Maryland School Assessment-Reading:

Grades 3, 5, and 8

Technical Report:

2003 Administration

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The 2003 *Maryland School Assessment-Reading Technical Report* is the product of many individuals at Harcourt Assessment, Inc., the National Psychometric Council, and the Maryland State Department of Education. These acknowledgements recognize those individuals who have made significant contributions to it.

Daeryong Seo and Michael James Young wrote and produced the chapters of this report and provided the principal data analyses with Husein Taherbhai. This report was also supported by Carl Hyman, Linda Hayes, Todd Garrard, Angela Respress, Kimberly Irving, Larry Pedersen, Louis Franco, Jessika Mathews, and Norma Martinez.

The National Psychometric Council for the 2003 *MSA-Reading* reviewed the contents and analyses of this report. Special acknowledgement is given to MSDE staff members for their input and guidance: Gary Heath, Ray Scott, Janet Bagsby, Martin Kehe, and John Merrill.

INTRODUCTION

The *Maryland School Assessment (MSA)* is a measure of students' reading and mathematics. It will eventually include the measure of students' science. The *MSA* replaced the *Maryland School Performance Assessment Program (MSPAP)* to meet the new federal test requirements of the *No Child Left Behind Act (NCLB)* that was reauthorized and renamed from the *Elementary and Secondary Education Act* in 2002.

New academic standards were designed to inform parents, teachers, and educators of what students actually learned in schools and to make schools accountable for teaching contents measured by the *MSA*. To this end, the Maryland State Department of Education (MSDE), in collaboration with hundreds of educators across the state and Harcourt Assessment, Inc. (Harcourt), developed a series of reading tests to measure students' achievement against the new academic standards.

The purpose of the 2003 *MSA-Reading Technical Report* is to provide users and other interested parties with a general overview and statistical results of the 2003 *MSA-Reading*.

The 2003 *Technical Report* is composed of four parts, and the first part contains the following information:

- General overview and purposes of the 2003 *MSA-Reading*
- Development and review of the 2003 *MSA-Reading*
- Test administration
- Item selection for scoring purposes
- Linking, equating, and scaling
- Standard setting
- Score interpretation
- Test validity
- Item Bank

The second part provides the 2003 *MSA-Reading* results for students in grades 3, 5, and 8. It contains information about the cutoff score and pass rate at each performance level for the 2003 reading tests.

The third part contains statistical summaries for the 2003 *MSA-Reading*. This part outlines the statistical and psychometric characteristics of the *MSA-Reading*.

Three appendices provide additional statistical results for the 2003 *MSA-Reading*. Appendix A contains scale score frequency distributions and histograms, Appendix B contains both classical and *item response theory (IRT*) item parameters, and Appendix C contains test blueprints for grades 3, 5, and 8.

1. OVERVIEW OF THE 2003 *MARYLAND SCHOOL ASSESSMENT-READING*

In 2002, the Maryland State Department of Education (MSDE) took an important step toward raising learning expectations for all students in public schools. The State Board of Education retired the *Maryland School Performance Assessment Program (MSPAP)* and adopted a new testing program known as the *Maryland School Assessment* (*MSA*). The 2003 *MSA* 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.

Beginning in March 2003, students in grades 3, 5, and 8 took the *MSA* in reading (*MSA-Reading*) and mathematics. Students in grade 10 took only the *MSA-Reading*, because high school mathematics achievement was measured by the *Maryland High School Assessment* in geometry. In addition, tests in reading and mathematics will be phased in for students in grades 4, 6, and 7 in February 2004.

1.1 Overview of the 2003 *MSA-Reading*

As can be seen from Table 1.1, the 2003 *MSA-Reading* field tests were designed to provide two kinds of information. First, *norm-referenced* information was provided by the items from the abbreviated form of the *Stanford Achievement Test Series, Tenth Edition (SAT10)*. The *SAT10* consisted of *Word Study*, *Reading Vocabulary*, and *Reading Comprehension* items. To produce *criterion-referenced* information, additional items, called augmented items, were written for the *Maryland Reading Standards* (*MRS*) in grades 3, 5, and 8 and organized under three reading processes: *General Reading*, *Literary Reading*, and *Informational Reading*.

