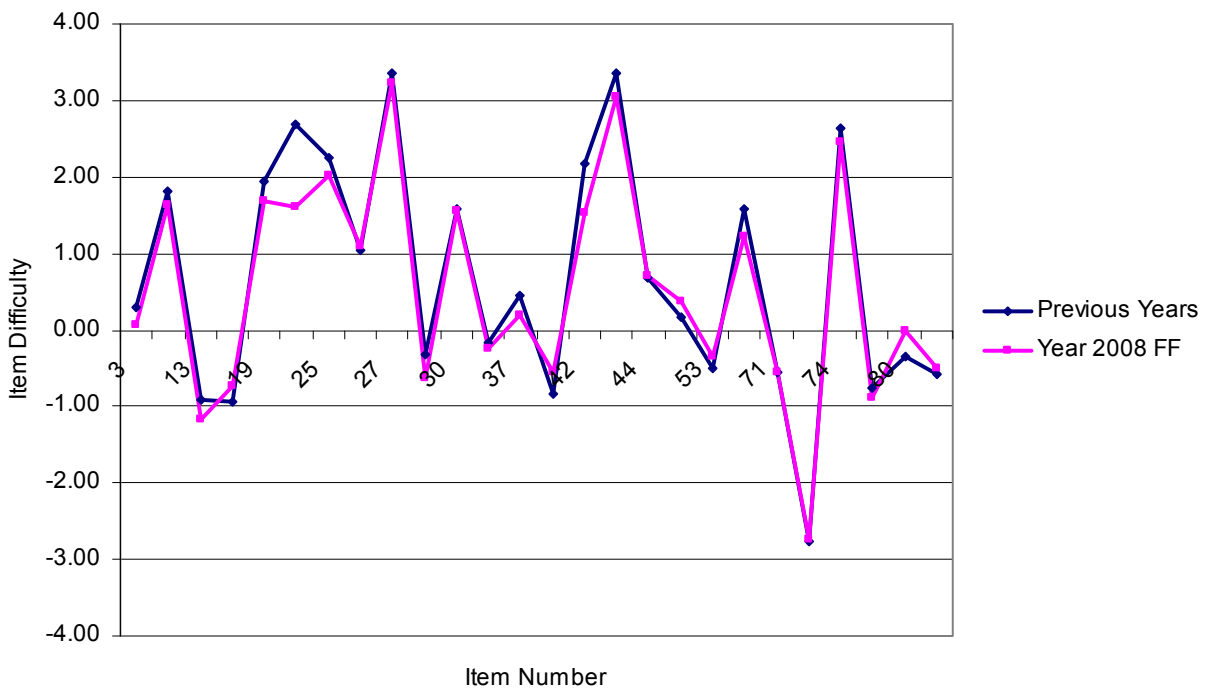
**Table 1.42 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 3 Form F**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2007	3	3509918	BCR_A	0.2848		
2007	4	3595500	BCR_B	1.8054	-1.5584	1.5584
2004	13	100000044160	SR	-0.9180		
2007	15	3488196	SR	-0.9320		
2006	19	3509941	BCR_A	1.9297		
2006	20	3595501	BCR_B	2.6804	-1.6811	1.6811
2007	25	3487972	SR	2.2419		
2007	26	3509922	BCR_A	1.0354		
2007	27	3595507	BCR_B	3.3509	-2.2387	2.2387
2004	28	100000044153	SR	-0.3189		
2005	30	100000044159	SR	1.5728		
2007	36	3510067	BCR_A	-0.1708		
2007	37	3595508	BCR_B	0.4429	-1.0657	1.0657
2007	41	3488069	SR	-0.8242		
2006	42	3509924	BCR_A	2.1650		
2006	43	3595509	BCR_B	3.3575	-1.6247	1.6247
2007	44	3488171	SR	0.6907		
2007	46	3488127	SR	0.1633		
2007	53	3510060	BCR_A	-0.4888		
2007	54	3595505	BCR_B	1.5699	-1.8116	1.8116
2007	71	3488033	SR	-0.5635		
2007	73	3509932	BCR_A	-2.7619		
2007	74	3595510	BCR_B	2.6430	-1.5825	1.5825
2007	75	3490561	SR	-0.7637		
2004	80	100000044162	SR	-0.3580		
2007	81	3490570	SR	-0.5833		
2008	3	3509918	BCR_A	0.0751		
2008	4	3595500	BCR_B	1.6361	-1.9848	1.9848
2008	13	100000044160	SR	-1.1760		
2008	15	3488196	SR	-0.7392		
2008	19	3509941	BCR_A	1.6893		
2008	20	3595501	BCR_B	1.6043	-1.8082	1.8082
2008	25	3487972	SR	2.0271		
2008	26	3509922	BCR_A	1.0913		
2008	27	3595507	BCR_B	3.2375	-2.1210	2.1210
2008	28	100000044153	SR	-0.6268		
2008	30	100000044159	SR	1.5483		
2008	36	3510067	BCR_A	-0.2338		
2008	37	3595508	BCR_B	0.1995	-0.9490	0.9490
2008	41	3488069	SR	-0.5610		

**Table 1.42 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2008	42	3509924	BCR_A	1.5222		
2008	43	3595509	BCR_B	3.0387	-2.0994	2.0994
2008	44	3488171	SR	0.7136		
2008	46	3488127	SR	0.3755		
2008	53	3510060	BCR_A	-0.3562		
2008	54	3595505	BCR_B	1.2218	-1.8923	1.8923
2008	71	3488033	SR	-0.5440		
2008	73	3509932	BCR_A	-2.7407		
2008	74	3595510	BCR_B	2.4652	-1.6902	1.6902
2008	75	3490561	SR	-0.8923		
2008	80	10000044162	SR	-0.0075		
2008	81	3490570	SR	-0.5043		

Note. Rasch item and step difficulties are on a common scale.

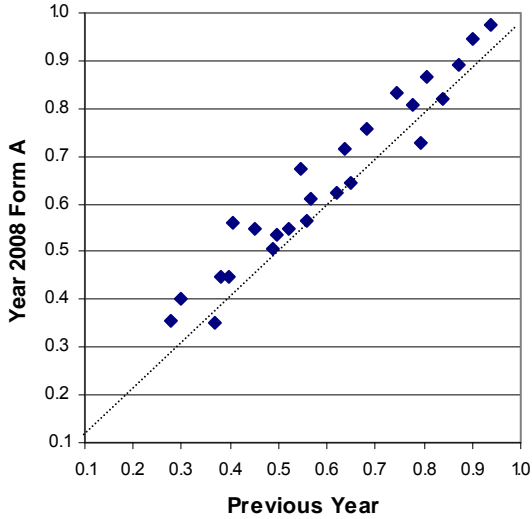


**Figure 1.2 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 3 Form F**

**Table 1.43 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 4 Form A**

Item CID	Previous Year	Year 08 Form A	Item CID	Previous Year	Year 08 Form A
<b>10000044148</b>	<b>0.28</b>	<b>0.35</b>	3548767	0.64	0.71
<b>3595498</b>	<b>0.41</b>	<b>0.56</b>	<b>3515807</b>	<b>0.79</b>	<b>0.73</b>
3487996	0.68	0.76	<b>3595533</b>	<b>0.37</b>	<b>0.35</b>
3488056	0.49	0.51	3488052	0.57	0.61
3488159	0.80	0.87	<b>3515886</b>	<b>0.45</b>	<b>0.55</b>
<b>10000044142</b>	<b>0.84</b>	<b>0.82</b>	<b>3595534</b>	<b>0.52</b>	<b>0.55</b>
<b>3595499</b>	<b>0.40</b>	<b>0.45</b>	3497876	0.62	0.62
3515737	0.75	0.83	3497869	0.78	0.81
<b>3515648</b>	<b>0.50</b>	<b>0.54</b>	<b>3515843</b>	<b>0.87</b>	<b>0.89</b>
<b>3595531</b>	<b>0.56</b>	<b>0.56</b>	<b>3595535</b>	<b>0.55</b>	<b>0.67</b>
10000044144	0.90	0.94	3497867	0.65	0.64
<b>3515823</b>	<b>0.38</b>	<b>0.45</b>			
<b>3595532</b>	<b>0.30</b>	<b>0.40</b>			
10000044149	0.94	0.98			

\*Bold-faced number indicates a BCR item.





**Table 1.44 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 4 Form A**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2005	100000044148	BCR_A	12,555	0.28	0.45	71.14	27.81	N/A	N/A	1.05
2005	3595498	BCR_B	12,555	0.82	0.48	32.66	50.27	15.64	N/A	1.43
2005	100000044142	BCR_A	12,716	0.84	0.37	14.47	84.01	N/A	N/A	1.52
2005	3595499	BCR_B	12,716	0.79	0.38	25.54	66.25	6.43	N/A	1.78
2007	3515648	BCR_A	30,402	0.50	0.50	49.35	49.60	N/A	N/A	1.05
2007	3595531	BCR_B	30,402	1.11	0.75	21.50	42.22	34.60	N/A	1.68
2006	3515823	BCR_A	2,847	0.38	0.49	57.96	37.97	N/A	N/A	4.07
2006	3595532	BCR_B	2,847	0.60	0.43	42.40	47.00	6.39	N/A	4.21
2007	3515807	BCR_A	30,402	0.79	0.41	16.33	79.31	N/A	N/A	4.36
2007	3595533	BCR_B	30,402	0.73	0.62	34.05	54.85	9.32	N/A	1.77
2007	3515886	BCR_A	30,402	0.45	0.50	51.99	44.94	N/A	N/A	3.07
2007	3595534	BCR_B	30,402	1.05	0.61	11.50	62.43	21.10	N/A	4.97
2006	3515843	BCR_A	2,847	0.87	0.34	11.87	87.04	N/A	N/A	1.09
2006	3595535	BCR_B	2,847	1.09	0.41	11.56	65.26	21.99	N/A	1.19
2008	100000044148	BCR_A	30,101	0.35	0.48	64.21	35.45	N/A	N/A	0.34
2008	3595498	BCR_B	30,101	1.12	0.59	10.82	64.20	24.12	N/A	0.87
2008	100000044142	BCR_A	30,101	0.82	0.38	16.89	82.01	N/A	N/A	1.10
2008	3595499	BCR_B	30,101	0.90	0.56	19.62	67.90	10.80	N/A	1.67
2008	3515648	BCR_A	30,101	0.54	0.50	45.53	53.57	N/A	N/A	0.89
2008	3595531	BCR_B	30,101	1.13	0.78	23.19	37.31	37.84	N/A	1.66
2008	3515823	BCR_A	30,101	0.45	0.50	54.09	44.61	N/A	N/A	1.30
2008	3595532	BCR_B	30,101	0.80	0.62	28.76	57.76	11.15	N/A	2.33
2008	3515807	BCR_A	30,101	0.73	0.45	24.02	72.76	N/A	N/A	3.23
2008	3595533	BCR_B	30,101	0.71	0.65	38.63	48.61	11.01	N/A	1.75
2008	3515886	BCR_A	30,101	0.55	0.50	43.74	54.60	N/A	N/A	1.66
2008	3595534	BCR_B	30,101	1.10	0.60	10.98	62.51	23.50	N/A	3.01
2008	3515843	BCR_A	30,101	0.89	0.31	10.31	88.99	N/A	N/A	0.70
2008	3595535	BCR_B	30,101	1.35	0.63	7.45	48.54	43.07	N/A	0.95

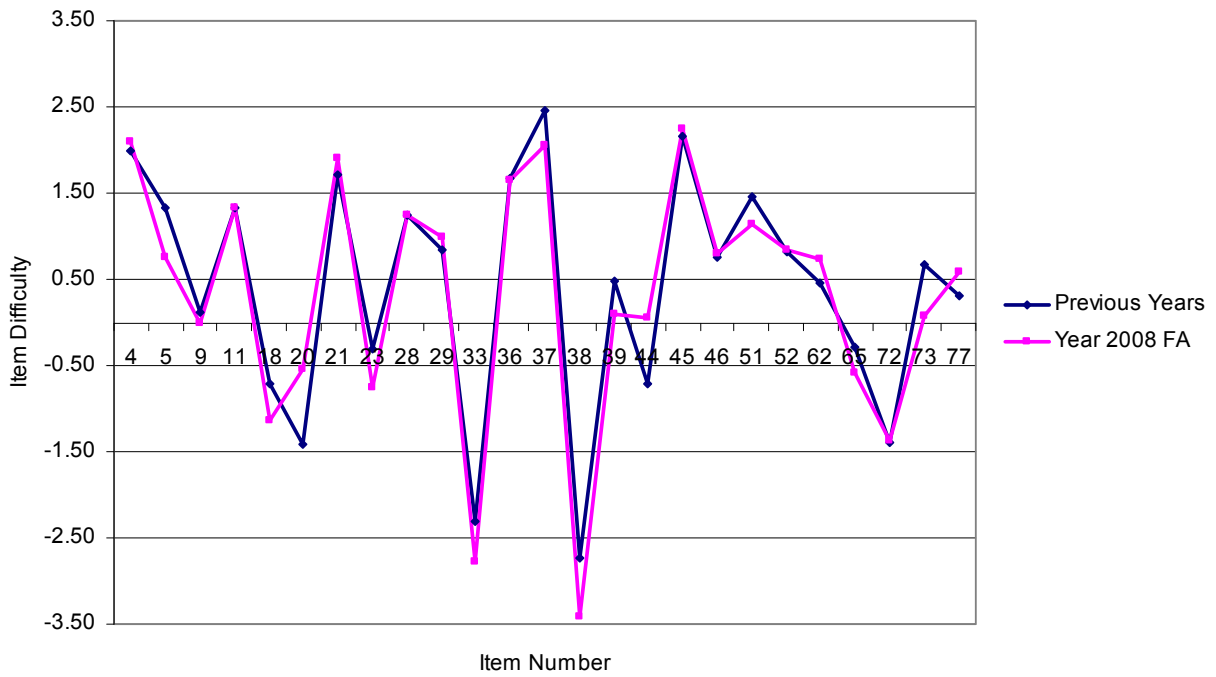
**Table 1.45 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 4 Form A**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2005	4	100000044148	BCR_A	1.9821		
2005	5	3595498	BCR_B	1.3312	-1.2756	1.2756
2007	9	3487996	SR	0.1099		
2007	11	3488056	SR	1.3257		
2007	18	3488159	SR	-0.7155		
2005	20	100000044142	BCR_A	-1.4172		
2005	21	3595499	BCR_B	1.7059	-2.2218	2.2218
2006	23	3515737	SR	-0.3146		
2007	28	3515648	BCR_A	1.2409		
2007	29	3595531	BCR_B	0.8470	-0.9809	0.9809
2004	33	100000044144	SR	-2.3078		
2006	36	3515823	BCR_A	1.6623		
2006	37	3595532	BCR_B	2.4625	-1.5870	1.5870
2004	38	100000044149	SR	-2.7383		
2007	39	3548767	SR	0.4693		
2007	44	3515807	BCR_A	-0.7079		
2007	45	3595533	BCR_B	2.1566	-1.7285	1.7285
2007	46	3488052	SR	0.7650		
2007	51	3515886	BCR_A	1.4586		
2007	52	3595534	BCR_B	0.8111	-1.9929	1.9929
2007	62	3497876	SR	0.4661		
2007	65	3497869	SR	-0.2823		
2006	72	3515843	BCR_A	-1.3830		
2006	73	3595535	BCR_B	0.6659	-1.9576	1.9576
2007	77	3497867	SR	0.3121		
2008	4	100000044148	BCR_A	2.0901		
2008	5	3595498	BCR_B	0.7602	-1.9452	1.9452
2008	9	3487996	SR	-0.0097		
2008	11	3488056	SR	1.3222		
2008	18	3488159	SR	-1.1317		
2008	20	100000044142	BCR_A	-0.5527		
2008	21	3595499	BCR_B	1.8966	-2.1375	2.1375
2008	23	3515737	SR	-0.7461		
2008	28	3515648	BCR_A	1.2519		
2008	29	3595531	BCR_B	0.9986	-0.6913	0.6913
2008	33	100000044144	SR	-2.7781		
2008	36	3515823	BCR_A	1.6468		
2008	37	3595532	BCR_B	2.0517	-1.6746	1.6746
2008	38	100000044149	SR	-3.4118		

**Table 1.45 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2008	39	3548767	SR	0.0868		
2008	44	3515807	BCR_A	0.0553		
2008	45	3595533	BCR_B	2.2474	-1.2842	1.2842
2008	46	3488052	SR	0.7943		
2008	51	3515886	BCR_A	1.1277		
2008	52	3595534	BCR_B	0.8438	-2.0080	2.0080
2008	62	3497876	SR	0.7358		
2008	65	3497869	SR	-0.5937		
2008	72	3515843	BCR_A	-1.3684		
2008	73	3595535	BCR_B	0.0715	-1.6202	1.6202
2008	77	3497867	SR	0.5913		

Note. Rasch item and step difficulties are on a common scale.

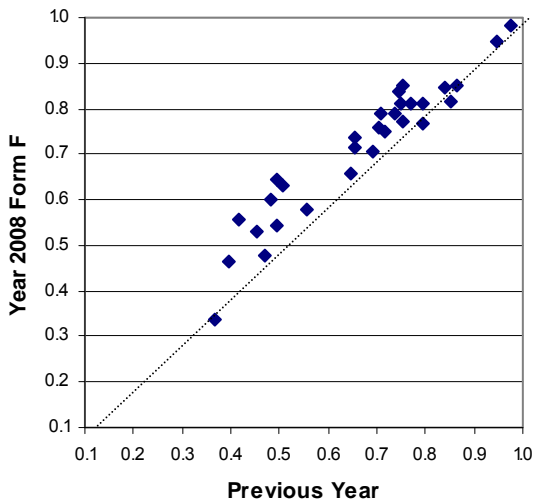


**Figure 1.3 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 4 Form A**

**Table 1.46 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 4 Form F**

Item CID	Previous Year	Year 08 Form F	Item CID	Previous Year	Year 08 Form F
100000044150	0.50	0.64	<b>3595537</b>	<b>0.48</b>	<b>0.60</b>
<b>3515595</b>	<b>0.77</b>	<b>0.81</b>	3488190	0.42	0.56
<b>3595536</b>	<b>0.47</b>	<b>0.48</b>	3488060	0.98	0.98
3497882	0.75	0.77	<b>3515807</b>	<b>0.79</b>	<b>0.77</b>
3497866	0.69	0.71	<b>3595533</b>	<b>0.37</b>	<b>0.34</b>
3515582	0.75	0.81	3490562	0.46	0.53
<b>100000044142</b>	<b>0.84</b>	<b>0.85</b>	3488019	0.65	0.66
<b>3595499</b>	<b>0.40</b>	<b>0.46</b>	<b>3515783</b>	<b>0.72</b>	<b>0.75</b>
3515737	0.75	0.84	<b>3595560</b>	<b>0.66</b>	<b>0.74</b>
<b>3515648</b>	<b>0.50</b>	<b>0.54</b>	3515935	0.75	0.85
<b>3595531</b>	<b>0.56</b>	<b>0.58</b>	3515785	0.65	0.71
3551599	0.85	0.82	3488189	0.79	0.81
3488180	0.86	0.85	3502604	0.74	0.79
3488166	0.71	0.79	<b>3515830</b>	<b>0.95</b>	<b>0.95</b>
<b>3515646</b>	<b>0.51</b>	<b>0.63</b>	<b>3595561</b>	<b>0.71</b>	<b>0.76</b>

Note. Bold-faced number indicates a BCR item.



**Table 1.47 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 4 Form F**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2007	3515595	BCR_A	30,103	0.77	0.42	21.63	77.13	N/A	N/A	1.25
2007	3595536	BCR_B	30,103	0.94	0.63	20.75	60.52	16.91	N/A	1.82
2005	100000044142	BCR_A	12,716	0.84	0.37	14.47	84.01	N/A	N/A	1.52
2005	3595499	BCR_B	12,716	0.79	0.38	25.54	66.25	6.43	N/A	1.78
2007	3515648	BCR_A	30,402	0.50	0.50	49.35	49.60	N/A	N/A	1.05
2007	3595531	BCR_B	30,402	1.11	0.75	21.50	42.22	34.60	N/A	1.68
2006	3515646	BCR_A	24,774	0.51	0.50	45.36	50.83	N/A	N/A	3.81
2006	3595537	BCR_B	24,774	0.96	0.63	37.80	19.80	38.27	N/A	4.13
2007	3515807	BCR_A	30,402	0.79	0.41	16.33	79.31	N/A	N/A	4.36
2007	3595533	BCR_B	30,402	0.73	0.62	34.05	54.85	9.32	N/A	1.77
2006	3515783	BCR_A	2,875	0.72	0.45	26.75	71.76	N/A	N/A	1.50
2006	3595560	BCR_B	2,875	1.31	0.59	22.64	20.28	55.34	N/A	1.74
2007	3515830	BCR_A	30,103	0.95	0.22	4.40	94.72	N/A	N/A	0.87
2007	3595561	BCR_B	30,103	1.41	0.59	4.28	47.72	46.68	N/A	1.32
2008	3515595	BCR_A	29,933	0.81	0.39	18.51	80.92	N/A	N/A	0.57
2008	3595536	BCR_B	29,933	0.96	0.65	21.99	57.21	19.33	N/A	1.46
2008	100000044142	BCR_A	29,933	0.85	0.36	14.33	84.76	N/A	N/A	0.91
2008	3595499	BCR_B	29,933	0.93	0.54	17.11	70.65	10.95	N/A	1.29
2008	3515648	BCR_A	29,933	0.54	0.50	44.90	54.36	N/A	N/A	0.74
2008	3595531	BCR_B	29,933	1.16	0.77	21.67	38.10	38.83	N/A	1.40
2008	3515646	BCR_A	29,933	0.63	0.48	36.05	62.96	N/A	N/A	0.99
2008	3595537	BCR_B	29,933	1.20	0.88	28.87	18.89	50.70	N/A	1.55
2008	3515807	BCR_A	29,933	0.77	0.42	22.05	76.76	N/A	N/A	1.18
2008	3595533	BCR_B	29,933	0.67	0.60	38.40	53.28	6.88	N/A	1.44
2008	3515783	BCR_A	29,933	0.75	0.43	24.51	74.97	N/A	N/A	0.51
2008	3595560	BCR_B	29,933	1.47	0.73	13.22	24.30	61.52	N/A	0.95
2008	3515830	BCR_A	29,933	0.95	0.22	4.81	94.72	N/A	N/A	0.47
2008	3595561	BCR_B	29,933	1.52	0.59	4.09	38.11	56.98	N/A	0.83