The 2003 *MSA-Reading* produced both norm-referenced and criterion-referenced scores for each student. While norm-referenced scores included only the *SAT10* items, both items selected from the *SAT10* and augmented items created for Maryland comprised criterion-referenced scores. Figure 1.1 shows a schematic of the *SAT10* and augmented items that produced these test scores.

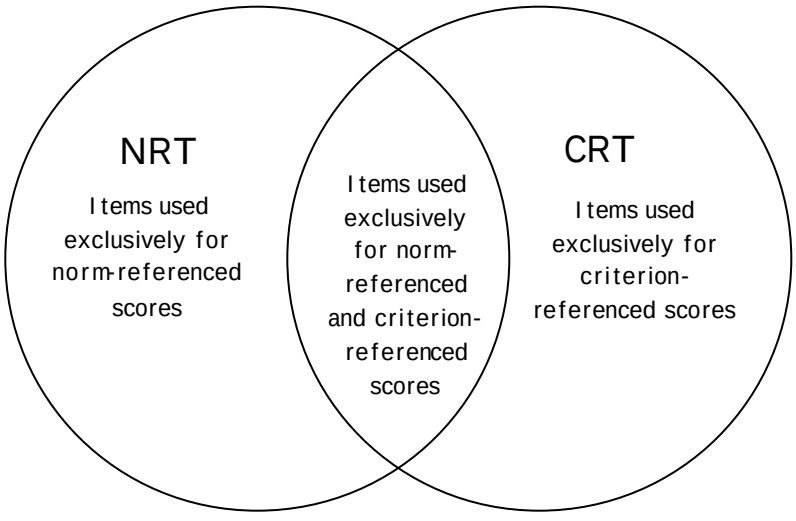


Figure 1.1 Schematic of the 2003 *MSA-Reading*

Table 1.1 The 2003 *MSA-Reading* Field Test Design: Grades 3, 5, and 8

Grade	Strand Title	*SAT10* / Augmented	Item Type	Number of Items	Total Points
3	Total NRT	*SAT10*	*SR*	70	70
	Word Study	*SAT10*	*SR*	20	20
	Reading Vocabulary	*SAT10*	*SR*	20	20
	Reading Comprehension	*SAT10*	*SR*	30	30
	Total CRT*	*SAT10*, Augmented	*SR*, *BCR*	45	57
	General Reading	*SAT10*	*SR*	15	15
	Literary Reading	*SAT10,* Augmented	*SR*, *BCR*	15	21
	Information Reading	*SAT10,* Augmented	*SR*, *BCR*	15	21
5	Total NRT	*SAT10*	*SR*	50	50
	Reading Vocabulary	*SAT10*	*SR*	20	20
	Reading Comprehension	*SAT10*	*SR*	30	30
	Total CRT*	*SAT10*, Augmented	*SR*, *BCR*	45	57
	General Reading	*SAT10*	*SR*	15	15
	Literary Reading	*SAT10,* Augmented	*SR*, *BCR*	15	21
	Informational Reading	*SAT10,* Augmented	*SR*, *BCR*	15	21
8	Total NRT	*SAT10*	*SR*	50	50
	Reading Vocabulary	*SAT10*	*SR*	20	20
	Reading Comprehension	*SAT10*	*SR*	30	30
	Total CRT*	*SAT10*, Augmented	*SR*, *BCR*	45	57
	General Reading	*SAT10*	*SR*	15	15
	Literary Reading	*SAT10,* Augmented	*SR*, *BCR*	15	21
	Informational Reading	*SAT10,* Augmented	*SR*, *BCR*	15	21

*Note: 1. CRT contains *SAT10* items
2. *SR* items are selected response items, and *BCR* items are brief constructed response items

1.2 Purposes/Uses of the 2003 *MSA-Reading*

By measuring students' achievement against the new academic standards, the 2003 *MSA-Reading* provides two main purposes. First, the *MSA-Reading* 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-Reading* serves as an accountability tool to measure performance levels of individual students, schools, and districts against the new academic standards.