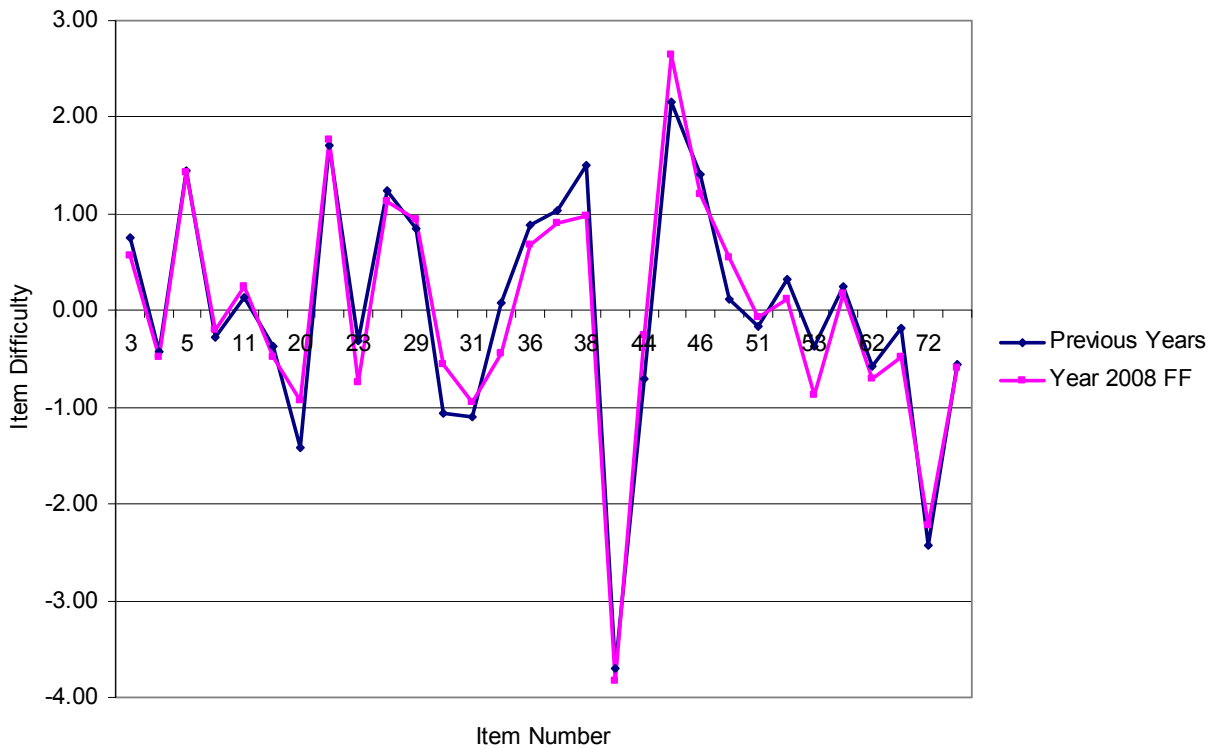
**Table 1.48 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 4 Form F**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2005	3	100000044150	SR	0.7558		
2007	4	3515595	BCR_A	-0.4303		
2007	5	3595536	BCR_B	1.4549	-1.7990	1.7990
2007	9	3497882	SR	-0.2727		
2007	11	3497866	SR	0.1421		
2006	18	3515582	SR	-0.3717		
2005	20	100000044142	BCR_A	-1.4172		
2005	21	3595499	BCR_B	1.7059	-2.2218	2.2218
2006	23	3515737	SR	-0.3146		
2007	28	3515648	BCR_A	1.2409		
2007	29	3595531	BCR_B	0.8470	-0.9809	0.9809
2007	30	3551599	SR	-1.0702		
2007	31	3488180	SR	-1.0902		
2007	35	3488166	SR	0.0725		
2006	36	3515646	BCR_A	0.8899		
2006	37	3595537	BCR_B	1.0287	0.2368	-0.2368
2007	38	3488190	SR	1.5096		
2007	39	3488060	SR	-3.7018		
2007	44	3515807	BCR_A	-0.7079		
2007	45	3595533	BCR_B	2.1566	-1.7285	1.7285
2007	46	3490562	SR	1.4108		
2007	48	3488019	SR	0.1111		
2006	51	3515783	BCR_A	-0.1707		
2006	52	3595560	BCR_B	0.3296	0.1176	-0.1176
2006	53	3515935	SR	-0.3748		
2006	54	3515785	SR	0.2464		
2007	62	3488189	SR	-0.5822		
2007	63	3502604	SR	-0.1844		
2007	72	3515830	BCR_A	-2.4304		
2007	73	3595561	BCR_B	-0.5629	-1.5858	1.5858
2008	3	100000044150	SR	0.5597		
2008	4	3515595	BCR_A	-0.4856		
2008	5	3595536	BCR_B	1.4296	-1.6389	1.6389
2008	9	3497882	SR	-0.1920		
2008	11	3497866	SR	0.2397		
2008	18	3515582	SR	-0.4742		
2008	20	100000044142	BCR_A	-0.9256		
2008	21	3595499	BCR_B	1.7707	-2.2938	2.2938
2008	23	3515737	SR	-0.7523		
2008	28	3515648	BCR_A	1.1199		
2008	29	3595531	BCR_B	0.9347	-0.8467	0.8467
2008	30	3551599	SR	-0.5617		
2008	31	3488180	SR	-0.9544		
2008	35	3488166	SR	-0.4437		

**Table 1.48 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2
2008	36	3515646	BCR_A	0.6734		
2008	37	3595537	BCR_B	0.8984	0.1686	-0.1686
2008	38	3488190	SR	0.9747		
2008	39	3488060	SR	-3.8290		
2008	44	3515807	BCR_A	-0.2581		
2008	45	3595533	BCR_B	2.6447	-1.7531	1.7531
2008	46	3490562	SR	1.1949		
2008	48	3488019	SR	0.5462		
2008	51	3515783	BCR_A	-0.0713		
2008	52	3595560	BCR_B	0.1264	-0.3519	0.3519
2008	53	3515935	SR	-0.8732		
2008	54	3515785	SR	0.1771		
2008	62	3488189	SR	-0.7062		
2008	63	3502604	SR	-0.4877		
2008	72	3515830	BCR_A	-2.2268		
2008	73	3595561	BCR_B	-0.5908	-1.3955	1.3955

Note. Rasch item and step difficulties are on a common scale.

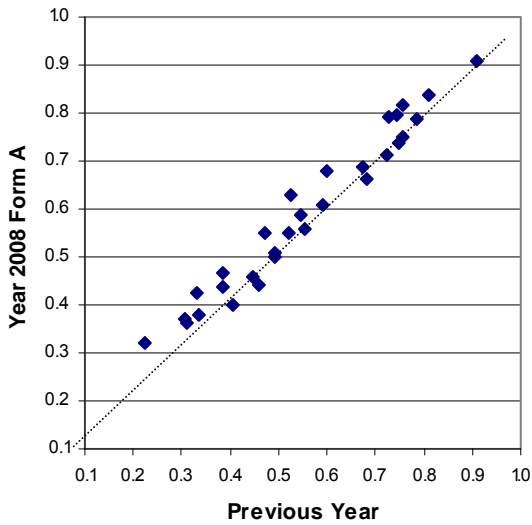


**Figure 1.4 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 4 Form F**

**Table 1.49 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 5 Form A**

Item CID	Previous Year	Year 08 Form A	Item CID	Previous Year	Year 08 Form A
3512642	0.53	0.63	<b>3595441</b>	<b>0.49</b>	<b>0.51</b>
<b>3511531</b>	<b>0.68</b>	<b>0.69</b>	3488431	0.75	0.74
<b>3595438</b>	<b>0.55</b>	<b>0.59</b>	<b>3556476</b>	<b>0.49</b>	<b>0.50</b>
3488390	0.39	0.44	<b>3595442</b>	<b>0.46</b>	<b>0.44</b>
3512622	0.60	0.68	3488241	0.91	0.91
3488506	0.41	0.40	100000043857	0.76	0.82
3488373	0.68	0.66	<b>3512618</b>	<b>0.45</b>	<b>0.46</b>
3512639	0.75	0.80	<b>3595443</b>	<b>0.52</b>	<b>0.55</b>
<b>3512615</b>	<b>0.78</b>	<b>0.79</b>	3512623	0.73	0.79
<b>3595439</b>	<b>0.47</b>	<b>0.55</b>	3488251	0.59	0.61
<b>3511336</b>	<b>0.33</b>	<b>0.43</b>	<b>3512564</b>	<b>0.31</b>	<b>0.36</b>
<b>3595440</b>	<b>0.34</b>	<b>0.38</b>	<b>3595444</b>	<b>0.22</b>	<b>0.32</b>
3488324	0.76	0.75	<b>3512644</b>	<b>0.30</b>	<b>0.37</b>
3488272	0.56	0.56	<b>3595445</b>	<b>0.39</b>	<b>0.47</b>
<b>3511258</b>	<b>0.81</b>	<b>0.84</b>	3488328	0.72	0.71

Note. Bold-faced number indicates a BCR or ECR item.





**Table 1.50 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 5 Form A**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2007	3511531	BCR_A	31,083	0.68	0.47	31.32	67.50	N/A	N/A	1.18
2007	3595438	BCR_B	31,083	1.10	0.65	15.05	56.26	26.61	N/A	2.08
2006	3512615	BCR_A	2,909	0.79	0.41	20.45	78.48	N/A	N/A	1.07
2006	3595439	BCR_B	2,909	0.94	0.45	22.28	58.58	17.81	N/A	1.34
2007	3511336	BCR_A	31,083	0.33	0.47	62.30	32.99	N/A	N/A	4.71
2007	3595440	BCR_B	31,083	0.67	0.70	40.87	39.81	13.67	N/A	5.66
2007	3511258	ECR_A	31,083	0.81	0.39	16.62	81.19	N/A	N/A	2.18
2007	3595441	ECR_B	31,083	1.48	0.69	2.53	51.80	37.02	7.44	1.22
2007	3556476	BCR_A	31,083	0.49	0.50	48.34	49.38	N/A	N/A	2.28
2007	3595442	BCR_B	31,083	0.92	0.92	43.72	14.72	38.42	N/A	3.15
2007	3512618	BCR_A	31,083	0.45	0.50	52.54	44.54	N/A	N/A	2.92
2007	3595443	BCR_B	31,083	1.05	0.52	7.46	72.60	15.99	N/A	3.95
2006	3512564	BCR_A	25,372	0.31	0.46	63.25	31.04	N/A	N/A	5.71
2006	3595444	BCR_B	25,372	0.44	0.38	51.10	39.71	2.34	N/A	6.85
2006	3512644	BCR_A	2,909	0.31	0.46	64.46	30.46	N/A	N/A	5.09
2006	3595445	BCR_B	2,909	0.77	0.54	37.85	36.51	20.25	N/A	5.40
2008	3511531	BCR_A	30,537	0.69	0.46	30.23	68.94	N/A	N/A	0.84
2008	3595438	BCR_B	30,537	1.18	0.68	13.87	51.27	33.19	N/A	1.68
2008	3512615	BCR_A	30,537	0.79	0.41	19.92	78.87	N/A	N/A	1.21
2008	3595439	BCR_B	30,537	1.10	0.68	16.68	53.00	28.44	N/A	1.88
2008	3511336	BCR_A	30,537	0.43	0.49	52.23	42.61	N/A	N/A	5.16
2008	3595440	BCR_B	30,537	0.75	0.75	36.69	37.97	18.73	N/A	6.61
2008	3511258	ECR_A	30,537	0.84	0.37	14.75	83.89	N/A	N/A	1.37
2008	3595441	ECR_B	30,537	1.53	0.64	2.35	45.36	46.42	4.87	1.00
2008	3556476	BCR_A	30,537	0.50	0.50	48.14	49.87	N/A	N/A	1.98
2008	3595442	BCR_B	30,537	0.88	0.91	45.22	15.45	36.46	N/A	2.86
2008	3512618	BCR_A	30,537	0.46	0.50	51.62	45.78	N/A	N/A	2.60
2008	3595443	BCR_B	30,537	1.10	0.54	6.93	69.45	20.45	N/A	3.17
2008	3512564	BCR_A	30,537	0.36	0.48	60.11	36.30	N/A	N/A	3.59
2008	3595444	BCR_B	30,537	0.64	0.60	39.26	51.16	6.54	N/A	3.04
2008	3512644	BCR_A	30,537	0.37	0.48	58.24	36.89	N/A	N/A	4.88
2008	3595445	BCR_B	30,537	0.93	0.76	26.55	42.24	25.38	N/A	5.83

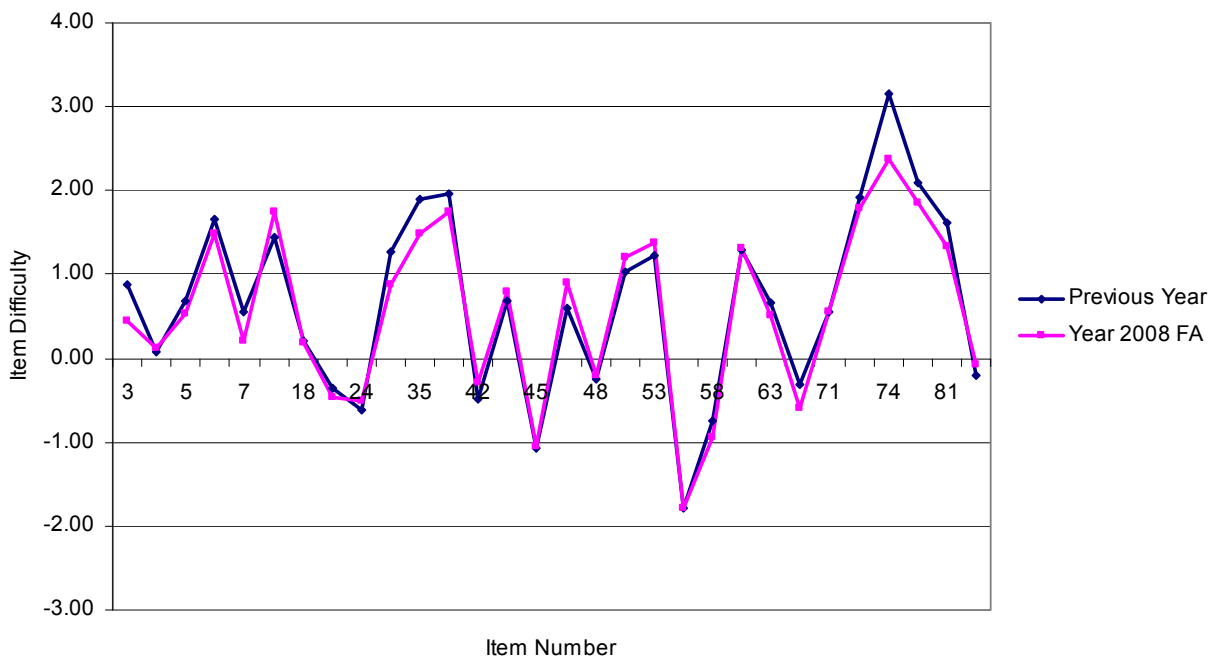
**Table 1.51 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008:  
Grade 5 Form A**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2006	3	3512642	SR	0.8800			
2007	4	3511531	BCR_A	0.0868			
2007	5	3595438	BCR_B	0.6862	-1.6106	1.6106	
2007	6	3488390	SR	1.6511			
2006	7	3512622	SR	0.5443			
2007	9	3488506	SR	1.4501			
2007	18	3488373	SR	0.2052			
2006	22	3512639	SR	-0.3576			
2006	24	3512615	BCR_A	-0.6075			
2006	25	3595439	BCR_B	1.2646	-1.6156	1.6156	
2007	35	3511336	BCR_A	1.8944			
2007	36	3595440	BCR_B	1.9609	-1.0144	1.0144	
2007	42	3488324	SR	-0.4943			
2007	44	3488272	SR	0.6757			
2007	45	3511258	ECR_A	-1.0768			
2007	46	3595441	ECR_B	0.6008	-3.6557	0.5929	3.0628
2007	48	3488431	SR	-0.2493			
2007	52	3556476	BCR_A	1.0216			
2007	53	3595442	BCR_B	1.2214	0.5363	-0.5363	
2007	57	3488241	SR	-1.7907			
2005	58	100000043857	SR	-0.7469			
2007	62	3512618	BCR_A	1.2891			
2007	63	3595443	BCR_B	0.6654	-2.4487	2.4487	
2006	64	3512623	SR	-0.3171			
2007	71	3488251	SR	0.5631			
2006	73	3512564	BCR_A	1.9152			
2006	74	3595444	BCR_B	3.1486	-1.8742	1.8742	
2006	80	3512644	BCR_A	2.0909			
2006	81	3595445	BCR_B	1.6112	-0.7732	0.7732	
2007	82	3488328	SR	-0.2008			
2008	3	3512642	SR	0.4442			
2008	4	3511531	BCR_A	0.1259			
2008	5	3595438	BCR_B	0.5335	-1.3908	1.3908	
2008	6	3488390	SR	1.4804			
2008	7	3512622	SR	0.2133			
2008	9	3488506	SR	1.7536			
2008	18	3488373	SR	0.1790			
2008	22	3512639	SR	-0.4690			
2008	24	3512615	BCR_A	-0.5151			
2008	25	3595439	BCR_B	0.8697	-1.4537	1.4537	
2008	35	3511336	BCR_A	1.4848			
2008	36	3595440	BCR_B	1.7477	-0.8139	0.8139	
2008	42	3488324	SR	-0.2851			
2008	44	3488272	SR	0.8010			

**Table 1.51 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	45	3511258	ECR_A	-1.0546			
2008	46	3595441	ECR_B	0.9016	-3.8291	0.2550	3.5741
2008	48	3488431	SR	-0.2130			
2008	52	3556476	BCR_A	1.2085			
2008	53	3595442	BCR_B	1.3852	0.6408	-0.6408	
2008	57	3488241	SR	-1.7928			
2008	58	100000043857	SR	-0.9439			
2008	62	3512618	BCR_A	1.3098			
2008	63	3595443	BCR_B	0.5173	-2.2969	2.2969	
2008	64	3512623	SR	-0.5862			
2008	71	3488251	SR	0.5581			
2008	73	3512564	BCR_A	1.7934			
2008	74	3595444	BCR_B	2.3824	-1.8916	1.8916	
2008	80	3512644	BCR_A	1.8531			
2008	81	3595445	BCR_B	1.3274	-0.9532	0.9532	
2008	82	3488328	SR	-0.0717			

Note. Rasch item and step difficulties are on a common scale.

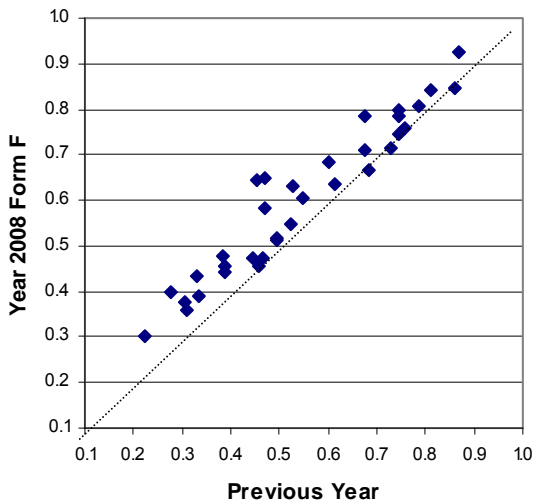


**Figure 1.5 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 5 Form A**

**Table 1.52 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 5 Form F**

Item CID	Previous Year	Year 08 Form F	Item CID	Previous Year	Year 08 Form F
3512642	0.53	0.63	3488324	0.76	0.76
<b>3511531</b>	<b>0.68</b>	<b>0.71</b>	<b>3511258</b>	<b>0.81</b>	<b>0.84</b>
<b>3595438</b>	<b>0.55</b>	<b>0.60</b>	<b>3595441</b>	<b>0.49</b>	<b>0.52</b>
3488390	0.39	0.44	3488431	0.75	0.75
3512622	0.60	0.69	<b>3556476</b>	<b>0.49</b>	<b>0.51</b>
3488356	0.75	0.78	<b>3595442</b>	<b>0.46</b>	<b>0.46</b>
3488373	0.68	0.67	3488418	0.39	0.46
100000043850	0.45	0.64	3488372	0.86	0.85
3512639	0.75	0.80	<b>3512618</b>	<b>0.45</b>	<b>0.47</b>
<b>3512615</b>	<b>0.78</b>	<b>0.81</b>	<b>3595443</b>	<b>0.52</b>	<b>0.55</b>
<b>3595439</b>	<b>0.47</b>	<b>0.58</b>	3488455	0.87	0.92
100000043855	0.28	0.40	3488299	0.61	0.63
3488377	0.73	0.71	3488457	0.47	0.48
3511542	0.47	0.65	<b>3512564</b>	<b>0.31</b>	<b>0.36</b>
<b>3511336</b>	<b>0.33</b>	<b>0.43</b>	<b>3595444</b>	<b>0.22</b>	<b>0.30</b>
<b>3595440</b>	<b>0.34</b>	<b>0.39</b>	<b>3512644</b>	<b>0.30</b>	<b>0.38</b>
3492137	0.68	0.79	<b>3595445</b>	<b>0.39</b>	<b>0.48</b>

Note. Bold-faced number indicates a BCR or ECR item.



**Table 1.53 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 5 Form F**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2007	3511531	BCR_A	31,083	0.68	0.47	31.32	67.50	N/A	N/A	1.18
2007	3595438	BCR_B	31,083	1.10	0.65	15.05	56.26	26.61	N/A	2.08
2006	3512615	BCR_A	2,909	0.79	0.41	20.45	78.48	N/A	N/A	1.07
2006	3595439	BCR_B	2,909	0.94	0.45	22.28	58.58	17.81	N/A	1.34
2007	3511336	BCR_A	31,083	0.33	0.47	62.30	32.99	N/A	N/A	4.71
2007	3595440	BCR_B	31,083	0.67	0.70	40.87	39.81	13.67	N/A	5.66
2007	3511258	ECR_A	31,083	0.81	0.39	16.62	81.19	N/A	N/A	2.18
2007	3595441	ECR_B	31,083	1.48	0.69	2.53	51.80	37.02	7.44	1.22
2007	3556476	BCR_A	31,083	0.49	0.50	48.34	49.38	N/A	N/A	2.28
2007	3595442	BCR_B	31,083	0.92	0.92	43.72	14.72	38.42	N/A	3.15
2007	3512618	BCR_A	31,083	0.45	0.50	52.54	44.54	N/A	N/A	2.92
2007	3595443	BCR_B	31,083	1.05	0.52	7.46	72.60	15.99	N/A	3.95
2006	3512564	BCR_A	25,372	0.31	0.46	63.25	31.04	N/A	N/A	5.71
2006	3595444	BCR_B	25,372	0.44	0.38	51.10	39.71	2.34	N/A	6.85
2006	3512644	BCR_A	2,909	0.31	0.46	64.46	30.46	N/A	N/A	5.09
2006	3595445	BCR_B	2,909	0.77	0.54	37.85	36.51	20.25	N/A	5.40
2008	3511531	BCR_A	30,289	0.71	0.45	28.53	70.81	N/A	N/A	0.66
2008	3595438	BCR_B	30,289	1.21	0.64	10.64	54.67	33.14	N/A	1.55
2008	3512615	BCR_A	30,289	0.81	0.39	18.37	80.80	N/A	N/A	0.82
2008	3595439	BCR_B	30,289	1.17	0.68	14.73	51.11	32.82	N/A	1.33
2008	3511336	BCR_A	30,289	0.43	0.50	52.72	43.34	N/A	N/A	3.94
2008	3595440	BCR_B	30,289	0.78	0.75	36.80	38.18	19.88	N/A	5.15
2008	3511258	ECR_A	30,289	0.84	0.37	14.61	84.11	N/A	N/A	1.28
2008	3595441	ECR_B	30,289	1.55	0.63	2.00	44.03	48.35	4.79	0.82
2008	3556476	BCR_A	30,289	0.51	0.50	46.75	51.39	N/A	N/A	1.86
2008	3595442	BCR_B	30,289	0.91	0.91	44.12	15.65	37.73	N/A	2.49
2008	3512618	BCR_A	30,289	0.47	0.50	50.34	47.19	N/A	N/A	2.47
2008	3595443	BCR_B	30,289	1.10	0.51	5.27	73.28	18.29	N/A	3.16
2008	3512564	BCR_A	30,289	0.36	0.48	61.06	35.98	N/A	N/A	2.97
2008	3595444	BCR_B	30,289	0.61	0.60	42.73	48.52	6.13	N/A	2.62
2008	3512644	BCR_A	30,289	0.38	0.49	58.06	37.86	N/A	N/A	4.08
2008	3595445	BCR_B	30,289	0.96	0.75	25.48	43.48	26.09	N/A	4.95

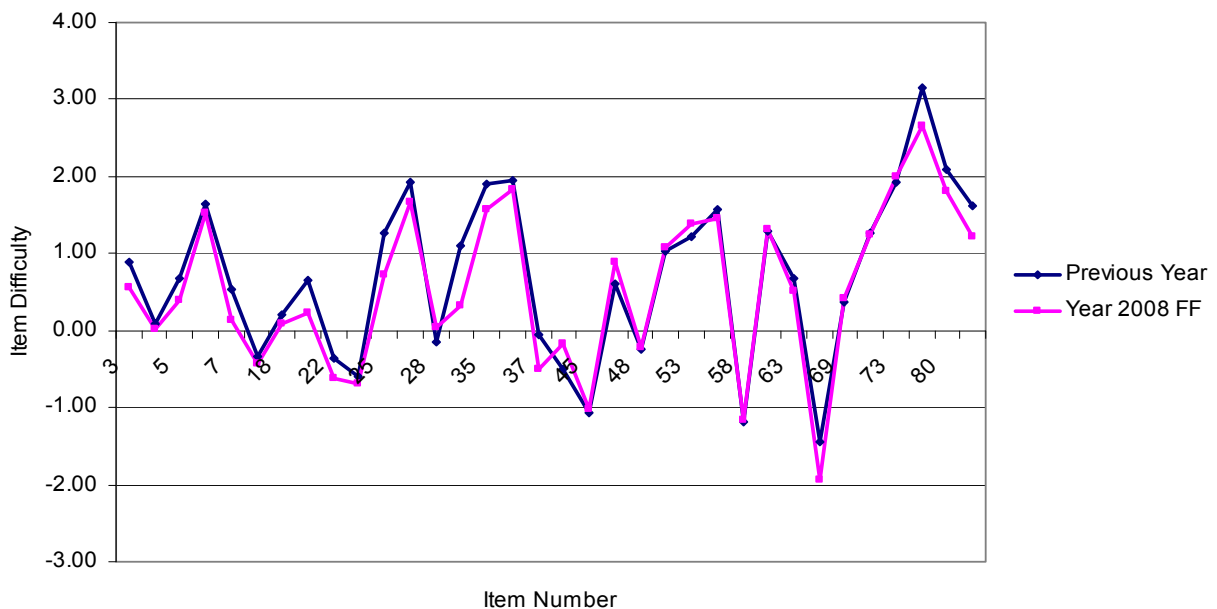
**Table 1.54 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008:  
Grade 5 Form F**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2006	3	3512642	SR	0.8800			
2007	4	3511531	BCR_A	0.0868			
2007	5	3595438	BCR_B	0.6862	-1.6106	1.6106	
2007	6	3488390	SR	1.6511			
2006	7	3512622	SR	0.5443			
2007	9	3488356	SR	-0.3278			
2007	18	3488373	SR	0.2052			
2004	20	100000043850	SR	0.6431			
2006	22	3512639	SR	-0.3576			
2006	24	3512615	BCR_A	-0.6075			
2006	25	3595439	BCR_B	1.2646	-1.6156	1.6156	
2005	27	100000043855	SR	1.9156			
2007	28	3488377	SR	-0.1395			
2006	34	3511542	SR	1.1084			
2007	35	3511336	BCR_A	1.8944			
2007	36	3595440	BCR_B	1.9609	-1.0144	1.0144	
2007	37	3492137	SR	-0.0612			
2007	42	3488324	SR	-0.4943			
2007	45	3511258	ECR_A	-1.0768			
2007	46	3595441	ECR_B	0.6008	-3.6557	0.5929	3.0628
2007	48	3488431	SR	-0.2493			
2007	52	3556476	BCR_A	1.0216			
2007	53	3595442	BCR_B	1.2214	0.5363	-0.5363	
2007	57	3488418	SR	1.5806			
2007	58	3488372	SR	-1.1803			
2007	62	3512618	BCR_A	1.2891			
2007	63	3595443	BCR_B	0.6654	-2.4487	2.4487	
2007	64	3488455	SR	-1.4423			
2007	69	3488299	SR	0.3609			
2007	70	3488457	SR	1.2553			
2006	73	3512564	BCR_A	1.9152			
2006	74	3595444	BCR_B	3.1486	-1.8742	1.8742	
2006	80	3512644	BCR_A	2.0909			
2006	81	3595445	BCR_B	1.6112	-0.7732	0.7732	
2008	3	3512642	SR	0.5603			
2008	4	3511531	BCR_A	0.0205			
2008	5	3595438	BCR_B	0.3914	-1.6521	1.6521	
2008	6	3488390	SR	1.5155			
2008	7	3512622	SR	0.1285			
2008	9	3488356	SR	-0.4351			
2008	18	3488373	SR	0.0790			
2008	20	100000043850	SR	0.2307			
2008	22	3512639	SR	-0.6094			
2008	24	3512615	BCR_A	-0.6929			

**Table 1.54 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	25	3595439	BCR_B	0.7258	-1.4307	1.4307	
2008	27	100000043855	SR	1.6760			
2008	28	3488377	SR	0.0327			
2008	34	3511542	SR	0.3252			
2008	35	3511336	BCR_A	1.5666			
2008	36	3595440	BCR_B	1.8347	-0.9254	0.9254	
2008	37	3492137	SR	-0.5120			
2008	42	3488324	SR	-0.1710			
2008	45	3511258	ECR_A	-1.0108			
2008	46	3595441	ECR_B	0.8865	-4.0748	0.2841	3.7906
2008	48	3488431	SR	-0.2189			
2008	52	3556476	BCR_A	1.0842			
2008	53	3595442	BCR_B	1.3952	0.5900	-0.5900	
2008	57	3488418	SR	1.4472			
2008	58	3488372	SR	-1.1720			
2008	62	3512618	BCR_A	1.3102			
2008	63	3595443	BCR_B	0.5023	-2.7505	2.7505	
2008	64	3488455	SR	-1.9347			
2008	69	3488299	SR	0.4209			
2008	70	3488457	SR	1.2508			
2008	73	3512564	BCR_A	2.0012			
2008	74	3595444	BCR_B	2.6476	-1.6217	1.6217	
2008	80	3512644	BCR_A	1.8003			
2008	81	3595445	BCR_B	1.2216	-0.9532	0.9532	

Note. Rasch item and step difficulties are on a common scale.

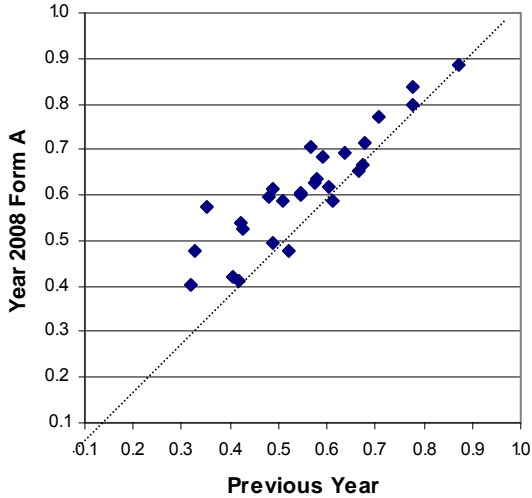


**Figure 1.6 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 5 Form F**

**Table 1.55 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 6 Form A**

Item CID	Previous Year	Year 08 Form A	Item CID	Previous Year	Year 08 Form A
3488264	0.59	0.68	3492095	0.78	0.80
3492143	0.71	0.77	<b>3516333</b>	<b>0.60</b>	<b>0.62</b>
<b>3517004</b>	<b>0.87</b>	<b>0.89</b>	<b>3595449</b>	<b>0.61</b>	<b>0.59</b>
<b>3595446</b>	<b>0.58</b>	<b>0.63</b>	3516929	0.66	0.65
3516909	0.51	0.59	3516906	0.55	0.60
<b>3516627</b>	<b>0.52</b>	<b>0.48</b>	3492142	0.58	0.63
<b>3595447</b>	<b>0.42</b>	<b>0.41</b>	<b>3517013</b>	<b>0.35</b>	<b>0.57</b>
3488482	0.78	0.84	<b>3595450</b>	<b>0.57</b>	<b>0.71</b>
100000043862	0.49	0.61	3516375	0.55	0.61
3488383	0.67	0.67	<b>3516616</b>	<b>0.40</b>	<b>0.42</b>
3488516	0.64	0.69	<b>3595451</b>	<b>0.49</b>	<b>0.50</b>
100000043865	0.43	0.53	3488508	0.68	0.71
<b>3516363</b>	<b>0.33</b>	<b>0.48</b>	<b>3516913</b>	<b>0.32</b>	<b>0.40</b>
<b>3595448</b>	<b>0.48</b>	<b>0.60</b>	<b>3595452</b>	<b>0.42</b>	<b>0.54</b>

Note. Bold-faced number indicates a BCR or ECR item.





**Table 1.56 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 6 Form A**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2007	3517004	ECR_A	31,258	0.87	0.34	11.36	87.09	N/A	N/A	1.55
2007	3595446	ECR_B	31,258	1.74	0.96	8.71	29.21	34.34	25.26	2.49
2007	3516627	BCR_A	31,558	0.52	0.50	41.66	52.17	N/A	N/A	6.17
2007	3595447	BCR_B	31,558	0.83	0.61	22.28	59.64	11.87	N/A	6.21
2006	3516363	BCR_A	3,289	0.33	0.47	62.51	32.75	N/A	N/A	4.74
2006	3595448	BCR_B	3,289	0.96	0.56	28.79	36.58	29.61	N/A	5.02
2007	3516333	BCR_A	31,558	0.60	0.49	37.60	60.47	N/A	N/A	1.93
2007	3595449	BCR_B	31,558	1.22	0.81	21.05	29.56	46.28	N/A	3.11
2007	3517013	BCR_A	31,558	0.35	0.48	61.55	35.38	N/A	N/A	3.07
2007	3595450	BCR_B	31,558	1.13	0.63	10.29	59.17	27.04	N/A	3.50
2007	3516616	BCR_A	31,258	0.40	0.49	55.22	40.43	N/A	N/A	4.34
2007	3595451	BCR_B	31,258	0.98	0.60	14.55	63.53	17.12	N/A	4.80
2006	3516913	BCR_A	3,242	0.32	0.47	60.33	31.89	N/A	N/A	7.77
2006	3595452	BCR_B	3,242	0.85	0.51	26.56	45.96	19.25	N/A	8.24
2008	3517004	ECR_A	31,060	0.89	0.32	10.64	88.75	N/A	N/A	0.61
2008	3595446	ECR_B	31,060	1.90	0.94	6.39	25.68	34.10	32.13	1.70
2008	3516627	BCR_A	31,060	0.48	0.50	45.52	47.79	N/A	N/A	6.69
2008	3595447	BCR_B	31,060	0.82	0.67	24.69	52.36	14.89	N/A	8.06
2008	3516363	BCR_A	31,060	0.48	0.50	49.59	47.80	N/A	N/A	2.61
2008	3595448	BCR_B	31,060	1.19	0.79	19.70	34.36	42.49	N/A	3.45
2008	3516333	BCR_A	31,060	0.62	0.49	36.21	61.65	N/A	N/A	2.14
2008	3595449	BCR_B	31,060	1.18	0.77	18.93	37.54	40.05	N/A	3.48
2008	3517013	BCR_A	31,060	0.57	0.49	39.91	57.45	N/A	N/A	2.65
2008	3595450	BCR_B	31,060	1.41	0.64	5.19	42.07	49.52	N/A	3.22
2008	3516616	BCR_A	31,060	0.42	0.49	51.83	42.07	N/A	N/A	6.11
2008	3595451	BCR_B	31,060	0.99	0.62	12.77	61.83	18.73	N/A	6.67
2008	3516913	BCR_A	31,060	0.40	0.49	57.57	40.49	N/A	N/A	1.94
2008	3595452	BCR_B	31,060	1.08	0.66	15.60	55.54	26.28	N/A	2.59

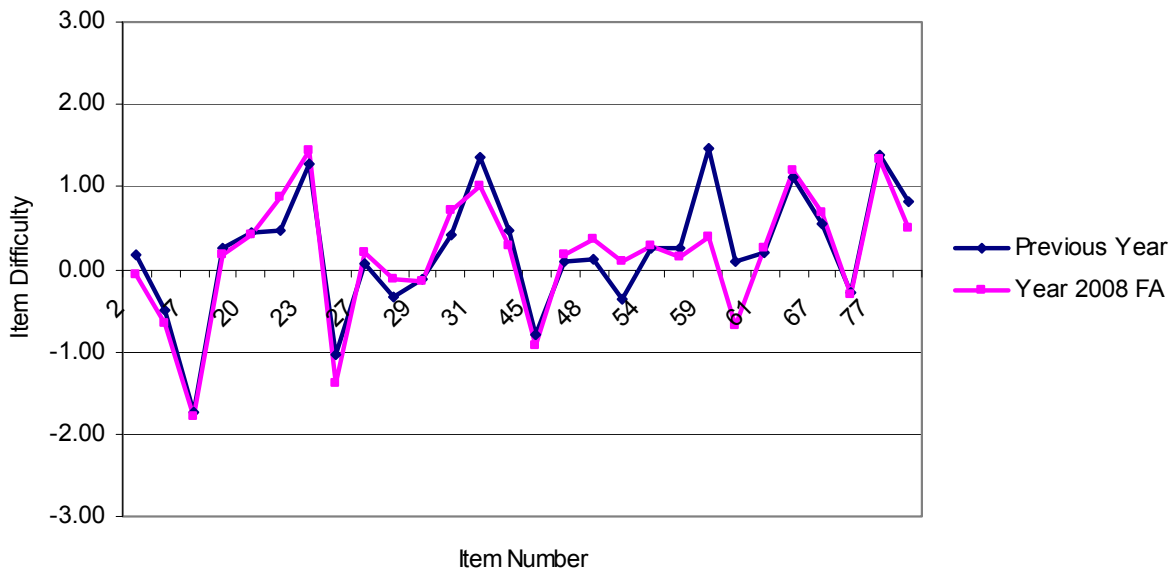
**Table 1.57 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008:  
Grade 6 Form A**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2007	2	3488264	SR	0.1759			
2007	4	3492143	SR	-0.4941			
2007	7	3517004	ECR_A	-1.7238			
2007	8	3595446	ECR_B	0.2493	-1.6097	0.1701	1.4396
2006	20	3516909	SR	0.4392			
2007	22	3516627	BCR_A	0.4728			
2007	23	3595447	BCR_B	1.265	-1.8927	1.8927	
2007	24	3488482	SR	-1.0259			
2004	27	100000043862	SR	0.0659			
2007	28	3488383	SR	-0.3426			
2007	29	3488516	SR	-0.1093			
2004	30	100000043865	SR	0.4276			
2006	31	3516363	BCR_A	1.3465			
2006	32	3595448	BCR_B	0.4811	-0.7058	0.7058	
2007	45	3492095	SR	-0.8005			
2007	47	3516333	BCR_A	0.1031			
2007	48	3595449	BCR_B	0.1124	-0.4274	0.4274	
2006	50	3516929	SR	-0.3587			
2006	54	3516906	SR	0.2547			
2007	58	3492142	SR	0.2448			
2007	59	3517013	BCR_A	1.4674			
2007	60	3595450	BCR_B	0.0865	-1.7954	1.7954	
2006	61	3516375	SR	0.1983			
2007	66	3516616	BCR_A	1.1174			
2007	67	3595451	BCR_B	0.5414	-1.8777	1.8777	
2007	71	3488508	SR	-0.2951			
2006	77	3516913	BCR_A	1.3788			
2006	78	3595452	BCR_B	0.8083	-1.1885	1.1885	
2008	2	3488264	SR	-0.0691			
2008	4	3492143	SR	-0.6658			
2008	7	3517004	ECR_A	-1.7891			
2008	8	3595446	ECR_B	0.1623	-1.8832	0.1382	1.7450
2008	20	3516909	SR	0.4042			
2008	22	3516627	BCR_A	0.8724			
2008	23	3595447	BCR_B	1.4503	-1.6293	1.6293	
2008	24	3488482	SR	-1.3990			
2008	27	100000043862	SR	0.2101			
2008	28	3488383	SR	-0.1255			

**Table 1.57 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	29	3488516	SR	-0.1462			
2008	30	100000043865	SR	0.7262			
2008	31	3516363	BCR_A	0.9993			
2008	32	3595448	BCR_B	0.2927	-0.7056	0.7056	
2008	45	3492095	SR	-0.9261			
2008	47	3516333	BCR_A	0.1802			
2008	48	3595449	BCR_B	0.3670	-0.8656	0.8656	
2008	50	3516929	SR	0.0810			
2008	54	3516906	SR	0.2864			
2008	58	3492142	SR	0.1475			
2008	59	3517013	BCR_A	0.4029			
2008	60	3595450	BCR_B	-0.6751	-1.4158	1.4158	
2008	61	3516375	SR	0.2607			
2008	66	3516616	BCR_A	1.2017			
2008	67	3595451	BCR_B	0.6973	-2.0205	2.0205	
2008	71	3488508	SR	-0.3096			
2008	77	3516913	BCR_A	1.3284			
2008	78	3595452	BCR_B	0.4845	-1.4754	1.4754	

Note. Rasch item and step difficulties are on a common scale.

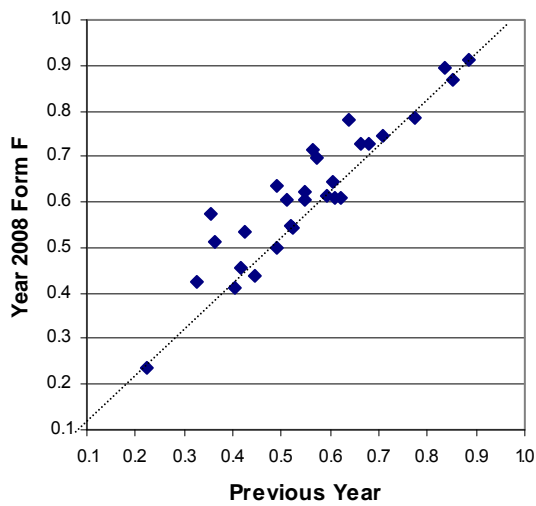


**Figure 1.7 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 6 Form A**

**Table 1.58 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 6 Form F**

Item CID	Previous Year	Year 08 Form F	Item CID	Previous Year	Year 08 Form F
3488502	0.89	0.91	<b>3516333</b>	<b>0.60</b>	<b>0.64</b>
<b>3516923</b>	<b>0.64</b>	<b>0.78</b>	<b>3595449</b>	<b>0.61</b>	<b>0.61</b>
<b>3595453</b>	<b>0.36</b>	<b>0.51</b>	3516929	0.66	0.73
3516361	0.57	0.70	3516906	0.55	0.61
3492088	0.85	0.87	3488256	0.62	0.61
3516909	0.51	0.61	<b>3517013</b>	<b>0.35</b>	<b>0.57</b>
<b>3516627</b>	<b>0.52</b>	<b>0.54</b>	<b>3595450</b>	<b>0.57</b>	<b>0.71</b>
<b>3595447</b>	<b>0.42</b>	<b>0.46</b>	3516375	0.55	0.62
3488441	0.52	0.55	<b>3516616</b>	<b>0.40</b>	<b>0.41</b>
100000043862	0.49	0.64	<b>3595451</b>	<b>0.49</b>	<b>0.50</b>
3488263	0.77	0.79	3488508	0.68	0.73
3488500	0.84	0.89	<b>3516327</b>	<b>0.44</b>	<b>0.44</b>
100000043865	0.43	0.53	<b>3595455</b>	<b>0.59</b>	<b>0.61</b>
<b>3516628</b>	<b>0.22</b>	<b>0.24</b>	3488257	0.71	0.74
<b>3595454</b>	<b>0.32</b>	<b>0.42</b>			

Note. Bold-faced number indicates a BCR or ECR item.



**Table 1.59 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 6 Form F**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2006	3516923	ECR_A	3,222	0.64	0.48	32.25	63.84	N/A	N/A	3.91
2006	3595453	ECR_B	3,222	1.09	0.53	26.23	36.84	24.58	7.70	4.66
2007	3516627	BCR_A	31,558	0.52	0.50	41.66	52.17	N/A	N/A	6.17
2007	3595447	BCR_B	31,558	0.83	0.61	22.28	59.64	11.87	N/A	6.21
2006	3516628	BCR_A	3,262	0.22	0.42	74.77	22.23	N/A	N/A	3.00
2006	3595454	BCR_B	3,262	0.65	0.54	49.57	29.28	17.78	N/A	3.37
2007	3516333	BCR_A	31,558	0.60	0.49	37.60	60.47	N/A	N/A	1.93
2007	3595449	BCR_B	31,558	1.22	0.81	21.05	29.56	46.28	N/A	3.11
2007	3517013	BCR_A	31,558	0.35	0.48	61.55	35.38	N/A	N/A	3.07
2007	3595450	BCR_B	31,558	1.13	0.63	10.29	59.17	27.04	N/A	3.50
2007	3516616	BCR_A	31,258	0.40	0.49	55.22	40.43	N/A	N/A	4.34
2007	3595451	BCR_B	31,258	0.98	0.60	14.55	63.53	17.12	N/A	4.80
2007	3516327	BCR_A	31,558	0.44	0.50	52.43	44.43	N/A	N/A	3.14
2007	3595455	BCR_B	31,558	1.19	0.75	16.87	39.74	39.45	N/A	3.94
2008	3516923	ECR_A	30,292	0.78	0.41	20.62	78.05	N/A	N/A	1.33
2008	3595453	ECR_B	30,292	1.54	0.97	12.95	34.55	30.65	19.50	2.35
2008	3516627	BCR_A	30,292	0.54	0.50	39.91	54.38	N/A	N/A	5.71
2008	3595447	BCR_B	30,292	0.92	0.66	19.43	55.77	17.86	N/A	6.94
2008	3516628	BCR_A	30,292	0.24	0.42	75.14	23.53	N/A	N/A	1.33
2008	3595454	BCR_B	30,292	0.85	0.74	34.13	43.11	20.86	N/A	1.90
2008	3516333	BCR_A	30,292	0.64	0.48	33.86	64.22	N/A	N/A	1.92
2008	3595449	BCR_B	30,292	1.22	0.77	17.67	36.74	42.52	N/A	3.08
2008	3517013	BCR_A	30,292	0.57	0.49	40.26	57.29	N/A	N/A	2.45
2008	3595450	BCR_B	30,292	1.43	0.62	4.20	42.79	50.03	N/A	2.98
2008	3516616	BCR_A	30,292	0.41	0.49	52.82	41.03	N/A	N/A	6.15
2008	3595451	BCR_B	30,292	1.00	0.61	11.55	63.02	18.55	N/A	6.88
2008	3516327	BCR_A	30,292	0.44	0.50	54.40	43.67	N/A	N/A	1.92
2008	3595455	BCR_B	30,292	1.23	0.74	16.34	39.59	41.47	N/A	2.60