1.3 The *Voluntary State Curriculum*

Federal law requires that states align their tests with their state content standards. The 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.

While 100% of the standards should be tested, it was not the case that every indicator would necessarily be tested each year. Consequently, the *VSC* specified curricular indicators and objectives that contributed directly to measuring content standards, which were aligned to the *Maryland School Assessment (MSA).*

1.4 Development and Review of the 2003 *MSA-Reading*

Developing the 2003 *MSA-Reading* was a complex process. It required a great deal of involvement from the MSDE, Harcourt, and local school systems. In addition, teachers, administrators, and content specialists from all over Maryland were recruited for different test development committees. These individuals reviewed test forms and items to ensure that they measured students' knowledge and skills fairly and without bias. Table 1.2 identifies which groups were responsible for developing the 2003 *MSA-Reading*.

National Psychometric Council

The National Psychometric Council (NPC) took a major role in reviewing and recommending to the MSDE on the development and implementation of the *MSA-Reading* program. For example, they made recommendations to the MSDE on issues, such as test blueprints, field test design, item analysis, item selection for scoring purposes, linking, equating, and scaling issues, standard setting, and other relevant statistical and psychometric issues. They recommended guidelines and accommodations for students with physical disabilities or limited English proficiency. The MSDE adopted their guidelines and recommendations.

Content Review Committee

During the item review process, the Content Review Committee members were briefed on the item review process. They ensured that the *MSA-Reading* was appropriately difficult and fair. Committee members were either specialists in reading 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 view 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 reading 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.

Table 1.2 The 2003 *MSA-Reading* Responsibility for the Test Development

Development of the 2003 *MSA-Reading*	Primary Responsibility
Development of Preliminary Blueprints and Item Specifications	Harcourt; MSDE; NPC
Development of Preliminary Brief Constructed Response Rubrics	MSDE
Item Writing	Harcourt
Item Review	Harcourt; MSDE; NPC; Content Review Committee
Bias Review	Harcourt; MSDE; Bias Review Committee
Construction of Field Test Forms	Harcourt; MSDE
Modification of Special Forms	Harcourt; MSDE
Review of Special Forms	MSDE
Pre-Field Test Training Workshops	Harcourt; MSDE; LEAs
Field Test Administrations	MSDE; LEAs
Construction of Operational Test Forms	Harcourt; MSDE; NPC
Review of Operational Test Forms	MSDE
Final Construction of Operational Test Forms	Harcourt; MSDE
Setting Standards for the 2003 *MSA-Reading*	Standard Setting Committee for the 2003 *MSA-Reading;* CTB/McGraw-Hill

1.5 Structure of the 2003 *MSA-Reading*

Structure of Field Test Forms

The 2003 *MSA-Reading* was composed of the *SAT10*, augmented, and field test items for future augmentation. The design of the *MSA-Reading* was to spiral a relatively large number of Maryland augmented field test items into multiple test forms for each grade in the 2003 test administration. Thus, the 2003 *MSA-Reading* produced 7 test forms. For forms 1 through 6, the order of the *SAT10* items was the same across the test forms, and Maryland-specific (augmented) items were located after the *SAT10* items. Form 7 was a re-ordered version of form 1 and was produced in order to examine the effects of item re-ordering. The descriptive statistics of each field test form can be found in section 1.8.

Types of Items

The 2003 *MSA-Reading* contains two types of items: *selected response* (*SR*) and *brief constructed response* (*BCR*) items. *SR* items required students to select a correct answer from several alternatives. For 2003 *MSA-Reading*, students selected an answer from four alternatives. Each *SR* item was scored as right or wrong.