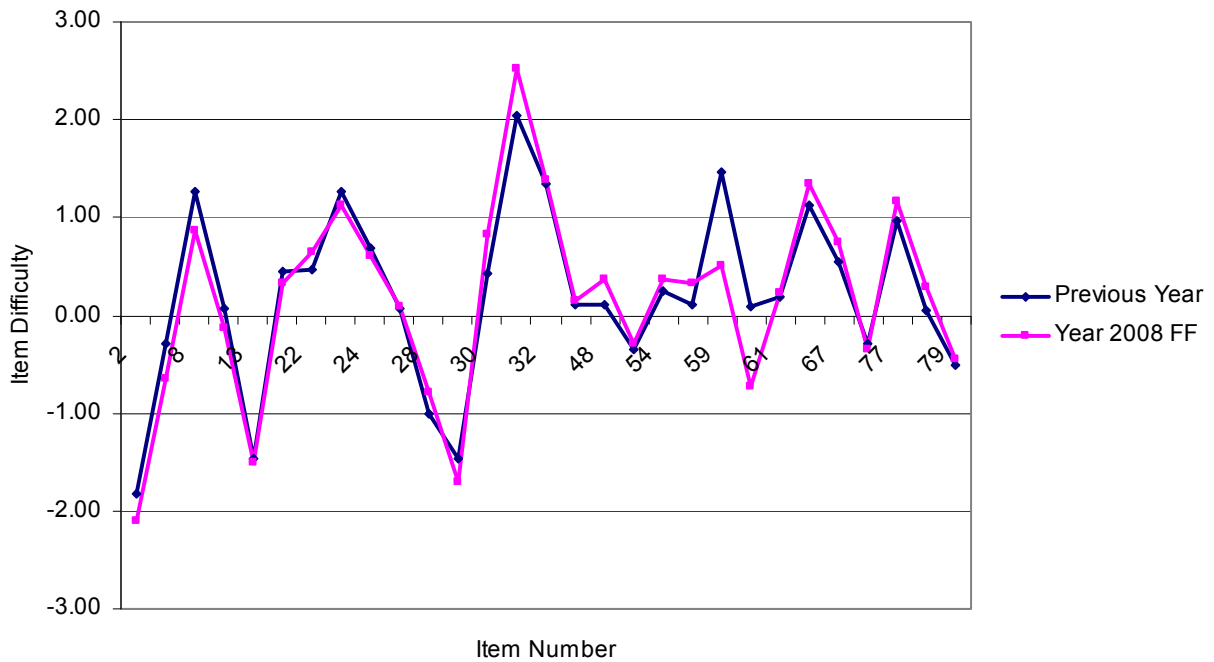
**Table 1.60 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008:  
Grade 6 Form F**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2007	2	3488502	SR	-1.8311			
2006	7	3516923	ECR_A	-0.2947			
2006	8	3595453	ECR_B	1.2607	-1.5377	0.0087	1.5290
2006	12	3516361	SR	0.0635			
2007	13	3492088	SR	-1.4568			
2006	20	3516909	SR	0.4392			
2007	22	3516627	BCR_A	0.4728			
2007	23	3595447	BCR_B	1.2650	-1.8927	1.8927	
2007	24	3488441	SR	0.6901			
2004	27	100000043862	SR	0.0659			
2007	28	3488263	SR	-1.0090			
2007	29	3488500	SR	-1.4601			
2004	30	100000043865	SR	0.4276			
2006	31	3516628	BCR_A	2.0518			
2006	32	3595454	BCR_B	1.3363	-0.4520	0.4520	
2007	47	3516333	BCR_A	0.1031			
2007	48	3595449	BCR_B	0.1124	-0.4274	0.4274	
2006	50	3516929	SR	-0.3587			
2006	54	3516906	SR	0.2547			
2007	58	3488256	SR	0.1076			
2007	59	3517013	BCR_A	1.4674			
2007	60	3595450	BCR_B	0.0865	-1.7954	1.7954	
2006	61	3516375	SR	0.1983			
2007	66	3516616	BCR_A	1.1174			
2007	67	3595451	BCR_B	0.5414	-1.8777	1.8777	
2007	71	3488508	SR	-0.2951			
2007	77	3516327	BCR_A	0.9630			
2007	78	3595455	BCR_B	0.0487	-0.9977	0.9977	
2007	79	3488257	SR	-0.5082			
2008	2	3488502	SR	-2.1067			
2008	7	3516923	ECR_A	-0.6537			
2008	8	3595453	ECR_B	0.8626	-1.8458	0.2499	1.5959
2008	12	3516361	SR	-0.1306			
2008	13	3492088	SR	-1.4997			
2008	20	3516909	SR	0.3274			
2008	22	3516627	BCR_A	0.6439			
2008	23	3595447	BCR_B	1.1168	-1.7979	1.7979	
2008	24	3488441	SR	0.6093			
2008	27	100000043862	SR	0.0847			

**Table 1.60 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	28	3488263	SR	-0.7946			
2008	29	3488500	SR	-1.7078			
2008	30	100000043865	SR	0.8249			
2008	31	3516628	BCR_A	2.5157			
2008	32	3595454	BCR_B	1.3822	-1.0038	1.0038	
2008	47	3516333	BCR_A	0.1568			
2008	48	3595449	BCR_B	0.3681	-0.8854	0.8854	
2008	50	3516929	SR	-0.2890			
2008	54	3516906	SR	0.3667			
2008	58	3488256	SR	0.3246			
2008	59	3517013	BCR_A	0.5063			
2008	60	3595450	BCR_B	-0.7280	-1.6251	1.6251	
2008	61	3516375	SR	0.2299			
2008	66	3516616	BCR_A	1.3552			
2008	67	3595451	BCR_B	0.7537	-2.0793	2.0793	
2008	71	3488508	SR	-0.3522			
2008	77	3516327	BCR_A	1.1725			
2008	78	3595455	BCR_B	0.2939	-0.8570	0.8570	
2008	79	3488257	SR	-0.4530			

Note. Rasch item and step difficulties are on a common scale.

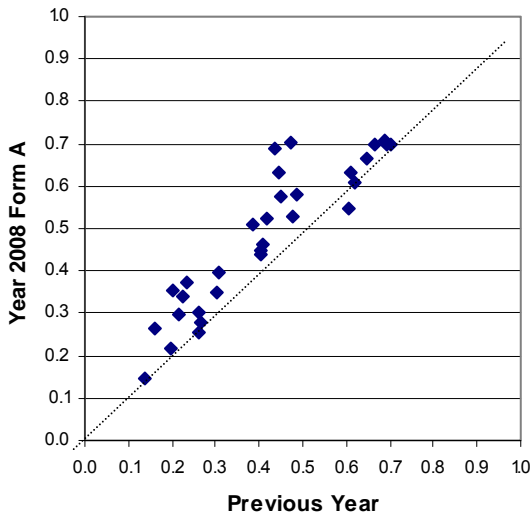


**Figure 1.8 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 6 Form F**

**Table 1.61 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 7 Form A**

Item CID	Previous Year	Year 08 Form A	Item CID	Previous Year	Year 08 Form A
<b>100000043334</b>	<b>0.41</b>	<b>0.46</b>	<b>100000043347</b>	<b>0.43</b>	<b>0.69</b>
<b>3595363</b>	<b>0.49</b>	<b>0.58</b>	<b>3595366</b>	<b>0.22</b>	<b>0.30</b>
3487667	0.26	0.25	<b>3517876</b>	<b>0.14</b>	<b>0.14</b>
3517863	0.61	0.63	<b>100000043353</b>	<b>0.44</b>	<b>0.63</b>
<b>100000043348</b>	<b>0.23</b>	<b>0.38</b>	100000043338	0.23	0.34
<b>3595364</b>	<b>0.16</b>	<b>0.26</b>	<b>3517673</b>	<b>0.65</b>	<b>0.67</b>
<b>100000043345</b>	<b>0.20</b>	<b>0.35</b>	<b>3564020</b>	<b>0.40</b>	<b>0.44</b>
<b>3547779</b>	<b>0.60</b>	<b>0.55</b>	<b>3487649</b>	<b>0.20</b>	<b>0.22</b>
<b>3517645</b>	<b>0.69</b>	<b>0.71</b>	<b>3517654</b>	<b>0.48</b>	<b>0.53</b>
<b>100000043351</b>	<b>0.47</b>	<b>0.70</b>	<b>100000043343</b>	<b>0.39</b>	<b>0.51</b>
<b>3517646</b>	<b>0.69</b>	<b>0.70</b>	<b>3517878</b>	<b>0.31</b>	<b>0.39</b>
<b>3595365</b>	<b>0.67</b>	<b>0.70</b>	<b>3595367</b>	<b>0.42</b>	<b>0.52</b>
<b>3547642</b>	<b>0.70</b>	<b>0.70</b>	3517691	0.62	0.61
<b>3487560</b>	<b>0.27</b>	<b>0.28</b>	<b>3492156</b>	<b>0.30</b>	<b>0.35</b>
<b>3517725</b>	<b>0.26</b>	<b>0.30</b>	3488830	0.45	0.58
<b>3564022</b>	<b>0.40</b>	<b>0.45</b>			

Note. Bold-faced number indicates a BCR or ECR item.





**Table 1.62 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 7 Form A**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2004	100000043334	BCR_A	11,522	0.41	0.49	45.83	40.74	N/A	N/A	13.43
2004	3595363	BCR_B	11,522	0.97	0.81	19.17	34.46	31.42	N/A	14.95
2004	100000043348	ECR_A	11,667	0.23	0.42	53.88	23.32	N/A	N/A	22.80
2004	3595364	ECR_B	11,667	0.48	0.83	40.76	18.01	7.90	4.75	28.58
2006	3517646	BCR_A	39,533	0.69	0.46	26.37	69.30	N/A	N/A	4.33
2006	3595365	BCR_B	39,533	1.33	0.56	15.67	25.53	53.93	N/A	4.87
2007	3517725	BCR_A	32,264	0.26	0.44	69.73	26.31	N/A	N/A	3.96
2007	3564022	BCR_B	32,264	0.81	0.80	38.54	31.61	24.57	N/A	5.28
2004	100000043347	ECR_A	11,522	0.43	0.50	47.12	43.40	N/A	N/A	9.49
2004	3595366	ECR_B	11,522	0.65	0.65	33.27	44.84	9.82	0.02	12.06
2007	3517673	ECR_A	32,264	0.65	0.48	31.06	64.74	N/A	N/A	4.20
2007	3564020	ECR_B	32,264	1.21	0.60	3.39	66.72	23.79	2.30	3.80
2006	3517878	BCR_A	3,382	0.31	0.46	61.00	30.81	N/A	N/A	8.19
2006	3595367	BCR_B	3,382	0.84	0.52	28.92	44.23	19.72	N/A	7.13
2008	100000043334	BCR_A	31,804	0.46	0.50	47.50	46.14	N/A	N/A	6.36
2008	3595363	BCR_B	31,804	1.16	0.82	19.66	30.65	42.56	N/A	7.12
2008	100000043348	ECR_A	31,804	0.38	0.48	55.55	37.50	N/A	N/A	6.95
2008	3595364	ECR_B	31,804	0.79	0.97	40.42	27.36	13.19	8.52	10.51
2008	3517646	BCR_A	31,804	0.70	0.46	20.75	69.96	N/A	N/A	9.29
2008	3595365	BCR_B	31,804	1.40	0.84	12.37	13.85	62.98	N/A	10.80
2008	3517725	BCR_A	31,804	0.30	0.46	66.31	30.25	N/A	N/A	3.44
2008	3564022	BCR_B	31,804	0.89	0.84	36.49	28.55	30.46	N/A	4.50
2008	100000043347	ECR_A	31,804	0.69	0.46	27.92	69.10	N/A	N/A	2.99
2008	3595366	ECR_B	31,804	0.89	0.55	15.77	69.82	9.23	0.35	4.82
2008	3517673	ECR_A	31,804	0.67	0.47	30.66	66.66	N/A	N/A	2.68
2008	3564020	ECR_B	31,804	1.32	0.64	2.53	62.97	27.49	4.78	2.22
2008	3517878	BCR_A	31,804	0.39	0.49	56.63	39.41	N/A	N/A	3.96
2008	3595367	BCR_B	31,804	1.04	0.72	18.74	47.86	28.28	N/A	5.12

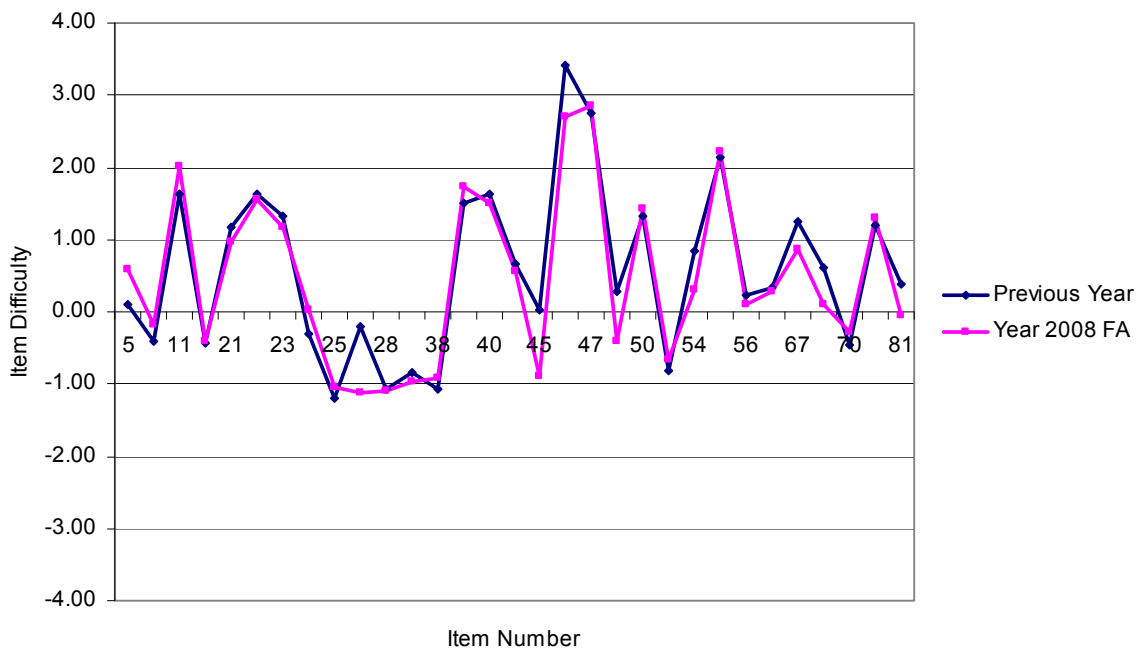
**Table 1.63 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008:  
Grade 7 Form A**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2004	5	100000043334	BCR_A	0.1146			
2004	6	3595363	BCR_B	-0.4087	-0.8011	0.8011	
2007	11	3487667	SR	1.6316			
2006	17	3517863	SR	-0.4238			
2004	21	100000043348	ECR_A	1.1693			
2004	22	3595364	ECR_B	1.6364	-0.8377	0.1101	0.7277
2004	23	100000043345	SPR	1.3162			
2007	24	3547779	SPR	-0.3069			
2007	25	3517645	SPR	-1.1974			
2004	26	100000043351	SPR	-0.2036			
2006	28	3517646	BCR_A	-1.0705			
2006	29	3595365	BCR_B	-0.8516	-0.3784	0.3784	
2007	38	3547642	SPR	-1.0604			
2007	39	3487560	SPR	1.5135			
2007	40	3517725	BCR_A	1.6208			
2007	41	3564022	BCR_B	0.6682	-0.6977	0.6977	
2004	45	100000043347	ECR_A	0.0200			
2004	46	3595366	ECR_B	3.4100	-4.0297	-1.0782	5.1079
2007	47	3517876	SPR	2.7529			
2005	48	100000043353	SPR	0.2680			
2004	50	100000043338	SR	1.3356			
2007	53	3517673	ECR_A	-0.8144			
2007	54	3564020	ECRB	0.8436	-4.4403	0.7733	3.667
2007	55	3487649	SPR	2.1304			
2006	56	3517654	SPR	0.2314			
2004	57	100000043343	SPR	0.3333			
2006	67	3517878	BCR_A	1.2445			
2006	68	3595367	BCR_B	0.6133	-1.1739	1.1739	
2006	70	3517691	SR	-0.4573			
2007	78	3492156	SPR	1.2034			
2007	81	3488830	SR	0.3784			
2008	5	100000043334	BCR_A	0.5824			
2008	6	3595363	BCR_B	-0.1895	-0.6819	0.6819	
2008	11	3487667	SR	2.0203			
2008	17	3517863	SR	-0.4038			
2008	21	100000043348	ECR_A	0.9563			
2008	22	3595364	ECR_B	1.5655	-1.0967	0.3318	0.7649
2008	23	100000043345	SPR	1.1777			
2008	24	3547779	SPR	0.0174			
2008	25	3517645	SPR	-1.0518			
2008	26	100000043351	SPR	-1.1112			

**Table 1.63 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	28	3517646	BCR_A	-1.0888			
2008	29	3595365	BCR_B	-0.9755	0.1389	-0.1389	
2008	38	3547642	SPR	-0.9166			
2008	39	3487560	SPR	1.7398			
2008	40	3517725	BCR_A	1.5045			
2008	41	3564022	BCR_B	0.5585	-0.4548	0.4548	
2008	45	100000043347	ECR_A	-0.8969			
2008	46	3595366	ECR_B	2.6970	-4.3825	0.7102	3.6723
2008	47	3517876	SPR	2.8645			
2008	48	100000043353	SPR	-0.4017			
2008	50	100000043338	SR	1.4392			
2008	53	3517673	ECR_A	-0.6673			
2008	54	3564020	ECR_B	0.2954	-4.5319	1.1166	3.4153
2008	55	3487649	SPR	2.2096			
2008	56	3517654	SPR	0.1125			
2008	57	100000043343	SPR	0.2698			
2008	67	3517878	BCR_A	0.8699			
2008	68	3595367	BCR_B	0.1013	-1.4432	1.4432	
2008	70	3517691	SR	-0.2784			
2008	78	3492156	SPR	1.2944			
2008	81	3488830	SR	-0.0385			

Note. Rasch item and step difficulties are on a common scale.

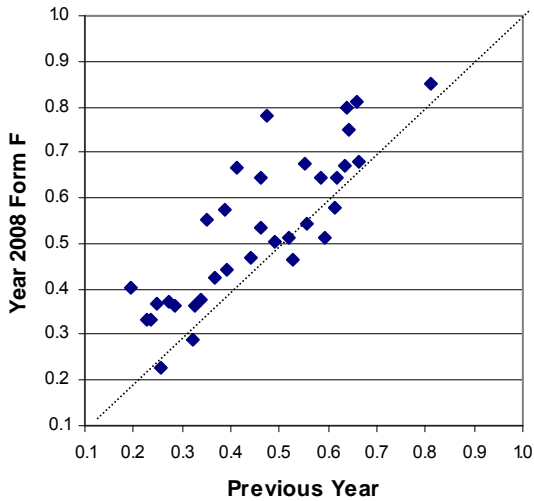


**Figure 1.9 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 7 Form A**

**Table 1.64 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 7 Form F**

Item CID	Previous Year	Year 08 Form F	Item CID	Previous Year	Year 08 Form F
<b>100000043335</b>	<b>0.55</b>	<b>0.68</b>	<b>3547487</b>	<b>0.66</b>	<b>0.81</b>
<b>3595368</b>	<b>0.64</b>	<b>0.80</b>	<b>3564031</b>	<b>0.23</b>	<b>0.33</b>
100000043349	0.29	0.36	<b>100000043354</b>	<b>0.19</b>	<b>0.40</b>
3517739	0.81	0.85	<b>100000043356</b>	<b>0.47</b>	<b>0.78</b>
<b>3487765</b>	<b>0.37</b>	<b>0.42</b>	100000043338	0.23	0.33
<b>3595369</b>	<b>0.49</b>	<b>0.50</b>	<b>3517648</b>	<b>0.63</b>	<b>0.67</b>
<b>100000043344</b>	<b>0.32</b>	<b>0.36</b>	<b>3564027</b>	<b>0.58</b>	<b>0.64</b>
<b>3513631</b>	<b>0.56</b>	<b>0.55</b>	<b>3492169</b>	<b>0.34</b>	<b>0.38</b>
<b>3487596</b>	<b>0.27</b>	<b>0.37</b>	<b>100000043342</b>	<b>0.41</b>	<b>0.67</b>
<b>100000043350</b>	<b>0.39</b>	<b>0.57</b>	<b>3492165</b>	<b>0.44</b>	<b>0.47</b>
<b>3517610</b>	<b>0.53</b>	<b>0.47</b>	3487747	0.26	<b>0.23</b>
<b>3595370</b>	<b>0.59</b>	<b>0.51</b>	<b>3517708</b>	<b>0.46</b>	<b>0.54</b>
<b>3513630</b>	<b>0.66</b>	<b>0.68</b>	<b>3595372</b>	<b>0.64</b>	<b>0.75</b>
<b>100000043360</b>	<b>0.35</b>	<b>0.55</b>	3517691	0.62	0.64
<b>100000048821</b>	<b>0.46</b>	<b>0.64</b>	3487615	0.62	0.58
<b>3595371</b>	<b>0.25</b>	<b>0.37</b>	<b>3487734</b>	<b>0.39</b>	<b>0.44</b>
3491634	0.32	0.29	3487898	0.52	0.51

Note. Bold-faced number indicates a BCR or ECR item.



**Table 1.65 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 7 Form F**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2004	100000043335	BCR_A	11,667	0.55	0.50	40.45	55.30	N/A	N/A	4.25
2004	3595368	BCR_B	11,667	1.28	0.87	22.01	16.18	55.90	N/A	5.91
2007	3487765	ECR_A	2,174	0.37	0.48	61.68	36.66	N/A	N/A	1.66
2007	3595369	ECR_B	2,174	1.47	0.83	11.32	27.92	49.95	6.30	4.51
2006	3517610	BCR_A	26,296	0.53	0.50	41.23	52.59	N/A	N/A	6.18
2006	3595370	BCR_B	26,296	1.18	0.60	21.22	25.73	46.34	N/A	6.71
2005	100000048821	BCR_A	13,390	0.46	0.50	46.37	46.10	N/A	N/A	7.53
2005	3595371	BCR_B	13,390	0.49	0.36	40.75	48.72	0.25	N/A	10.28
2005	3547487	ECR_A	13,123	0.66	0.47	26.34	65.77	N/A	N/A	7.89
2005	3564031	ECR_B	13,123	0.70	0.33	26.59	59.00	5.49	0.14	8.78
2007	3517648	ECR_A	32,000	0.63	0.48	33.25	63.38	N/A	N/A	3.37
2007	3564027	ECR_B	32,000	1.75	0.91	10.73	11.83	55.75	17.37	4.32
2006	3517708	BCR_A	39,533	0.46	0.50	42.03	46.00	N/A	N/A	11.97
2006	3595372	BCR_B	39,533	1.28	0.55	8.46	30.49	48.83	N/A	12.23
2008	100000043335	BCR_A	31,048	0.68	0.47	31.33	67.52	N/A	N/A	1.15
2008	3595368	BCR_B	31,048	1.60	0.68	9.17	18.32	70.69	N/A	1.81
2008	3487765	ECR_A	31,048	0.42	0.49	54.91	42.35	N/A	N/A	2.75
2008	3595369	ECR_B	31,048	1.51	0.80	8.29	29.32	50.91	6.52	4.96
2008	3517610	BCR_A	31,048	0.47	0.50	42.44	46.52	N/A	N/A	11.05
2008	3595370	BCR_B	31,048	1.03	0.82	19.49	32.13	35.38	N/A	13.01
2008	100000048821	BCR_A	31,048	0.64	0.48	32.78	64.45	N/A	N/A	2.78
2008	3595371	BCR_B	31,048	0.73	0.52	25.16	66.09	3.70	N/A	5.05
2008	3547487	ECR_A	31,048	0.81	0.39	15.98	81.16	N/A	N/A	2.86
2008	3564031	ECR_B	31,048	0.99	0.59	12.72	66.68	15.44	0.61	4.55
2008	3517648	ECR_A	31,048	0.67	0.47	31.60	66.88	N/A	N/A	1.52
2008	3564027	ECR_B	31,048	1.93	0.85	8.05	9.71	57.19	22.85	2.19
2008	3517708	BCR_A	31,048	0.54	0.50	40.50	53.54	N/A	N/A	5.97
2008	3595372	BCR_B	31,048	1.50	0.71	6.96	23.67	63.34	N/A	6.03

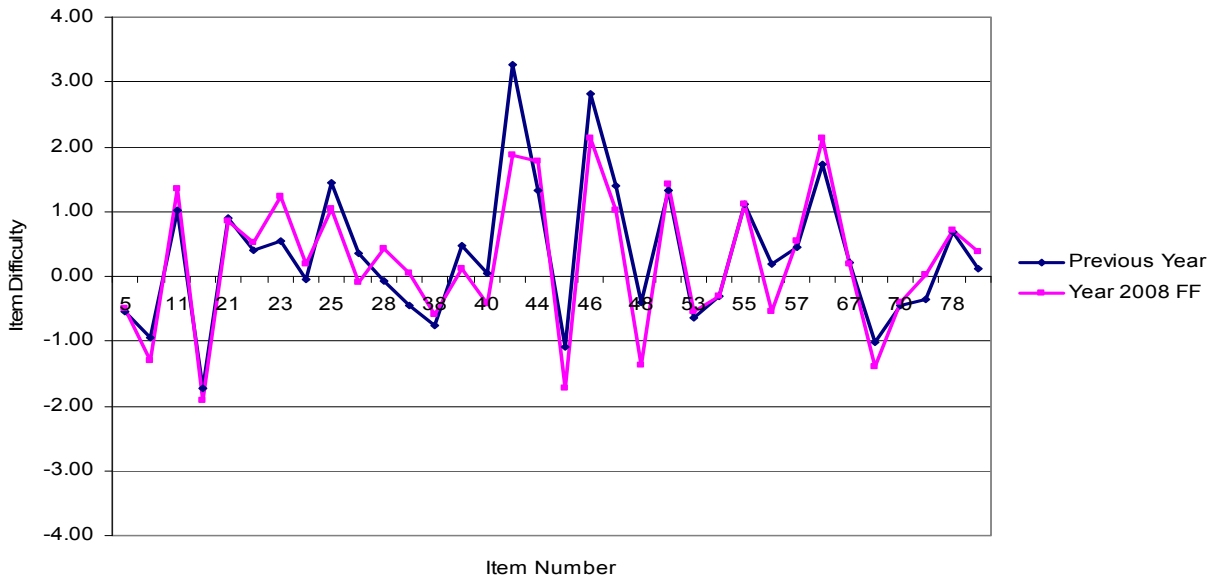
**Table 1.66 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008:  
Grade 7 Form F**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step	Step	Step
					0-1	1-2	2-3
2004	5	10000043335	BCR_A	-0.5514			
2004	6	3595368	BCR_B	-0.9471	0.3475	-0.3475	
2004	11	10000043349	SR	1.0088			
2006	17	3517739	SR	-1.7201			
2007	21	3487765	ECR_A	0.8991			
2007	22	3595369	ECR_B	0.3951	-2.2571	-0.9019	3.1591
2004	23	10000043344	SPR	0.5443			
2007	24	3513631	SPR	-0.0385			
2007	25	3487596	SPR	1.4391			
2004	26	10000043350	SPR	0.3584			
2006	28	3517610	BCR_A	-0.0757			
2006	29	3595370	BCR_B	-0.4436	-0.3197	0.3197	
2007	38	3513630	SPR	-0.7477			
2004	39	10000043360	SPR	0.4777			
2005	40	10000048821	BCR_A	0.0541			
2005	41	3595371	BCR_B	3.2743	-3.4653	3.4653	
2007	44	3491634	SR	1.3329			
2005	45	3547487	ECR_A	-1.0830			
2005	46	3564031	ECR_B	2.8049	-3.8213	0.685	3.1363
2004	47	10000043354	SPR	1.4061			
2004	48	10000043356	SPR	-0.4084			
2004	50	10000043338	SR	1.3356			
2007	53	3517648	ECR_A	-0.6275			
2007	54	3564027	ECR_B	-0.3188	-0.9499	-1.4821	2.432
2007	55	3492169	SPR	1.1052			
2004	56	10000043342	SPR	0.1876			
2007	57	3492165	SPR	0.4581			
2007	63	3487747	SR	1.7313			
2006	67	3517708	BCR_A	0.2039			
2006	68	3595372	BCR_B	-1.0212	-0.8667	0.8667	
2006	70	3517691	SR	-0.4573			
2007	72	3487615	SR	-0.3572			
2007	78	3487734	SPR	0.6953			
2007	81	3487898	SR	0.1081			
2008	5	10000043335	BCR_A	-0.4860			
2008	6	3595368	BCR_B	-1.3132	-0.2339	0.2339	
2008	11	10000043349	SR	1.3415			
2008	17	3517739	SR	-1.9183			
2008	21	3487765	ECR_A	0.8535			
2008	22	3595369	ECR_B	0.5148	-2.3799	-0.9260	3.3059
2008	23	10000043344	SPR	1.2419			
2008	24	3513631	SPR	0.1934			
2008	25	3487596	SPR	1.0476			
2008	26	10000043350	SPR	-0.0987			

**Table 1.66 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	28	3517610	BCR_A	0.4313			
2008	29	3595370	BCR_B	0.0359	-0.7781	0.7781	
2008	38	3513630	SPR	-0.5989			
2008	39	100000043360	SPR	0.1275			
2008	40	100000048821	BCR_A	-0.4202			
2008	41	3595371	BCR_B	1.8783	-2.7673	2.7673	
2008	44	3491634	SR	1.7720			
2008	45	3547487	ECR_A	-1.7311			
2008	46	3564031	ECR_B	2.1381	-4.0409	0.4064	3.6345
2008	47	100000043354	SPR	1.0066			
2008	48	100000043356	SPR	-1.3785			
2008	50	100000043338	SR	1.4309			
2008	53	3517648	ECR_A	-0.5360			
2008	54	3564027	ECR_B	-0.3195	-0.8801	-1.5639	2.4440
2008	55	3492169	SPR	1.1079			
2008	56	100000043342	SPR	-0.5518			
2008	57	3492165	SPR	0.5345			
2008	63	3487747	SR	2.1330			
2008	67	3517708	BCR_A	0.1843			
2008	68	3595372	BCR_B	-1.3958	-0.7581	0.7581	
2008	70	3517691	SR	-0.4072			
2008	72	3487615	SR	0.0253			
2008	78	3487734	SPR	0.7117			
2008	81	3487898	SR	0.3744			

Note. Rasch item and step difficulties are on a common scale.

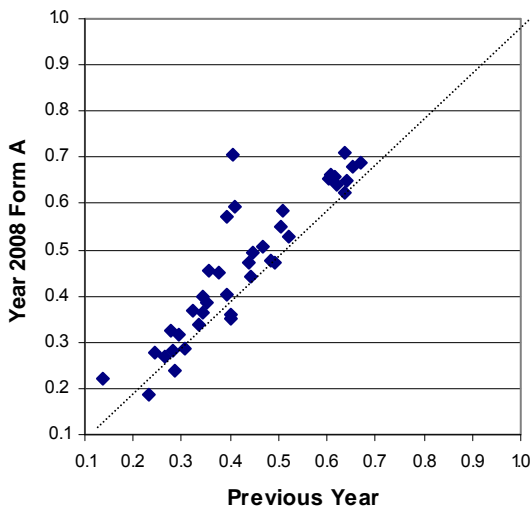


**Figure 1.10 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 7 Form F**

**Table 1.67 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 8 Form A**

Item CID	Previous Year	Year 08 Form A	Item CID	Previous Year	Year 08 Form A
<b>3514013</b>	<b>0.44</b>	<b>0.47</b>	<b>3514607</b>	<b>0.26</b>	<b>0.27</b>
<b>3564107</b>	<b>0.64</b>	<b>0.65</b>	<b>3564112</b>	<b>0.24</b>	<b>0.28</b>
3500150	0.49	0.47	<b>3514118</b>	<b>0.09</b>	<b>0.10</b>
100000043330	0.38	0.45	<b>3564113</b>	<b>0.40</b>	<b>0.35</b>
100000043305	0.62	0.64	3487539	0.64	0.63
<b>3514702</b>	<b>0.28</b>	<b>0.33</b>	100000043311	0.40	0.36
<b>3564108</b>	<b>0.34</b>	<b>0.40</b>	3487525	0.47	0.50
<b>3513650</b>	<b>0.30</b>	<b>0.32</b>	3487540	0.60	0.65
<b>3514064</b>	<b>0.14</b>	<b>0.22</b>	<b>100000043313</b>	<b>0.41</b>	<b>0.59</b>
<b>3500166</b>	<b>0.34</b>	<b>0.34</b>	<b>000003595405</b>	<b>0.41</b>	<b>0.70</b>
<b>100000043325</b>	<b>0.39</b>	<b>0.57</b>	<b>3513638</b>	<b>0.34</b>	<b>0.36</b>
3514595	0.65	0.68	<b>3487542</b>	<b>0.45</b>	<b>0.49</b>
<b>3514267</b>	<b>0.35</b>	<b>0.39</b>	3514136	0.61	0.66
<b>3564110</b>	<b>0.62</b>	<b>0.66</b>	3487568	0.23	0.19
<b>3514263</b>	<b>0.51</b>	<b>0.58</b>	100000043304	0.28	0.28
<b>3487907</b>	<b>0.36</b>	<b>0.45</b>	<b>3500162</b>	<b>0.29</b>	<b>0.24</b>
100000043320	0.49	0.47	<b>3514079</b>	<b>0.31</b>	<b>0.29</b>
<b>3514117</b>	<b>0.32</b>	<b>0.37</b>	<b>3514669</b>	<b>0.51</b>	<b>0.55</b>
<b>3564111</b>	<b>0.39</b>	<b>0.40</b>	<b>3564114</b>	<b>0.63</b>	<b>0.71</b>
<b>3492059</b>	<b>0.44</b>	<b>0.44</b>	3487912	0.52	0.53
<b>3487708</b>	<b>0.67</b>	<b>0.69</b>			

Note. Bold-faced number indicates a BCR or ECR item.





**Table 1.68 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 8 Form A**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2007	3514013	BCR_A	32,836	0.44	0.50	52.34	43.85	N/A	N/A	3.81
2007	3564107	BCR_B	32,836	1.28	0.69	8.61	44.62	41.81	N/A	4.95
2007	3514702	ECR_A	32,836	0.28	0.45	65.83	27.65	N/A	N/A	6.51
2007	3564108	ECR_B	32,836	1.03	1.13	34.62	26.39	11.53	17.81	9.66
2007	3514267	BCR_A	32,836	0.35	0.48	61.06	34.98	N/A	N/A	3.96
2007	3564110	BCR_B	32,836	1.23	0.67	8.73	49.35	36.87	N/A	5.05
2007	3514117	BCR_A	32,836	0.32	0.47	57.53	32.11	N/A	N/A	10.36
2007	3564111	BCR_B	32,836	0.79	0.78	30.13	34.58	21.97	N/A	13.31
2007	3514607	ECR_A	32,836	0.26	0.44	64.57	26.32	N/A	N/A	9.12
2007	3564112	ECR_B	32,836	0.73	1.05	49.17	12.90	15.01	10.02	12.89
2007	3514118	BCR_A	32,836	0.09	0.29	86.62	9.37	N/A	N/A	4.00
2007	3564113	BCR_B	32,836	0.80	0.49	18.42	72.25	3.99	N/A	5.34
2004	100000043313	ECR_A	12,814	0.41	0.49	54.68	40.92	N/A	N/A	4.39
2004	3595405	ECR_B	12,814	1.22	1.24	39.17	14.88	16.93	24.55	4.47
2007	3514669	BCR_A	32,836	0.51	0.50	41.96	50.51	N/A	N/A	7.53
2007	3564114	BCR_B	32,836	1.27	0.81	14.37	26.39	50.27	N/A	8.97
2008	3514013	BCR_A	32,318	0.47	0.50	50.68	47.01	N/A	N/A	2.31
2008	3564107	BCR_B	32,318	1.30	0.69	9.99	43.80	43.23	N/A	2.99
2008	3514702	ECR_A	32,318	0.33	0.47	63.84	32.64	N/A	N/A	3.53
2008	3564108	ECR_B	32,318	1.19	1.16	31.56	29.23	11.81	22.18	5.22
2008	3514267	BCR_A	32,318	0.39	0.49	58.52	38.52	N/A	N/A	2.95
2008	3564110	BCR_B	32,318	1.31	0.69	9.29	42.85	44.24	N/A	3.61
2008	3514117	BCR_A	32,318	0.37	0.48	55.99	36.99	N/A	N/A	7.02
2008	3564111	BCR_B	32,318	0.81	0.79	33.68	34.31	23.22	N/A	8.79
2008	3514607	ECR_A	32,318	0.27	0.44	64.36	26.92	N/A	N/A	8.72
2008	3564112	ECR_B	32,318	0.83	1.09	45.22	13.36	17.36	11.71	12.35
2008	3514118	BCR_A	32,318	0.10	0.30	85.89	10.08	N/A	N/A	4.02
2008	3564113	BCR_B	32,318	0.71	0.54	28.88	62.68	3.92	N/A	4.52
2008	100000043313	ECR_A	32,318	0.59	0.49	37.84	59.42	N/A	N/A	2.74
2008	3595405	ECR_B	32,318	2.11	1.04	6.88	17.26	22.06	49.94	3.86
2008	3514669	BCR_A	32,318	0.55	0.50	39.26	55.18	N/A	N/A	5.56
2008	3564114	BCR_B	32,318	1.42	0.76	10.25	23.93	59.06	N/A	6.75

**Table 1.69 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008:  
Grade 8 Form A**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2007	3	3514013	BCR_A	0.3616			
2007	4	3564107	BCR_B	-1.1070	-1.441	1.441	
2007	6	3500150	SR	0.1886			
2004	8	100000043330	SR	0.3962			
2005	14	100000043305	SR	-0.8459			
2007	15	3514702	ECR_A	1.2761			
2007	16	3564108	ECR_B	0.6901	-0.7491	0.5272	0.2219
2007	17	3513650	SPR	1.1558			
2007	18	3514064	SPR	2.7466			
2007	19	3500166	SPR	0.8807			
2004	20	100000043325	SPR	0.1969			
2006	21	3514595	SR	-0.8658			
2007	23	3514267	BCR_A	0.8169			
2007	24	3564110	BCR_B	-0.9309	-1.4936	1.4936	
2007	25	3514263	SPR	-0.1362			
2007	26	3487907	SPR	0.7793			
2005	31	100000043320	SR	-0.1139			
2007	34	3514117	BCR_A	0.9736			
2007	35	3564111	BCR_B	0.4284	-0.8203	0.8203	
2007	36	3492059	SPR	0.2971			
2007	37	3487708	SPR	-1.0544			
2007	39	3514607	ECR_A	1.2953			
2007	40	3564112	ECR_B	1.2629	0.1082	-0.8532	0.745
2007	44	3514118	BCR_A	2.8471			
2007	45	3564113	BCR_B	1.0451	-2.7281	2.7281	
2007	48	3487539	SR	-0.7390			
2004	49	100000043311	SR	0.2180			
2007	50	3487525	SR	0.0654			
2007	51	3487540	SR	-0.5923			
2004	54	100000043313	ECR_A	0.1847			
2004	55	3595405	ECR_B	0.1002	0.1515	-0.3165	0.165
2007	56	3513638	SPR	0.8689			
2007	57	3487542	SPR	0.2362			
2006	59	3514136	SR	-0.6909			
2007	65	3487568	SR	1.3814			
2005	68	100000043304	SR	1.0478			
2007	74	3500162	SPR	1.2222			
2007	75	3514079	SPR	1.2068			
2007	76	3514669	BCR_A	-0.1522			
2007	77	3564114	BCR_B	-0.8897	-0.4608	0.4608	
2007	78	3487912	SR	-0.1038			
2008	3	3514013	BCR_A	0.2481			
2008	4	3564107	BCR_B	-0.9490	-1.2905	1.2905	
2008	6	3500150	SR	0.3158			

**Table 1.69 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	8	100000043330	SR	0.3621			
2008	14	100000043305	SR	-0.7079			
2008	15	3514702	ECR_A	1.0904			
2008	16	3564108	ECR_B	0.5369	-0.8401	0.6976	0.1425
2008	17	3513650	SPR	1.1148			
2008	18	3514064	SPR	1.6566			
2008	19	3500166	SPR	0.9813			
2008	20	100000043325	SPR	-0.4508			
2008	21	3514595	SR	-0.854			
2008	23	3514267	BCR_A	0.6788			
2008	24	3564110	BCR_B	-0.9812	-1.2635	1.2635	
2008	25	3514263	SPR	-0.3795			
2008	26	3487907	SPR	0.3647			
2008	31	100000043320	SR	0.2581			
2008	34	3514117	BCR_A	0.7094	-0.7657	0.7657	
2008	35	3564111	BCR_B	0.5525			
2008	36	3492059	SPR	0.3553			
2008	37	3487708	SPR	-1.0683			
2008	39	3514607	ECR_A	1.3428			
2008	40	3564112	ECR_B	1.1814	0.0189	-0.8420	0.8231
2008	44	3514118	BCR_A	2.969			
2008	45	3564113	BCR_B	1.6151	-2.4991	2.4991	
2008	48	3487539	SR	-0.6178			
2008	49	100000043311	SR	0.8435			
2008	50	3487525	SR	0.0551			
2008	51	3487540	SR	-0.7102			
2008	54	100000043313	ECR_A	-0.5034			
2008	55	3595405	ECR_B	-1.1346	-0.08613	0.1993	0.6621
2008	56	3513638	SPR	0.8139			
2008	57	3487542	SPR	0.0986			
2008	59	3514136	SR	-0.891			
2008	65	3487568	SR	2.0087			
2008	68	100000043304	SR	1.3321			
2008	74	3500162	SPR	1.5218			
2008	75	3514079	SPR	1.2839			
2008	76	3514669	BCR_A	-0.255			
2008	77	3564114	BCR_B	-1.2583	-0.4250	0.4250	
2008	78	3487912	SR	-0.0934			

Note. Rasch item and step difficulties are on a common scale.

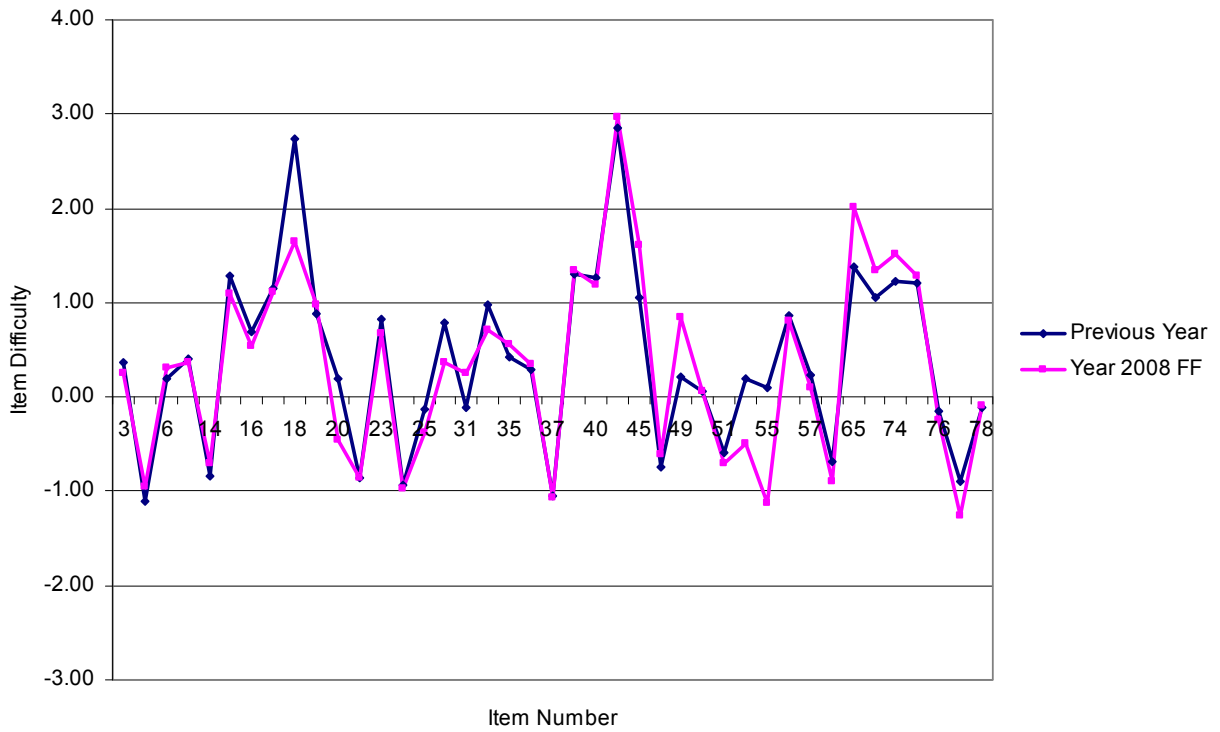
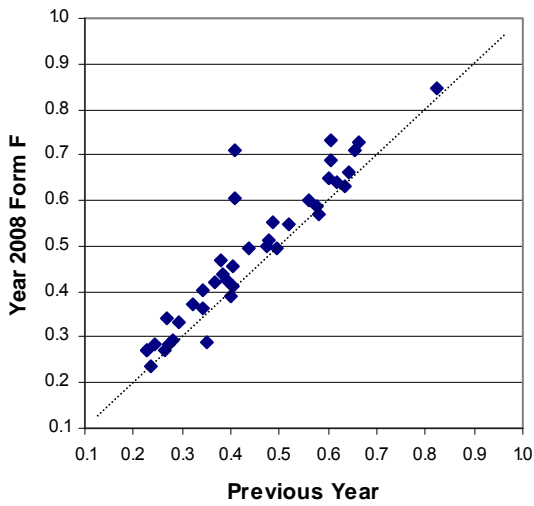


Figure 1.11 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 8 Form A

**Table 1.70 P-Value Comparisons of Core Items for Previous Year vs. Year 2008: Grade 8 Form F**

Item CID	Previous Year	Year 08 Form F	Item CID	Previous Year	Year 08 Form F
<b>3514013</b>	<b>0.44</b>	<b>0.49</b>	<b>3519804</b>	<b>0.23</b>	<b>0.27</b>
<b>3564107</b>	<b>0.64</b>	<b>0.66</b>	<b>3514607</b>	<b>0.26</b>	<b>0.27</b>
3487526	0.56	0.60	<b>3564112</b>	<b>0.24</b>	<b>0.28</b>
100000043330	0.38	0.47	<b>3514266</b>	<b>0.29</b>	<b>0.33</b>
100000043305	0.62	0.64	<b>3564120</b>	<b>0.49</b>	<b>0.50</b>
<b>3514283</b>	<b>0.34</b>	<b>0.40</b>	3487539	0.64	0.63
<b>3564116</b>	<b>0.49</b>	<b>0.55</b>	3487901	0.82	0.85
<b>3492049</b>	<b>0.58</b>	<b>0.59</b>	3487540	0.60	0.65
<b>100000043307</b>	<b>0.27</b>	<b>0.28</b>	<b>100000043313</b>	<b>0.41</b>	<b>0.61</b>
<b>3514162</b>	<b>0.37</b>	<b>0.42</b>	<b>3595405</b>	<b>0.41</b>	<b>0.71</b>
<b>3487563</b>	<b>0.40</b>	<b>0.46</b>	<b>3487913</b>	<b>0.34</b>	<b>0.36</b>
3514595	0.65	0.71	<b>3514167</b>	<b>0.52</b>	<b>0.55</b>
<b>3514217</b>	<b>0.23</b>	<b>0.27</b>	3514136	0.61	0.69
<b>3595406</b>	<b>0.38</b>	<b>0.44</b>	3492047	0.27	0.34
<b>3513648</b>	<b>0.58</b>	<b>0.57</b>	100000043304	0.28	0.29
<b>100000043314</b>	<b>0.35</b>	<b>0.29</b>	<b>3487721</b>	<b>0.47</b>	<b>0.50</b>
3500154	0.66	0.73	<b>3492052</b>	<b>0.24</b>	<b>0.24</b>
<b>3514117</b>	<b>0.32</b>	<b>0.37</b>	<b>3514709</b>	<b>0.48</b>	<b>0.51</b>
<b>3564111</b>	<b>0.39</b>	<b>0.43</b>	<b>3595408</b>	<b>0.61</b>	<b>0.73</b>
<b>3514114</b>	<b>0.41</b>	<b>0.41</b>	3487672	0.40	0.39

Note. Bold-faced number indicates a BCR or ECR item.