BCR items required students to answer a question with a couple of words, a sentence, or a more elaborated way. For the 2003 *MSA-Reading*, these items were scored on a general rubric with maximum values between 0 and 3.

1.6 Test Administration

Test Administration Preparation and Materials

Pre-test workshops were held across the state prior to field test. These workshops provided the representatives of all the local school divisions with an overview of the tests' content, security expectations, and procedures for completing the answer documents. They also considered the receipt, distribution, and return of test materials.

For the test examiner, Harcourt provided the following materials:

- Test examiner's manual
- One set of pre-printed student ID labels and one set of generic ID labels
- Paper bands for used Answer Books
- Student roster

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

- Test Book
- Answer Book
- Two No.2 pencils with erasers (by school or student)

Two test-related manuals were developed for the administration of the 2003 *MSA-Reading*: Test Administration and Coordination Manual (TACM) and Examiner's Manual for Test Administration (EMTA). The TACM was developed and distributed by Harcourt. This manual provided Local Accountability Coordinators (LACs) and building level School Test

Coordinators (STCs) with information about the administration, packaging, and return of test materials. The TACM also described any issues specific to grades 3, 5, and 8. One TACM was produced for all administrations in grades 3, 5, and 8. The TACM was distributed one per school at the pre-test workshops and again included in the shipping materials.

The EMTA was developed for each grade by Harcourt and provided directions for administering the 2003 *MSA-Reading* at each grade level.

Test Administration Schedule

Specific dates were designated for each content area test. For the 2003 *MSA-Reading*, students were required to take Part I on Day 1 and Part II on Day 2, and the primary testing days were as follows:

- Test materials delivered to schools (including pre-print Student ID labels) — February 10-14, 2003
- Reading testing days — March 3 and 4, 2003
- Make-up testing days — March 7, and 10-14, 2003

Sessions were scheduled at any convenient time during the school day, but testing had to be scheduled to allow sufficient time to complete the test. Table 1.3 shows timing sessions allowed for the 2003 *MSA-Reading*.

Table 1. 3 The 2003 *MSA-Reading* Timing Sessions: Grades 3, 5, and 8

Grade	Form	Session				
		1	2	3	4	5
3	1-6	Q1-Q20	Q21-Q40	Q41-Q70	Q71-Q80	Q81-Q90
		12 min.	14 min.	30 min.	35 min.	35 min.
	7	Q1-Q20	Q21-Q40	Q41-Q55	Q56-Q65	Q66-Q90
		12 min.	14 min.	45 min.	20 min.	60 min.
5	1-6	Q1-Q20	Q21-Q50	Q51-Q60	Q61-Q70	
		14 min.	30 min.	35 min.	35 min.	
		Q1-Q20	Q21-Q35	Q36-Q50	Q51-Q60	Q61-Q70
		26 min.	40 min.	25 min.	35 min.	20 min.
8	1-6	Q1-Q20	Q21-Q50	Q51-Q60	Q61-Q70	
		14 min.	30 min.	35 min.	35 min.	
	7	Q1-Q20	Q21-Q40	Q41-Q50	Q51-Q60	Q61-Q70
		26 min.	50 min.	20 min.	35 min.	20 min.

If a student was absent on the testing days, a make-up test was administered on one of the subsequent days within the testing window (March 7 or March 10-14, 2003). 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 with LACs to determine the testing schedule to be followed. The LACs addressed all questions to the Assessment Branch in the Division of Planning, Results, and Information Management at the MSDE.

Any student who was tested during the make-up period was to continue to test in the original book. There was not a separate make-up book. Therefore, if a student unexpectedly left during a session, the remaining time was noted so that the student might finish that section with the proper amount of remaining time. If a known absence occurred during testing, splitting a session between original testing and make-up testing was avoided.

During the administration of the 2003 *MSA-Reading*, the MSDE had testing monitors in selected schools observing administration procedures and testing conditions. All monitors had identification cards for security purposes.