**Table 1.71 Score-Point Distribution Comparisons of Constructed Response Core Items for Previous Year vs. Year 2008: Grade 8 Form F**

Year	Item CID	Item Type	N	Mean	SD	Score-Point Distribution (%)				
						0	1	2	3	Omit
2007	3514013	BCR_A	32,836	0.44	0.50	52.34	43.85	N/A	N/A	3.81
2007	3564107	BCR_B	32,836	1.28	0.69	8.61	44.62	41.81	N/A	4.95
2007	3514283	ECR_A	32,480	0.34	0.48	59.90	34.44	N/A	N/A	5.66
2007	3564116	ECR_B	32,480	1.46	1.00	7.40	46.89	15.97	22.24	7.52
2006	3514217	BCR_A	27,033	0.23	0.42	69.06	22.59	N/A	N/A	8.35
2006	3595406	BCR_B	27,033	0.76	0.35	20.82	70.44	2.98	N/A	5.76
2007	3514117	BCR_A	32,836	0.32	0.47	57.53	32.11	N/A	N/A	10.36
2007	3564111	BCR_B	32,836	0.79	0.78	30.13	34.58	21.97	N/A	13.31
2007	3514607	ECR_A	32,836	0.26	0.44	64.57	26.32	N/A	N/A	9.12
2007	3564112	ECR_B	32,836	0.73	1.05	49.17	12.90	15.01	10.02	12.89
2007	3514266	BCR_A	32,480	0.29	0.46	65.02	29.33	N/A	N/A	5.66
2007	3564120	BCR_B	32,480	0.99	0.72	18.94	47.63	25.58	N/A	7.85
2004	100000043313	ECR_A	12,814	0.41	0.49	54.68	40.92	N/A	N/A	4.39
2004	3595405	ECR_B	12,814	1.22	1.24	39.17	14.88	16.93	24.55	4.47
2006	3514709	BCR_A	3,524	0.48	0.50	43.59	47.90	N/A	N/A	8.51
2006	3595408	BCR_B	3,524	1.21	0.57	15.66	29.80	45.72	N/A	8.83
2008	3514013	BCR_A	31,743	0.49	0.50	48.17	49.33	N/A	N/A	2.50
2008	3564107	BCR_B	31,743	1.33	0.68	8.79	42.53	45.03	N/A	3.65
2008	3514283	ECR_A	31,743	0.40	0.49	57.00	40.29	N/A	N/A	2.71
2008	3564116	ECR_B	31,743	1.66	0.98	4.88	44.63	18.27	28.15	4.07
2008	3514217	BCR_A	31,743	0.27	0.45	69.32	27.26	N/A	N/A	3.42
2008	3595406	BCR_B	31,743	0.88	0.46	12.33	77.51	5.13	N/A	5.04
2008	3514117	BCR_A	31,743	0.37	0.48	55.13	37.26	N/A	N/A	7.61
2008	3564111	BCR_B	31,743	0.85	0.79	29.01	35.61	24.73	N/A	10.65
2008	3514607	ECR_A	31,743	0.27	0.44	63.81	27.07	N/A	N/A	9.13
2008	3564112	ECR_B	31,743	0.85	1.09	43.12	14.36	17.20	11.97	13.35
2008	3514266	BCR_A	31,743	0.33	0.47	60.28	33.36	N/A	N/A	6.36
2008	3564120	BCR_B	31,743	0.99	0.73	18.71	46.43	26.39	N/A	8.47
2008	100000043313	ECR_A	31,743	0.61	0.49	36.74	60.54	N/A	N/A	2.72
2008	3595405	ECR_B	31,743	2.13	1.02	5.65	18.15	22.85	49.58	3.77
2008	3514709	BCR_A	31,743	0.51	0.50	44.72	51.12	N/A	N/A	4.16
2008	3595408	BCR_B	31,743	1.46	0.71	7.94	28.05	59.09	N/A	4.93

**Table 1.72 Rasch Item and Step Difficulty Comparisons of Core Items for Previous Year vs. Year 2008:  
Grade 8 Form F**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2007	3	3514013	BCR_A	0.3616			
2007	4	3564107	BCR_B	-1.1070	-1.441	1.441	
2007	6	3487526	SR	-0.248			
2004	8	100000043330	SR	0.3962			
2005	14	100000043305	SR	-0.8459			
2007	15	3514283	ECR_A	0.8146			
2007	16	3564116	ECR_B	-0.2444	-2.2962	1.2817	1.0145
2007	17	3492049	SPR	-0.475			
2004	18	100000043307	SPR	0.8978			
2007	19	3514162	SPR	0.6672			
2007	20	3487563	SPR	0.3705			
2006	21	3514595	SR	-0.8658			
2006	23	3514217	BCR_A	1.5404			
2006	24	3595406	BCR_B	1.3259	-2.8804	2.8804	
2007	25	3513648	SPR	-0.4536			
2004	26	100000043314	SPR	0.4873			
2007	31	3500154	SR	-0.9169			
2007	34	3514117	BCR_A	0.9736			
2007	35	3564111	BCR_B	0.4284	-0.8203	0.8203	
2007	36	3514114	SPR	0.4361			
2007	37	3519804	SPR	1.6473			
2007	39	3514607	ECR_A	1.2953			
2007	40	3564112	ECR_B	1.2629	0.1082	-0.8532	0.745
2007	44	3514266	BCR_A	1.2135			
2007	45	3564120	BCR_B	-0.0476	-1.3415	1.3415	
2007	48	3487539	SR	-0.739			
2007	49	3487901	SR	-1.9759			
2007	51	3487540	SR	-0.5923			
2004	54	100000043313	ECR_A	0.1847			
2004	55	3595405	ECR_B	0.1002	0.1515	-0.3165	0.165
2007	56	3487913	SPR	0.8646			
2006	57	3514167	SPR	-0.3444			
2006	59	3514136	SR	-0.6909			
2007	66	3492047	SR	1.2678			
2005	68	100000043304	SR	1.0478			
2007	74	3487721	SPR	0.0882			
2007	75	3492052	SPR	1.4609			
2006	76	3514709	BCR_A	-0.0807			
2006	77	3595408	BCR_B	-0.7861	-0.5116	0.5116	
2007	78	3487672	SR	0.5495			
2008	3	3514013	BCR_A	0.1635			
2008	4	3564107	BCR_B	-1.032	-1.3764	1.3764	
2008	6	3487526	SR	-0.3818			
2008	8	100000043330	SR	0.4265			

**Table 1.72 (continued)**

Year	Item Seq. No.	Item CID	Item Type	Item Difficulty	Step 0-1	Step 1-2	Step 2-3
2008	14	100000043305	SR	-0.6357			
2008	15	3514283	ECR_A	0.7105			
2008	16	3564116	ECR_B	-0.5004	-2.4388	1.4114	1.0274
2008	17	3492049	SPR	-0.391			
2008	18	100000043307	SPR	1.4660			
2008	19	3514162	SPR	0.6185			
2008	20	3487563	SPR	0.4457			
2008	21	3514595	SR	-0.9382			
2008	23	3514217	BCR_A	1.5261			
2008	24	3595406	BCR_B	0.9485	-2.9439	2.9439	
2008	25	3513648	SPR	-0.2496			
2008	26	100000043314	SPR	1.3827			
2008	31	3500154	SR	-1.1600			
2008	34	3514117	BCR_A	0.8458			
2008	35	3564111	BCR_B	0.4167	-0.8046	0.8046	
2008	36	3514114	SPR	0.5733			
2008	37	3519804	SPR	1.5395			
2008	39	3514607	ECR_A	1.4386			
2008	40	3564112	ECR_B	1.2339	-0.1186	-0.6697	0.7883
2008	44	3514266	BCR_A	1.123			
2008	45	3564120	BCR_B	0.0332	-1.3167	1.3167	
2008	48	3487539	SR	-0.592			
2008	49	3487901	SR	-2.1555			
2008	51	3487540	SR	-0.6311			
2008	54	100000043313	ECR_A	-0.4647			
2008	55	3595405	ECR_B	-1.1127	-1.1482	0.3695	0.7787
2008	56	3487913	SPR	0.9479			
2008	57	3514167	SPR	-0.2011			
2008	59	3514136	SR	-0.9134			
2008	66	3492047	SR	1.0119			
2008	68	100000043304	SR	1.3763			
2008	74	3487721	SPR	0.1096			
2008	75	3492052	SPR	1.7004			
2008	76	3514709	BCR_A	0.0865			
2008	77	3595408	BCR_B	-1.3311	-0.7377	0.7377	
2008	78	3487672	SR	0.7592			

Note. Rasch item and step difficulties are on a common scale.



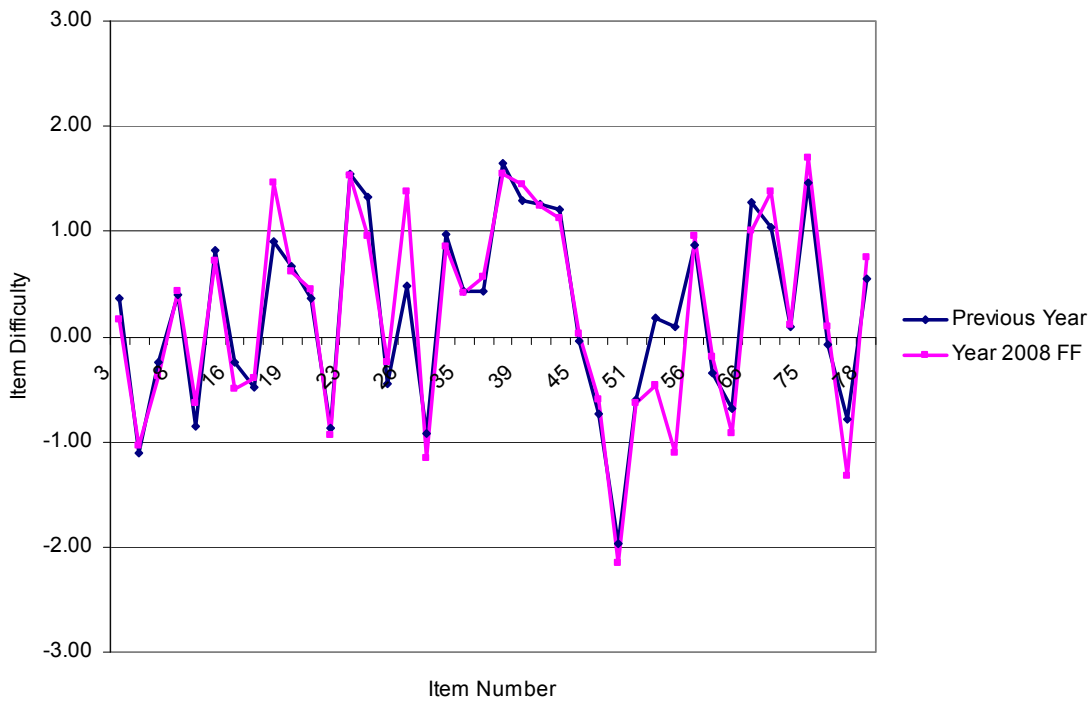


Figure 1.12 Rasch Item Difficulty Comparisons of Core Items for Previous Year vs. Year 2008: Grade 8 Form F

## 1.8 Field Test Analyses

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 2008 Maryland item bank. Information on the item bank can be found in section 1.14, *Item Bank Construction*. 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 ethnic 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 2008 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 2008 field test items on the base scale (i.e., the 2006 scale), each field test item was freely calibrated by fixing Rasch item and step parameters of the 2008 operational items that had been already placed on the base scale during the 2008 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 2008 Maryland item bank. These field test items are eligible to be used as operational items in subsequent years.

## 1.9 Operational Test Construction Using the Rasch Model

The selection of items to be included in the final operational test forms of the 2008 MSA-Math required a careful consideration based on test blueprints, classical item analyses, *DIF* analyses, and IRT analyses. Specifically, the IRT method played a major role in constructing the 2008 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 2008 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 2008 MSA-Math. The following table and figure show an example of the 2008 MSA-Math operational test form construction using the IRT method. Further information on other grades can be obtained from MSDE.

**Table 1.73 An Example of the 2008 Math Operational Test Construction Using the Rasch Model**

Item Type	P-value	a	b1	b2	B3
BCR_A	0.67	1.00	0.0868		
BCR_B	0.55	1.00	-0.9244	2.2968	
BCR_A	0.34	1.00	1.8944		
BCR_B	0.34	1.00	0.9465	2.9753	
BCR_A	0.44	1.00	1.2891		
BCR_B	0.52	1.00	-1.7833	3.1141	
BCR_A	0.50	1.00	1.0216		
BCR_B	0.45	1.00	1.7577	0.6851	
BCR_A	0.31	1.00	1.9152		
BCR_B	0.22	1.00	1.2744	5.0228	
BCR_A	0.79	1.00	-0.6075		
BCR_B	0.47	1.00	-0.351	2.8802	
BCR_A	0.3	1.00	2.0909		
BCR_B	0.39	1.00	0.838	2.3844	
ECR_A	0.84	1.00	-1.0768		
ECR_B	0.52	1.00	-3.0549	1.1937	3.6636
SR	0.67	1.00	0.2030		
SR	0.78	1.00	-0.3310		
SR	0.71	1.00	0.0148		
SR	0.82	1.00	-1.0845		
SR	0.41	1.00	1.5483		
SR	0.40	1.00	1.5795		
SR	0.64	1.00	0.6342		
SR	0.57	1.00	0.8118		
SR	0.84	1.00	-1.1516		
SR	0.80	1.00	-0.5779		
SR	0.62	1.00	0.5383		
SR	0.86	1.00	-0.9093		
SR	0.82	1.00	-0.6898		
SR	0.51	1.00	0.6218		
SR	0.63	1.00	0.1746		
SR	0.91	1.00	-1.2550		
SR	0.85	1.00	-1.1293		
SR	0.68	1.00	0.2895		
SR	0.81	1.00	-0.6828		
SR	0.56	1.00	0.6094		
SR	0.87	1.00	-1.3086		
SR	0.63	1.00	0.4633		
SR	0.81	1.00	-0.9892		
SR	0.75	1.00	-0.5025		
SR	0.92	1.00	-1.7042		
SR	0.89	1.00	-1.6381		
SR	0.40	1.00	1.6552		
SR	0.64	1.00	0.2606		
SR	0.66	1.00	0.2013		

**Table 1.73 (continued)**

Type	P-value	a	b1	b2	b3
SR	0.38	1.00	1.6628		
SR	0.53	1.00	0.8800		
SR	0.74	1.00	-0.2530		
SR	0.75	1.00	-0.3576		
SR	0.46	1.00	0.9885		
SR	0.83	1.00	-1.0402		
SR	0.58	1.00	0.6766		
SR	0.52	1.00	0.8102		
SR	0.42	1.00	1.4497		
SR	0.91	1.00	-1.7982		
SR	0.72	1.00	-0.2014		
SR	0.60	1.00	0.5575		
SR	0.69	1.00	-0.0616		
SR	0.64	1.00	-0.3909		
SR	0.75	1.00	-0.8175		
SR	0.78	1.00	-0.6839		
SR	0.85	1.00	1.2809		
SR	0.53	1.00	0.0910		
SR	0.77	1.00	-0.2779		
SR	0.56	1.00	0.4459		

Note. a: item discrimination; b1: step 1 difficulty; b2: step 2 difficulty; b3: step 3 difficulty

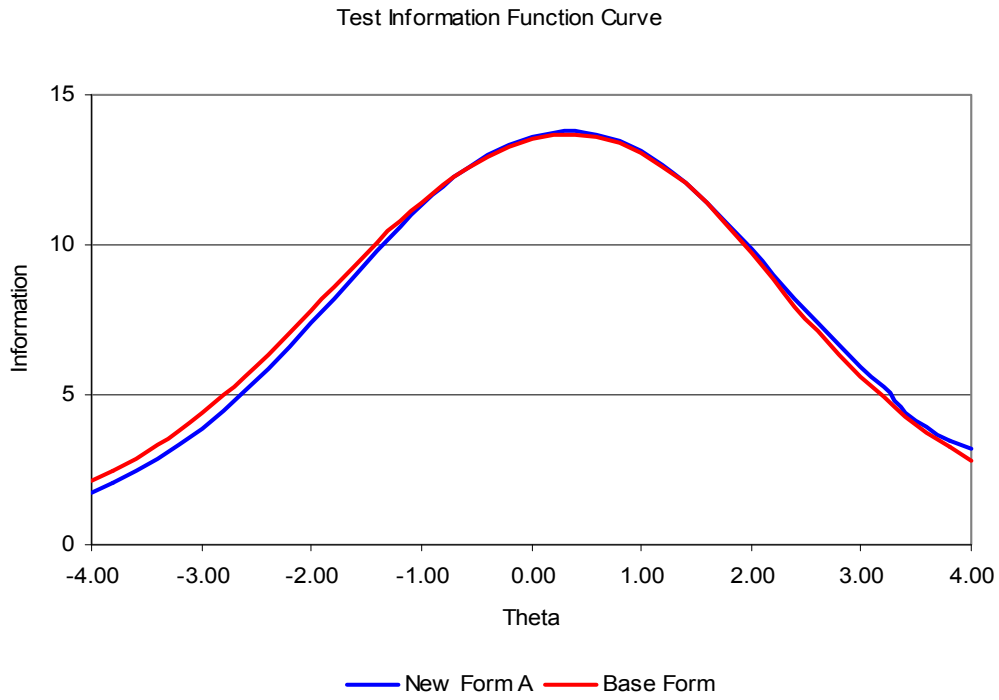


Figure 1.13 Test Information Curves of Base Form vs. Current Year’s Math Operational Test Form

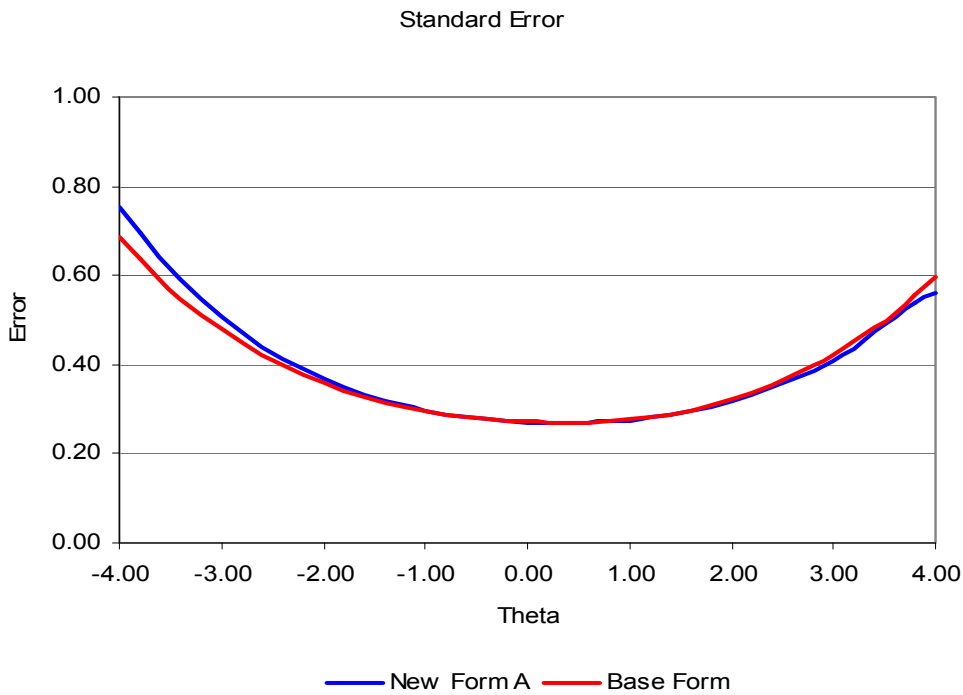


Figure 1.14 Standard Errors of Base Form vs. Current Year’s Math Operational Test Form

## 1.10 Linking, Equating, and Scaling Procedures of the 2008 MSA-Math

The 2008 MSA-Math was calibrated, equated, and scaled by fixing the item parameters of the operational items which appeared on the 2008 and previous operational forms (i.e., Rasch item fixed method). This means that Rasch item difficulty parameters on the common scale of the 2006 assessment were carried and fixed during the linking and equating process. It should be noted that Rasch recalibration due to IRT model transition (i.e., from 3-PL to the Rasch) was conducted using the 2006 MSA-Math data. Detailed information on the 2006 Rasch recalibration can be obtained in the 2007 MSA-Math technical report.

### Stratified Random Sampling Procedures

To select equating samples, a stratified random sampling method was applied to the 2008 state examinee population. To verify that the sample was representative of the statewide examinee population in terms of school district, gender, and ethnicity, the distributions of LEA, gender, and ethnicity of the 2008 sample were compared with those of the 2008 population. Appendix A, *The 2008 MSA-Math Stratified Random Sampling*, provides the results of 2008 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

Robust z values 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 (SCDE, 2001):

- Try not to include items with an absolute value of robust  $z$  exceeding 1.645. In addition, 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.
- Try to maintain that the ratio of the standard deviations between two operational forms is in the 90 to 110 percent range.
- Try to maintain the correlation between two operational forms is greater than .95.

### **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

After examining the robust  $z$  and correlations from 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 2008 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 1.56 through 1.61 include Rasch item difficulties used for calculating robust  $z$  values, correlation coefficients, and standard deviations. Figures 1.14 through 1.37 depict item difficulty plots between current and previous years. It should be noted that the item difficulties of the 2008 operational forms were obtained from independent calibration, and those of previous assessments were on a common scale (i.e., linked to the 2006 assessment).