Testing Accommodations

Testing accommodations for Special Education students, English Language Learners (ELL), and students with disabilities covered under Section 504 had to be approved and documented according to the procedures and requirements outlined in the document entitled "Requirements for Accommodating, Excusing, and Exempting Students in Maryland Assessment Programs," as revised December 20, 2002. (A copy of the most recent edition of this document is available electronically on the LAC and STC web pages at http://docushare.msde.state.md.us)

No accommodations were 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 the MSDE upon request.

Braille and Large-Print Test Books

The 2003 *MSA-Reading* was administered to those requiring large-print Test Books and Answer Books or Braille Test Books. For both large-print and Braille Test Books, students' responses were transcribed into the regular Answer Books following testing. The pre-printed student ID labels were affixed to the regular Answer Books containing the transcribed responses.

Once the grades 3, 5, and 8 reading answers had been transcribed, large-print and Braille Test Books were returned for scoring with the regular materials. The Answer Books were bounded with a paper band, and the bundle was labeled "transcribed books."

Security of Test Materials

The Test Books and all used Answer Books for the 2003 *MSA-Reading* were confidential and kept secure at all times. Unauthorized use, duplication, or reproduction of any or all portions of the assessment was prohibited.

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

- Secure materials: Test Books and Answer Books
- Non-secure materials: Test Administration and Coordination Manual, Examiner's Manual for Test Administration, unused Answer Books, return address labels, pre-printed student ID labels, and instructions for applying ID labels

Distribution of Materials

Different test forms were administered to students in each classroom participating in reading tests, and each test form was identified by a cover of a different color and number. In addition, the Test Books and Answer Books were spiraled within a classroom.

1.7 Scoring Procedures

Students' responses to *SR* items were machine-scored, and their responses to *BCR* items were individually read and scored by Harcourt in San Antonio.

Once received by Harcourt, 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* 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 were assigned zero points. Students' responses with multiple marks and blank responses (omits) were also assigned zero points.

Hand-Scored Items

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.

Scorer Qualifications

BCR items were scored by scorers who were trained to stringent requirements and procedures. All applicants for *MSA* scorer positions were required to provide resumes and documentation of completed higher education. They were required to have earned a four-year college degree or higher. As part of the initial recruiting and screening process, applicants responded to a writing prompt and several content specific, open-response questions. The writing sample ensured that all applicants were fluent in writing and reading standard English. If successful on the preliminary screening, applicants participated in introductory workshops. The purpose of these workshops was to familiarize the applicants with general processes and procedures for scoring

performance assessments and to provide a final screening activity before they were added to the overall pool of potential scorers for the *MSA* project.

From that pool, potential scorers were assigned to the *MSA* project. *MSA*-specific training and qualifying consisted of having each scorer respond to actual *MSA* items or prompts prior to actual training. Using anchor papers and training sets, scorers then internalized the standards and the scoring scale for the item they were to score and were given qualifying sets. Those who met the qualifying standard were then allowed to score.

Methodology for Scoring the 2003 *MSA-Reading BCR* Items

For the *MSA*, each domain/level had a room director to direct scoring activities. The room director worked closely with the training supervisor and the content training specialist. The room director conducted training to ensure that scorers became experts in their scoring assignment. The main job of the room director was to oversee the actual scoring of the papers, acting as the decision maker for situations in which questions arise during the scoring process. The room director was also responsible for the quality of the scoring within the room. For the *MSA-Reading* program, those who served as room directors were usually active members of the training material development team, worked with MSDE staff and selected Maryland teachers to finalize scoring guides and training materials, and benchmarked student work.

For each item, scorers were trained to use the same scale to ensure accurate, consistent, and reliable scoring. All *BCR* items received a 0-3 score point range from two independent scorers. Equal or adjacent scores were acceptable. Readers were trained on and scored one item at a time. If the two readers did not assign equal or adjacent scores, the response was routed to a team leader for a third, independent reading to resolve the anomalous scores.

The read-behind application was also used to monitor reader performance. The team leader was provided a random selection of responses from each reader, distributed randomly across all readers. Although it could be tailored for each reader, by default, three percent of all responses scored appeared in the read-behind application. The team leader could agree with the scores and confirm them, disagree and send them back to the reader, or change them.