**Table 1.74 Rasch Item Difficulties and Robust Z values for Previous Year vs. Year 2008: Grade 3**

Item Seq No.	Previous Year	Y2008 Form A	Robust Z	Item Seq No.	Previous Year	Y2008 Form F	Robust Z
1	0.9627	0.3194	-.3882	1	0.9627	0.5535	.0888
2	0.6288	-0.0080	-.3534	2	0.6288	0.0253	-.7873
5	0.0690	-0.6618	-.8566	5	0.0690	-0.5094	-.6741
6	1.0359	0.5974	.7083	6	1.0359	0.8025	.8815
7	0.5502	-0.4334	-2.2101	7	0.5502	-0.2661	-1.7468
8	1.8411	0.7657	-2.7016	8	1.8411	0.9343	-2.1549
14	0.8740	0.2276	-.4048	14	0.8740	0.2551	-.8567
16	0.3981	-0.3428	-.9107	16	0.3981	-0.3372	-1.3816
17	-0.0360	-0.5435	.3389	17	-0.0360	-0.428	.1664
21	-0.3305	-0.8089	.4947	21	-0.3305	-0.9223	-.7345
22	2.0077	1.2651	-.9198	22	2.0077	1.5038	-.3382
23	0.4123	-0.3835	-1.2047	23	0.4123	-0.1993	-.8238
24	1.2257	0.8441	1.0130	24	1.2257	1.0837	1.2936
32	1.2515	0.6779	-.0150	32	1.2515	0.8226	.0000
33	0.0425	-0.6008	-.3882	33	0.0425	-0.3706	.0712
45	2.4187	1.9288	.4331	45	2.4187	1.7525	-1.0700
47	-1.3667	-1.5247	2.2101	47	-1.3667	-1.571	1.0127
48	-2.1822	-2.4086	1.8439	48	-2.1822	-2.2903	1.4465
49	0.4861	-0.3019	-1.1629	49	0.4861	-0.1222	-.8089
50	-0.499	-0.9164	.8213	50	-0.499	-0.6287	1.3491
51	0.2953	-0.2755	.0000	51	0.2953	-0.0972	.1641
52	-0.6165	-1.2806	-.4995	52	-0.6165	-1.0805	-.1583
55	1.2952	0.7279	.0187	55	1.2952	0.946	.3594
56	-0.5906	-1.0743	.4663	56	-0.5906	-0.9372	.3711
62	0.9229	-0.0931	-2.3836	62	0.9229	0.0706	-1.9091
63	-0.2691	-0.7012	.7426	63	-0.2691	-0.5922	.4771
64	-0.6059	-1.1106	.3539	64	-0.6059	-0.9199	.5181
65	1.4814	0.9414	.1649	65	1.4814	1.2351	.8233
66	1.8021	1.1497	-.4369	66	1.8021	1.2884	-.3824
67	1.5719	1.0426	.2222	67	1.5719	1.2735	.5884
68	0.0473	-0.5327	-.0493	68	0.0473	-0.4908	-.4924
69	0.0444	-0.2374	1.5473	69	0.0444	-0.0897	1.3293
70	0.0993	-0.3569	.6136	70	0.0993	-0.2351	.4261
72	-0.6247	-1.4077	-1.1361	72	-0.6247	-1.2908	-1.0695
82	-0.5397	-1.0778	.1751	82	-0.5397	-0.804	.7422
41A	1.6971	0.9662	-.8572	18F	1.1953	0.7507	-.0708
44A	-0.4817	-0.1209	4.9878	29F	0.1797	0.091	1.5340
				31F	-0.2581	-1.0564	-1.6656
				76F	-0.1123	-0.5583	-.0771

*Note.* The 2008 item sequence number was used to indicate that it was the same item appearing across years.

*Note.* Each item parameter was generated with a live, stratified random sample (i.e., about 3,000 students) of the year.

*Note.* Item parameters of previous years were on the base scale.

*Note.* The 2008 items were independently calibrated with the 2008 stratified random sample.

**Form Statistics**

	Previous	2008	Previous	2008
Form Statistics	Base Form	Form A	Base Form	Form F
Mean	.414	-.155	.387	-.062
SD	1.007	.927	.966	.922

**Correlation and Standard Deviation Ratio**

	2008	2008
With Base Form	Form A	Form F
Correlation	.969	.975
SD Ratio	92%	95%

**Values Used for Robust Z Statistics**

	2008	2008
With Base Form	Form A	Form F
Mean Diff	-.569	-.449
Median Diff	-.571	-.429
IQR Diff	.252	.300

Based on correlation coefficients, SD ratios, robust z values, and item difficulty plot, none of the linking common items were dropped from the linking pool.

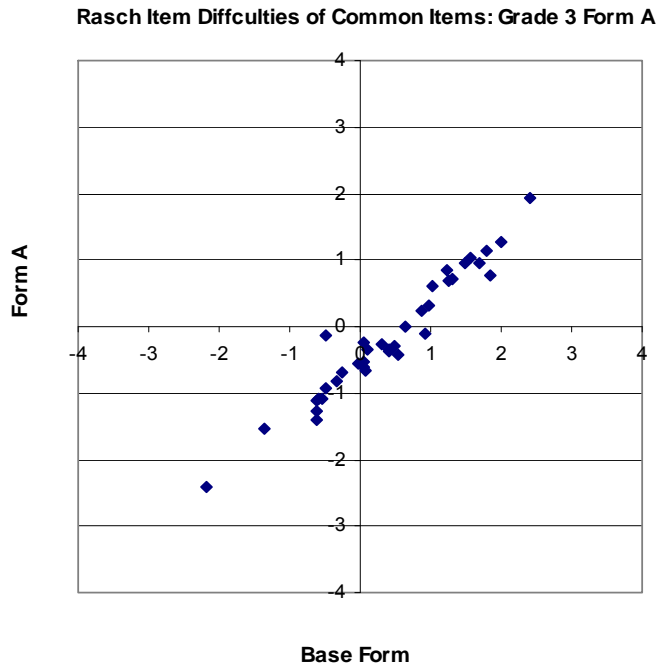


Figure 1.15 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 3 Form A

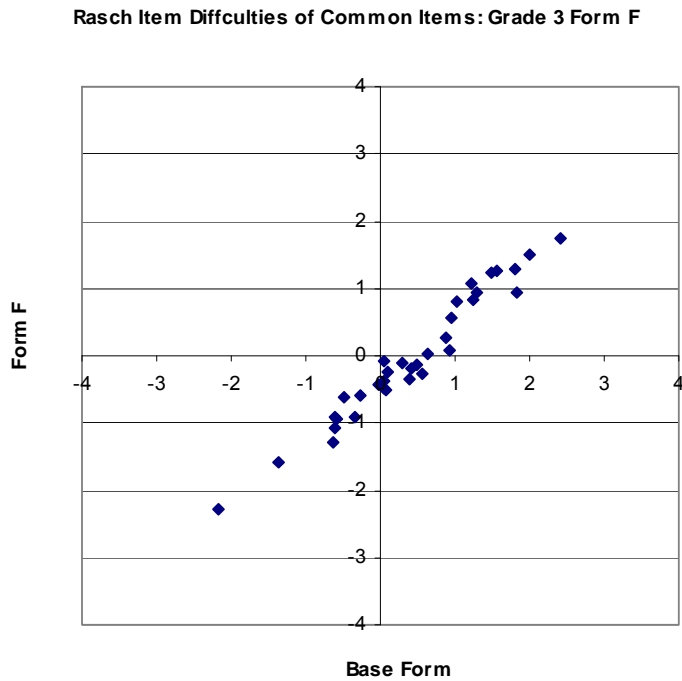


Figure 1.16 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 3 Form F



**Table 1.75 Rasch Item Difficulties and Robust Z values for Previous Year vs. Year 2008: Grade 4**

Item Seq No.	Previous Year	Y2008 Form A	Robust Z	Item Seq No.	Previous Year	Y2008 Form F	Robust Z
1	0.6241	0.5151	-.1921	1	0.6241	0.6010	-.3400
2	-0.7990	-0.8431	.0598	2	-0.7990	-0.9280	-.6862
6	0.1763	-0.2238	-1.3220	6	0.1763	-0.1258	-1.2521
7	-0.8522	-0.4349	1.8507	7	-0.8522	-0.4070	1.1909
8	-1.0550	-1.0645	.1941	8	-1.0550	-0.9838	-.0317
10	0.9009	0.7343	-.4157	10	0.9009	0.8346	-.4812
19	1.4979	1.2324	-.7996	19	1.4979	1.2857	-.9582
22	0.3940	0.3045	-.1164	22	0.3940	0.5280	.1736
24	0.5508	0.4526	-.1502	24	0.5508	0.6392	.0245
25	0.0797	0.0202	.0000	25	0.0797	0.1531	-.0245
26	1.7570	1.8603	.6319	26	1.7570	1.7678	-.2292
27	-0.1355	0.0119	.8031	27	-0.1355	0.0549	.3580
32	-0.9395	-0.508	1.9058	32	-0.9395	-0.4355	1.3831
34	-0.6701	-0.9158	-.7227	34	-0.6701	-0.8498	-.8519
<b>47</b>	<b>-0.1077</b>	<b>-1.392</b>	<b>-4.7541</b>	<b>47</b>	<b>-0.1077</b>	<b>-1.0369</b>	<b>-3.3021</b>
49	-0.9767	-.914	.4728	49	-0.9767	-0.7666	.4224
50	0.9291	1.233	1.4102	50	0.9291	1.2345	.7339
55	-0.4674	-0.2373	1.1241	55	-0.4674	-0.1348	.8228
56	0.7468	0.8174	.5050	56	0.7468	1.0148	.6116
64	-0.1060	-0.5082	-1.3302	64	-0.1060	-0.3638	-1.1072
<b>66</b>	<b>0.6282</b>	<b>1.3341</b>	<b>2.9709</b>	<b>66</b>	<b>0.6282</b>	<b>1.3538</b>	<b>2.1076</b>
67	-0.3619	-0.2492	.6684	67	-0.3619	-0.1754	.3452
69	0.5626	0.5772	.2876	69	0.5626	0.7624	.3887
70	-0.8464	-1.1027	-.7639	70	-0.8464	-1.0064	-.7875
71	-0.2943	-0.5289	-.6797	71	-0.2943	-0.3431	-.4240
78	-1.2169	-1.5502	-1.0628	78	-1.2169	-1.5338	-1.3004
79	-1.4589	-1.0696	1.7420	79	-1.4589	-1.3355	.1389
80	-0.0118	0.2558	1.2696	80	-0.0118	0.2411	.5623
81	-0.1831	-0.2465	-.0151	81	-0.1831	-0.0560	.1510
3A	-1.8595	-1.6742	.9502	33F	-2.1504	-2.5904	-1.7029
30A	0.1734	0.2351	.4704	57F	0.2215	0.0466	-.8362
31A	1.6228	1.4083	-.6016	65F	0.6346	1.0172	.9863
35A	-0.9677	-1.1138	-.3361	68F	0.6901	1.2236	1.4796
48A	-0.2051	-.625	-1.3985	77F	-2.0300	-2.2628	-1.0255
53A	1.1443	1.087	.0085				
54A	-0.7839	-0.8629	-.0757				
57A	-1.7626	-1.7364	.3326				
63A	-0.2743	-0.5841	-.9715				
68A	-0.6898	-1.0365	-1.1148				

*Note.* The 2008 item sequence number was used to indicate that it was the same item appearing across years.

*Note.* Each item parameter was generated with a live, stratified random sample (i.e., about 3,000 cases) of the year.

*Note.* Item parameters of previous years were on the base scale.

*Note.* The 2008 items were independently calibrated with the 2008 stratified random sample.

**Form Statistics**

	Previous	2008	Previous	2008
Form Statistics	Base Form	Form A	Base Form	Form F
Mean	-.134	-.188	-.126	-.076
SD	.899	.937	.919	1.041

**Correlation and Standard Deviation Ratio**

	2008	2008
With Base Form	Form A	Form F
Correlation	.937	.954
SD Ratio	104%	113%

**Values Used for Robust Z Statistics**

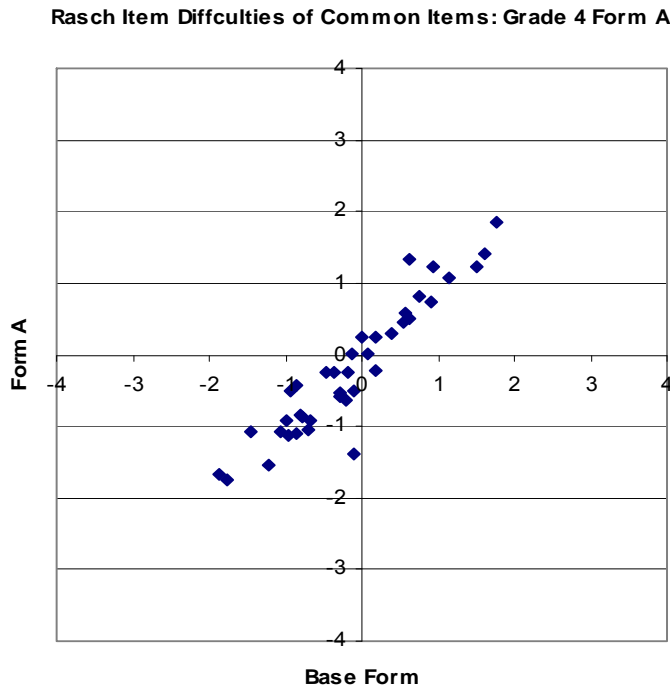
	2008	2008
With Base Form	Form A	Form F
Mean Diff	-.054	.050
Median Diff	-.060	.081
IQR Diff	.348	.413

Based on correlation coefficients, SD ratios, robust z values, and item difficulty plot, item number 47 and item number 66 appearing on both forms were dropped from the linking pool.

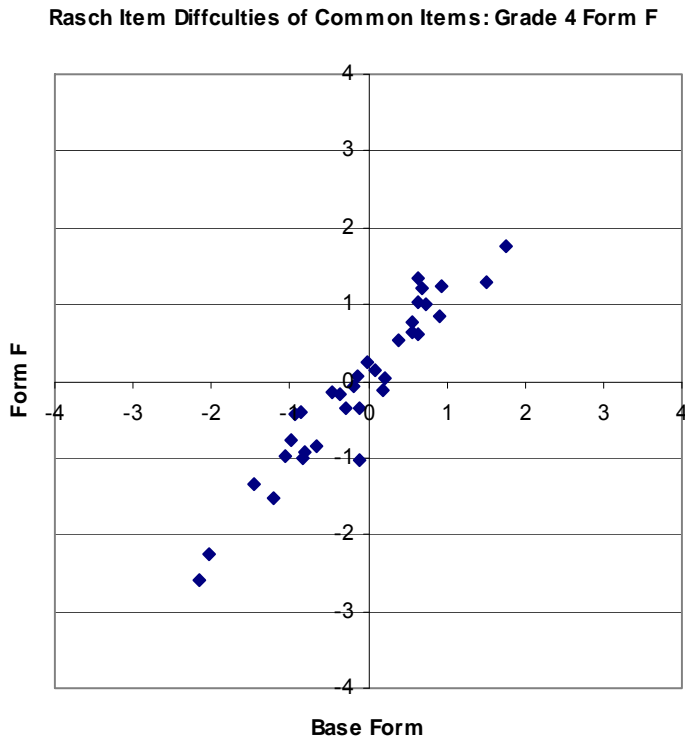
The following correlation coefficients and SD ratios were calculated after dropping those items:

	2008	2008
With Base Form	Form A	Form F
Correlation	.967	.971
SD Ratio	99%	110%





**Figure 1.17 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 4 Form A**



**Figure 1.18 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 4 Form F**



**Table 1.76 Rasch Item Difficulties and Robust Z values for Previous Year vs. Year 2008: Grade 5**

Item Seq No.	Previous Year	Y2008 Form A	Robust Z	Item Seq No.	Previous Year	Y2008 Form F	Robust Z
1	1.5795	1.3729	.0687	1	1.5795	1.4642	.9973
2	-1.0845	-1.5171	-1.0730	2	-1.0845	-1.7078	-1.7947
<b>8</b>	<b>-1.3086</b>	<b>-1.8697</b>	<b>-1.7222</b>	<b>8</b>	<b>-1.3086</b>	<b>-2.0437</b>	<b>-2.4092</b>
10	0.9885	0.7175	-.2566	10	0.9885	0.6722	-.1074
16	0.6094	0.5451	.7876	16	0.6094	0.4886	.9670
17	1.5483	1.3903	.3142	17	1.5483	1.3832	.7236
19	-0.9093	-1.247	-.5936	19	-0.9093	-1.2398	-.1855
21	0.4633	0.2138	-.1480	21	0.4633	0.1083	-.3201
23	0.335	0.0702	-.2253	23	0.335	-0.0192	-.3158
26	0.2030	-0.0926	-.3809	26	0.2030	-0.1482	-.2993
38	-0.3310	-0.6101	-.2975	38	-0.3310	-0.5541	.4048
39	-1.7042	-1.6243	1.5160	39	-1.7042	-1.9843	.0915
40	0.6218	0.41	.0424	40	0.6218	0.3084	-.0915
41	-1.2550	-1.4642	.0556	41	-1.2550	-1.6259	-.4075
43	-1.1293	-1.2141	.6840	43	-1.1293	-1.5220	-.5273
47	0.0148	-0.1031	.5168	47	0.0148	-0.1070	.9615
49	-0.6898	-1.2267	-1.5999	49	-0.6898	-1.3650	-2.0800
50	0.1746	0.2462	1.4741	50	0.1746	0.2196	1.8783
51	0.6342	0.3762	-.1910	51	0.6342	0.2927	-.2460
55	-0.6828	-0.9114	-.0424	55	-0.6828	-0.8574	.6713
56	0.2895	0.232	.8219	56	0.2895	0.3261	1.8321
59	-0.9892	-1.5528	-1.7348	59	-0.9892	-1.6263	-1.8706
60	-1.1516	-1.6428	-1.3690	60	-1.1516	-1.5510	-.5642
61	-0.5025	-.476	1.2448	61	-0.5025	-0.5900	1.1501
72	-0.5779	-0.6363	.8174	72	-0.5779	-0.7055	.9297
79	0.5383	0.3649	.2364	79	0.5383	0.4285	1.0275
83	-0.6839	-0.5865	1.6044	83	-0.6839	-0.7491	1.2726
20A	0.4459	0.3457	.6062	44F	1.6552	1.4963	.7576
27A	0.2606	-0.328	-1.8611	71F	-0.5862	-1.0184	-.7444
28A	1.0014	0.4695	-1.5746	<b>82F</b>	<b>-0.0098</b>	<b>0.2279</b>	<b>2.9374</b>
34A	0.8431	0.9251	1.5266				
37A	0.1548	0.0317	.4905				
69A	-1.6381	-2.0004	-.7179				
70A	-0.3862	-0.7863	-.9088				

*Note.* The 2008 item sequence number was used to indicate that it was the same item appearing across years.

*Note.* Each item parameter was generated with a live, stratified random sample (i.e., about 3,000 cases) of the year.

*Note.* Item parameters of previous years were on the base scale.

*Note.* The 2008 items were independently calibrated with the 2008 stratified random sample.

**Form Statistics**

Form Statistics	Previous Base Form	2008 Form A	Previous Base Form	2008 Form F
Mean	-.127	-.358	-.131	-.400
SD	.887	.941	.914	1.029

**Correlation and Standard Deviation Ratio**

	2008 Form A	2008 Form F
with Base Form		
Correlation	.978	.981
SD Ratio	106%	113%

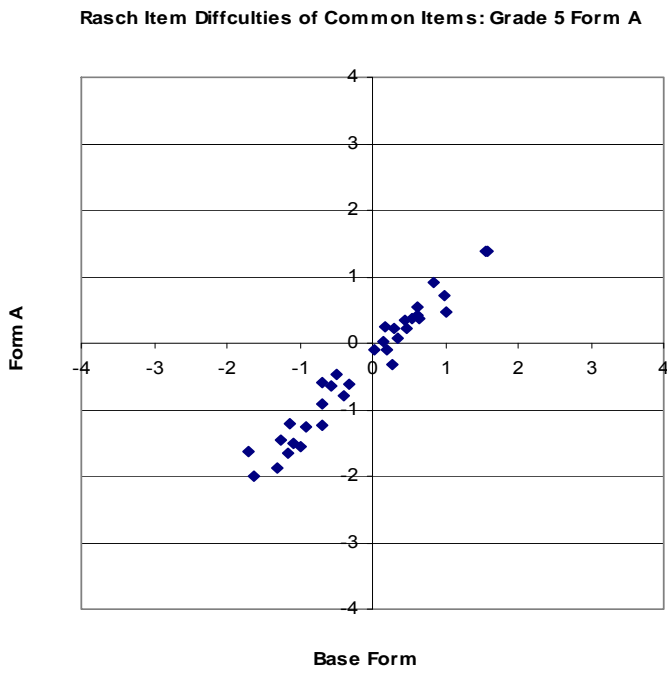
**Values Used for Robust Z Statistics**

	2008 Form A	2008 Form F
With Base Form		
Mean Diff	-.231	-.269
Median Diff	-.220	-.297
IQR Diff	.268	.246

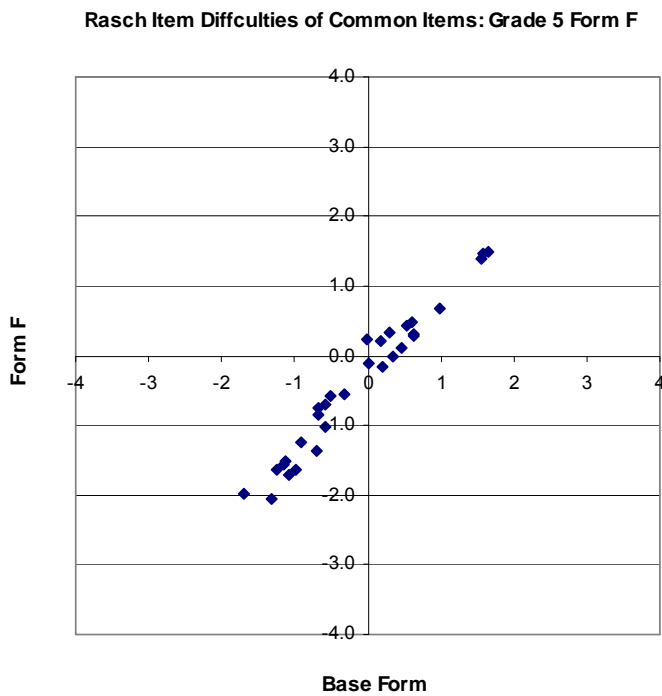
Item number 8 on both forms was dropped from the linking pool based on correlation coefficients, SR ratios, robust z values, and item difficulty plot. In addition, item number 82 appearing only on Form F was dropped from the linking pool.

The following correlation coefficients and SD ratios are based on dropping those items:

	2008 Form A	2008 Form F
With Base Form		
Correlation	.977	.986
SD Ratio	105%	110%



**Figure 1.19 Item Difficulty Plot Previous Year Form vs. Current Year (2008) Form: Grade 5 Form A**



**Figure 1.20 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 5 Form F**



**Table 1.77 Rasch Item Difficulties and Robust Z values for Previous Year vs. Year 2008: Grade 6**

Item Seq No.	Previous Year	Y2008 Form A	Robust Z	Item Seq No.	Previous Year	Y2008 Form F	Robust Z
3	0.6406	0.7247	0.5109	3	0.6406	0.8116	0.7992
5	0.1004	-0.1971	-0.7421	5	0.1004	-0.0962	-0.4821
6	-0.2844	-0.3424	0.0443	6	-0.2844	-0.4139	-0.2482
<b>9</b>	<b>-0.7278</b>	<b>-1.2812</b>	<b>-1.5824</b>	<b>9</b>	<b>-0.7278</b>	<b>-1.3886</b>	<b>-2.1001</b>
10	-1.4432	-2.0424	-1.7327	10	-1.4432	-2.0239	-1.8209
11	-0.4703	-0.8228	-0.9227	11	-0.4703	-0.5946	-0.2301
19	0.2409	0.0489	-0.3957	19	0.2409	0.0717	-0.3866
21	0.9203	0.9848	0.4466	21	0.9203	1.016	0.5368
25	0.3104	0.4554	0.7109	25	0.3104	0.4412	0.6591
26	-0.1396	-0.7605	-1.8040	26	-0.1396	-0.611	-1.4399
33	-0.8160	-0.7477	0.4590	33	-0.8160	-1.243	-1.2851
34	1.1378	0.9715	-0.3113	34	1.1378	1.0391	-0.1408
35	-1.4702	-1.3704	0.5625	35	-1.4702	-1.2929	0.8212
36	0.3674	0.2934	-0.0082	36	0.3674	0.3473	0.1332
37	0.5144	0.2905	-0.5004	37	0.5144	0.1308	-1.1339
38	-0.1849	-0.2603	-0.0128	38	-0.1849	-0.2668	-0.0823
44	0.4777	0.7114	1.0021	44	0.4777	0.7307	1.0851
46	-0.4091	-0.0928	1.2734	46	-0.4091	-0.1503	1.1053
49	-0.3209	-0.6227	-0.7562	49	-0.3209	-0.7805	-1.3988
51	1.2969	0.9484	-0.9096	51	1.2969	1.1211	-0.4096
<b>53</b>	<b>-0.7843</b>	<b>-0.2535</b>	<b>1.9777</b>	<b>53</b>	<b>-0.7843</b>	<b>0.0034</b>	<b>2.9488</b>
55	0.5885	.7390	0.7276	55	0.5885	0.8233	1.0216
56	0.1350	.2490	0.6101	56	0.1350	0.3402	0.9184
57	-0.4092	-.2690	0.6958	57	-0.4092	-0.0781	1.3573
62	0.6588	0.3472	-0.7884	62	0.6588	0.6026	0.0073
68	0.4071	0.5456	0.6896	68	0.4071	0.8175	1.6337
69	-1.3362	-1.1197	0.9457	69	-1.3362	-1.1645	0.8017
70	-1.8302	-1.4976	1.3269	70	-1.8302	-1.619	0.9394
80	0.6580	0.5226	-0.2098	80	0.6580	0.5976	-0.0073
1A	-1.2053	-1.5261	-0.8186	<b>4F</b>	<b>-1.2475</b>	<b>-1.9669</b>	<b>-2.3043</b>
12A	0.3254	0.2564	0.0082	45F	0.8786	0.8037	-0.0579
13A	1.7544	1.6555	-0.0900	<b>52F</b>	<b>-0.9336</b>	<b>-0.0217</b>	<b>3.3817</b>
52A	0.8563	0.9000	0.3783				
79A	-0.0894	-0.2357	-0.2456				

**Form Statistics**

Form Statistics	Previous Base Form	2008 Form A	Previous Base Form	2008 Form F
Mean	-.016	-.082	-.109	-.125
SD	.849	.871	.828	.915

**Correlation and Standard Deviation Ratio**

With Base Form	2008 Form A	2008 Form F
Correlation	.951	.913
SD Ratio	103%	111%

**Values Calculated for Robust Z Statistics**

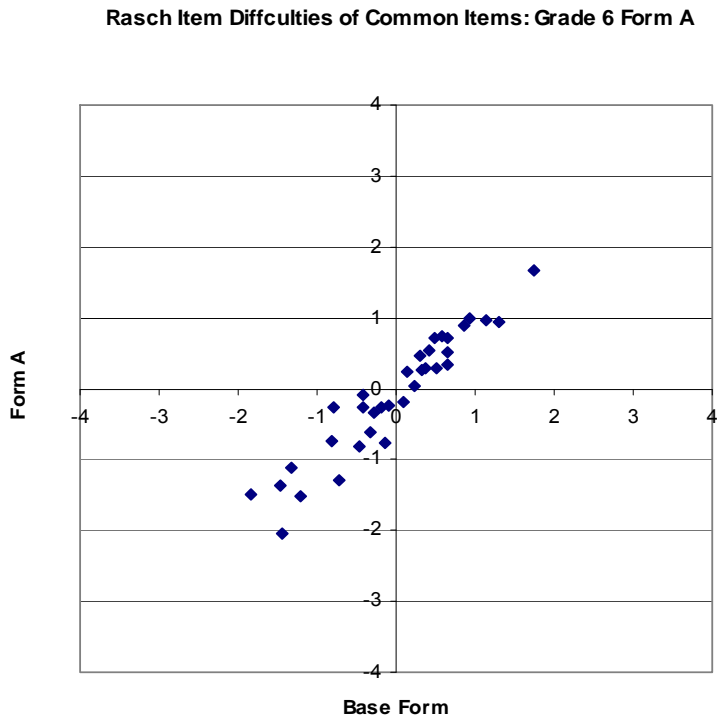
With Base Form	2008 Form A	2008 Form F
Mean Diff	-.067	-.017
Median Diff	-.072	-.058
IQR Diff	.412	.388

Based on correlation coefficients, SD ratios, robust z values, and item difficulty plot, item number 9 and item number 53 appearing on both forms were dropped from the linking pool. In addition, item number 4 and item number 52 appearing only on Form F were also dropped from the linking pool.

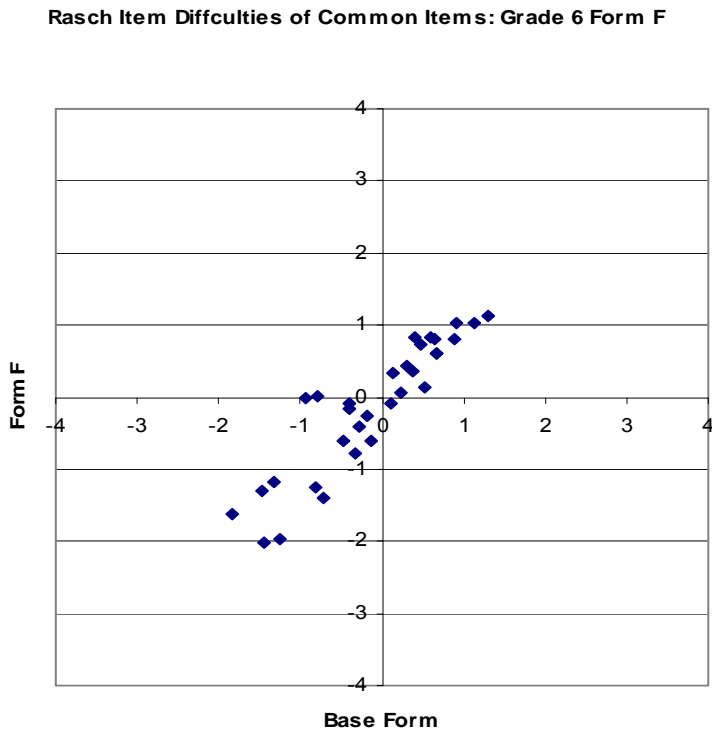
The following correlation coefficients and SD ratios were calculated after dropping those items:

With Base Form	2008 Form A	2008 Form F
Correlation	.960	.953
SD Ratio	102%	107%





**Figure 1.21 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 6 Form A**



**Figure 1.22 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 6 Form F**



**Table 1.78 Rasch Item Difficulties and Robust Z values for Previous Year vs. Year 2008: Grade 7**

Item Seq No.	Previous Year	Y2008 Form A	Robust Z	Item Seq No.	Previous Year	Y2008 Form F	Robust Z
1	1.0539	1.2292	1.7119	1	1.0539	1.4199	1.3700
2	0.4455	0.3246	0.0445	2	0.4455	0.4549	0.0599
3	0.1508	-0.0823	-0.5872	3	0.1508	-0.0543	-0.7282
4	-0.6420	-0.8084	-0.2117	4	-0.6420	-0.6828	-0.1246
7	-0.1398	-0.3921	-0.6952	7	-0.1398	-0.3344	-0.6896
8	-0.4706	-0.6741	-0.4205	8	-0.4706	-0.5354	-0.2127
9	0.3982	0.5276	1.4535	9	0.3982	0.6043	0.7825
10	-1.1551	-1.4001	-0.6541	10	-1.1551	-1.1752	-0.0485
12	-0.4683	-0.5263	0.3986	12	-0.4683	-0.1423	1.2230
18	-0.6359	-0.6882	0.4307	18	-0.6359	-0.9090	-0.9780
19	-1.1243	-1.6423	-2.1910	19	-1.1243	-1.4982	-1.3483
20	1.5825	1.3717	-0.4616	20	1.5825	1.5900	0.0529
27	0.9745	0.9564	0.6232	27	0.9745	1.2557	1.0585
30	-0.5147	-0.0711	3.2223	30	-0.5147	0.3714	3.2808
31	-2.6820	-2.8987	-0.4948	31	-2.6820	-2.8542	-0.6073
32	0.0227	-0.1885	-0.4639	32	0.0227	0.0290	0.0485
42	-0.6119	-0.7860	-0.2556	42	-0.6119	-0.9560	-1.2388
43	-0.4094	-0.4780	0.3389	43	-0.4094	-0.3685	0.1756
49	-1.6395	-1.7683	0.0000	49	-1.6395	-1.5025	0.5287
51	-0.0583	-0.0385	0.8365	51	-0.0583	0.0383	0.3802
52	-1.4991	-1.8734	-1.3820	52	-1.4991	-1.5797	-0.2708
64	0.0092	-0.0572	0.3513	64	0.0092	-0.1661	-0.6187
65	-0.4333	-0.6475	-0.4808	65	-0.4333	-0.4830	-0.1572
66	-0.2963	-0.7720	-1.9529	66	-0.2963	-0.7557	-1.6624
69	0.5231	0.1496	-1.3775	69	0.5231	0.4027	-0.4170
71	-0.7302	-0.6375	1.2469	71	-0.7302	-0.5134	0.8219
79	-1.4603	-1.0930	2.7928	79	-1.4603	-0.9985	1.7220
80	-0.5723	-0.3801	1.8071	80	-0.5723	-0.3812	0.7274
44A	0.6653	0.3420	-1.0949				
63A	0.5663	0.5768	0.7842				
72A	0.6673	0.7070	0.9486				

*Note.* The 2008 item sequence number was used to indicate that it was the same item appearing across years.

*Note.* Each item parameter was generated with a live, stratified random sample (i.e., about 3,000 cases) of the year.

*Note.* Item parameters of previous years were on the base scale.

*Note.* The 2008 items were independently calibrated with the 2008 stratified random sample.

**Form Statistics**

Form Statistics	Previous Base Form	2008 Form A	Previous Base Form	2008 Base Form
Mean	-.274	-.378	-.371	-.347
SD	.898	.942	.891	.965

**Correlation and Standard Deviation Ratio**

With Base Form	2008 Form A	2008 Form F
Correlation	.972	.956
SD Ratio	105%	108%

**Values Used for Robust Z Statistics**

With Base Form	2008 Form A	2008 Form F
Mean Diff	-.104	.024
Median Diff	-.129	-.007
IQR Diff	.240	.368

Based on correlation coefficients, SD ratios, robust z values, and item difficulty plot, none of the linking common items were dropped from the linking pool.

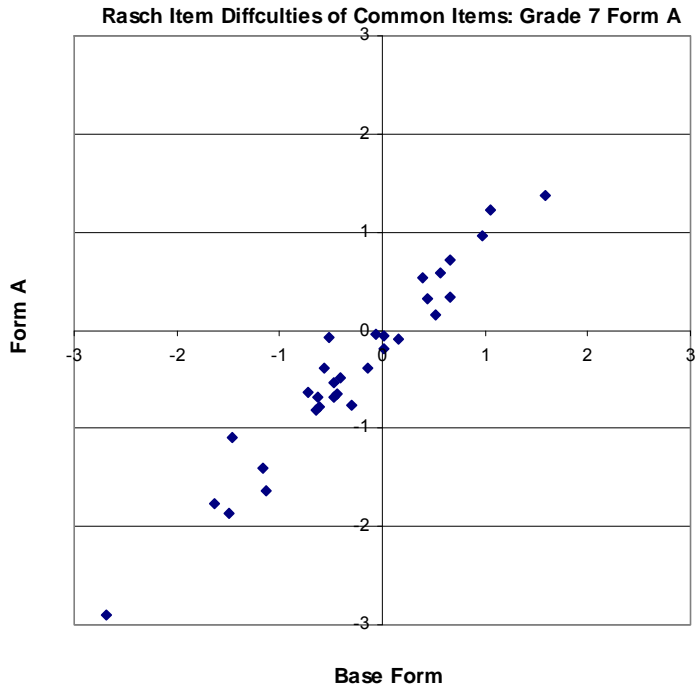


Figure 1.23 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 7 Form A

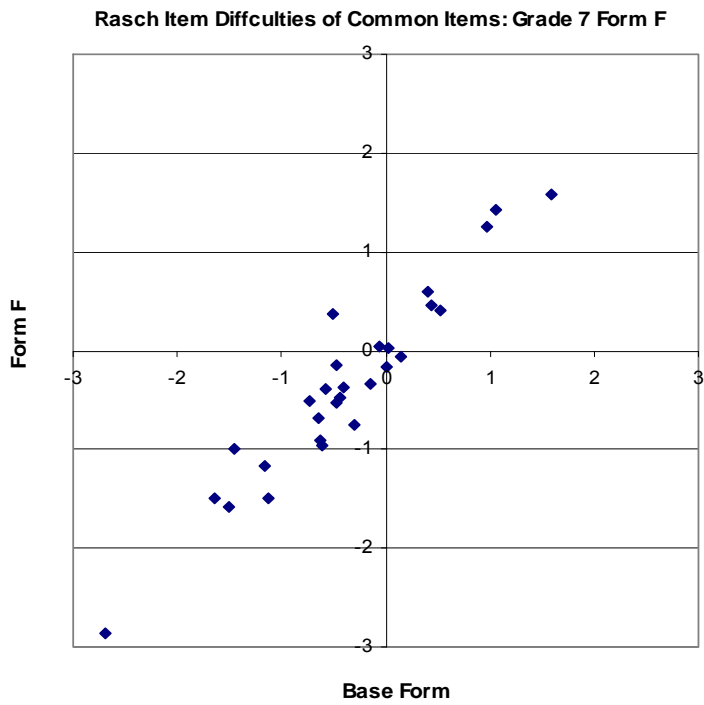


Figure 1.24 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 7 Form F



**Table 1.79 Rasch Item Difficulties and Robust Z values for Previous Year vs. Year 2008: Grade 8**

Item Seq No.	Previous Year	Y2008 Form A	Robust Z	Item Seq No.	Previous Year	Y2008 Form F	Robust Z
1	1.4965	1.1502	-1.1673	1	1.4965	1.4878	0.2847
2	-0.2177	-0.4902	-0.6306	2	-0.2177	-0.3137	-0.6109
5	-1.3613	-1.6955	-1.0793	5	-1.3613	-1.5966	-2.0400
7	-1.2003	-1.4260	-0.2902	7	-1.2003	-1.3021	-0.6704
22	-0.5815	-0.8619	-0.6880	22	-0.5815	-0.8328	-2.2041
32	1.0306	0.9439	0.7208	32	1.0306	1.0118	0.1811
33	0.5139	0.3190	-0.0662	33	0.5139	0.4341	-0.4447
38	-1.4001	-1.5560	0.2175	38	-1.4001	-1.4472	-0.1093
41	0.5661	-0.1043	-3.5246	41	0.5661	0.1477	-3.9184
42	-0.9380	-0.9479	1.2793	42	-0.9380	-0.9487	0.2642
43	-1.0563	-1.3020	-0.4357	43	-1.0563	-1.0821	0.1093
46	-0.2581	-0.4220	0.1600	46	-0.2581	-0.3181	-0.2416
47	-0.1085	-0.2010	0.6764	47	-0.1085	0.0413	1.9107
52	0.3257	0.4442	2.2132	52	0.3257	0.3892	1.0254
53	-0.6275	-0.8133	0.0000	53	-0.6275	-0.6853	-0.2190
58	0.2379	0.2491	1.4328	58	0.2379	0.5491	3.5665
64	1.2102	1.0693	0.3266	64	1.2102	1.2525	0.8079
67	-0.5330	-1.0868	-2.6765	67	-0.5330	-0.9848	-4.2610
79	-0.1424	-0.3414	-0.0960	79	-0.1424	-0.1102	0.7043
80	-1.3743	-1.1675	2.8554	80	-1.3743	-1.1695	2.4750
66A	1.8701	1.8947	1.5303	50F	-1.9767	-1.5105	5.1566
				65F	0.1391	0.0760	-0.2734

*Note.* The 2008 item sequence number was used to indicate that it was the same item appearing across years.

*Note.* Each item parameter was generated with a live, stratified random sample (i.e., about 3,000 cases) of the year.

*Note.* Item parameters of previous years were on the base scale.

*Note.* The 2008 items were independently calibrated with the 2008 stratified random sample.

**Form Statistics**

	Previous	2008	Previous	2008
Form Statistics	Base Form	Form A	Base Form	Form F
Mean	-.121	-.302	-.284	-.314
SD	.975	.996	.925	.918

**Correlation and Standard Deviation Ratio**

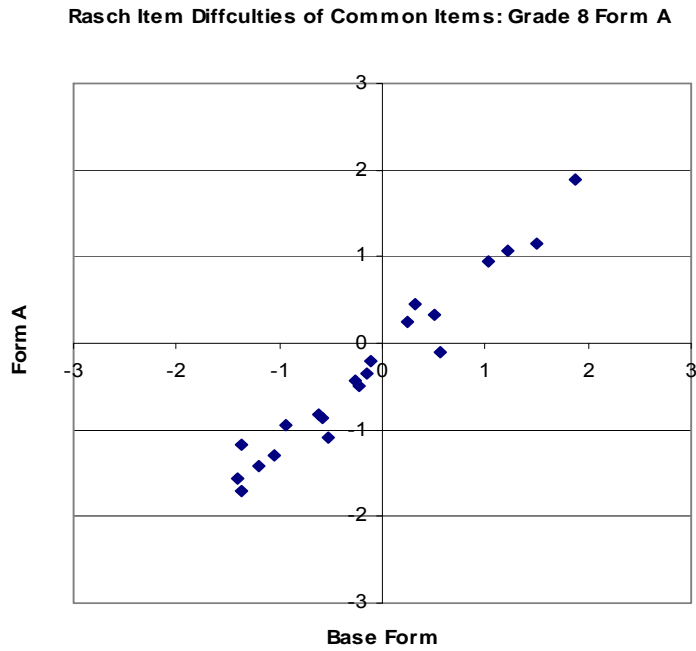
	2008	2008
With Base Form	Form A	Form F
Correlation	.979	.975
SD Ratio	102%	99%

**Values Used for Robust Z Statistics**

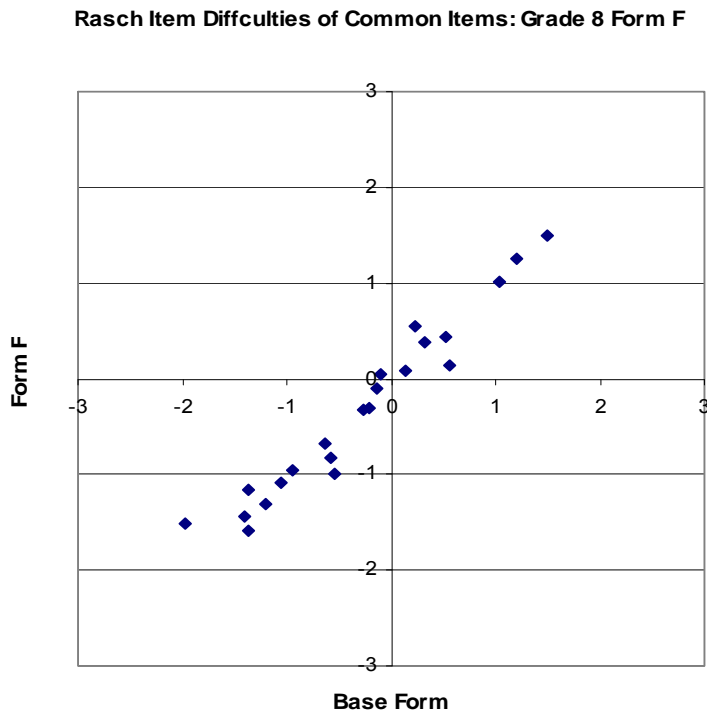
	2008	2008
With Base Form	Form A	Form F
Mean Diff	-.181	-.030
Median Diff	-.186	-.036
IQR Diff	.186	.132

Based on correlation coefficients, SD ratios, robust z values, and item difficulty plot, none of the linking common items were dropped from the linking pool.





**Figure 1.25 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 8 Form A**



**Figure 1.26 Item Difficulty Plot of Previous Year Form vs. Current Year (2008) Form: Grade 8 Form F**

### Reporting Scale Scores

In order to facilitate the use and interpretation of the results of the 2008 MSA-Math, the following formula was used to convert each student's ability or theta to the reporting scale score:

$$\text{ReportingAbilityScaleScore} = 32.8398 \cdot \text{theta} + 380.2954$$

$$\text{ReportingSE} = 32.8398 \cdot \text{SE}$$

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 2008 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 2008 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

## 1.11 Score Interpretation

To help provide appropriate interpretation of the 2008 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 1.10, *Linking, Equating, and Scaling Procedures*, the 2008 MSA-Math produced scale scores that ranged between 240 and 650. These scale scores have the same meaning within the same grade, but are not comparable across grade levels.

It should be noted that 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 2008 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.marylandpublicschools.org](http://www.marylandpublicschools.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 2008 MSA-Math score interpretation can be obtained from MSDE.

## 1.12 Test Validity of the 2008 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 2008 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 2008 MSA-Math blueprints provide extensive evidence regarding the alignment between the content in the 2008 MSA-Math and the VSC. It should be noted that the 2008 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 2008 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 2008 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 1.9, *Field Test Analyses*.

### ***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 2008 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 2008 Maryland item bank. More information on *DIF* analyses can be obtained in section 3.7, *Differential Item Functioning*.

### **Evidence from Internal Structure**

The 2008 MSA-Math has five math strands: *Algebra, Geometry and Measurement, Statistics and Probability, Numbers and Computations*, and *Process*. Tables 4.3 through 4.8 show the correlations among the math strands.

### 1.13 Unidimensionality Analyses of the 2008 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 2008 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 2008 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 2008 MSA-Math was met.

**Table 1.81 The 2008 MSA-Math Eigenvalues between the First and Second Components**

Grade	Form	Number of Items	First Eigenvalue	Second Eigenvalue
3	A	65	21.76	1.82
	F	65	22.36	1.95
4	A	64	22.74	1.99
	F	64	22.94	1.94
5	A	65	22.74	1.95
	F	65	23.27	1.94
6	A	62	23.99	1.74
	F	62	23.80	1.82
7	A	62	26.46	2.49
	F	62	27.10	2.27
8	A	62	25.18	2.17
	F	62	25.71	2.02

*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.

## 1.14 Item Bank Construction

The number of test forms to be constructed each year, and the need to replace items that would be released to the public, necessitated the availability of a large pool of items. The 2008 MSA-Math item bank continues to be maintained by Pearson in the form of computer files and paper copies. This enables test items to be readily available to both Pearson and MSDE staff for reference, test construction, test book design, and printing.

Pearson maintains a computerized statistical item bank to store supporting and identification information for each item. The information stored in this item bank for each item is as follows:

- CID
- Test administration year and season
- Test form
- Grade level
- Item type
- Item stem and options
- Passage code and title
- Subject code and description
- Process code and description
- Standard code and description
- Indicator code and description
- Objective code and description
- Item status
- Item statistics

It should be noted that each field test item of each form was calibrated by fixing each operational item with its operational Rasch items parameter. For example, all of the field test items of test forms A, B, C, D and E were independently calibrated after fixing the same items appearing on the five forms with the same operational item parameters, since each field test form belonged to the same operational form A. Item difficulties, step difficulties, and infit and outfit fit statistics of all the field test items were stored in the 2008 item bank.



## 1.15 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.