Training for Scoring Accuracy

The key to accurate scoring of *BCR* items is to train scorers appropriately. The following procedures were employed for training *MSA* project scorers.

Project-specific team leader training was conducted in the days immediately preceding scoring. Team leaders experienced in the scoring process helped train and retrain their team members. In addition, the logistics of the scoring sessions and the routines for resolution reading were discussed. All team leaders were also required to meet the qualifying standards set for the project. These standards were determined in conjunction with the MSDE.

Scorer training for *MSA* scoring began with an overview of the project and continued with the reading and discussion of selected student responses. The training utilized anchor sets, training sets, and qualifying sets, all of which contained MSDE reviewed and approved responses in addition to the *MSA* scoring rubric. Emphasis was placed on the scorer's understanding of how the responses differed from one another in quality and how each response represented the description of its score point as generalized in the scoring guidelines.

Inter-Rater Agreement

The scoring system generated many different kinds of internal monitoring reports that enabled accuracy of *MSA* scoring to be monitored. Teams produced the reports listing team scorers and providing the results of their scoring on an ongoing basis. Information on these reports included the number of responses read by the scorers during the period, the number and percentage of invalid responses (i.e., off-topic or blank responses, refusals to respond, responses in foreign languages), and the number of responses for which there was a subsequent reading. To illustrate, the number of responses with second reading provided data that allowed for reporting the number and percentage of responses with perfect agreement, the number and percentage of responses for which the first scorer was a point lower than the second scorer, the number and percentage of responses for which the first scorer was a point higher than the second scorer, and the number and percentage of responses differing by more than one score point.

In addition to the scorer reports described above, a daily order status report was generated each day to monitor the progress, logistically, of the overall scoring process through the system. This report was at the individual, team, and room levels, and showed, by order of completion and prompt, the number and percentage of responses for which first and second (check score) readings were required and completed for each item. These reports were available to team leaders, room directors, and training supervisors. They were also calculated and reported cumulatively for the day, the week, and the entire project. All reports were made available to the *MSA* supervisor every morning, and several of these monitoring reports could be called up online anytime throughout the scoring day. Statistical summaries of inter-rater reliability can be found in section 3.4.

1.8 Item Selection for Scoring Purposes

All items of the 2003 *MSA-Reading* were subjected to rigorous analyses for their properties. These analyses provided statistical information about test items that would be included as a part of scoring (operational) test forms. The following analyses were conducted:

- Overall statistical analyses for each field test form
- Classical item analyses for *SR* and *BCR* items
- *Differential item functioning* (*DIF*)
- "Not-reached" item analyses

Descriptive Statistical Analyses for Each Field Test Form

To ascertain whether or not each field test showed statistical abnormalities, descriptive statistics, such as means, standard deviations, standard errors, and reliability were used with the common items of each test form. These analyses also provided statistical information about determining whether later calibration and equating were successful for field test forms 1 through 7. As can be seen from Table 1.4, there are no significant differences across six test forms except form 7.

A separate check on test forms 1 and 7 to determine if the test form with the mixed item order (i.e., form 7) exhibited the same properties as the form with the original item order (i.e., form 1) indicated that there were some differences between form 1 and form 7. The "omits" in form 7 dominated the discussions with respect to the possible reason for higher mean of form 7. In addition, different session times allowed for form 7 gave some answers about the different means of these two forms.

Table 1.4 The 2003 *MSA-Reading* Descriptive Statistics for Each Field Test Form (Common Items)

Grade	Form	Number of Items	N	Mean	SD	Reliability	SEM
3	1	25	8,298	15.43	5.07	0.83	2.09
	2	25	8,615	15.58	5.06	0.83	2.09
	3	25	8,562	15.59	5.08	0.83	2.09
	4	25	8,466	15.59	5.11	0.83	2.11
	5	25	8,597	15.55	5.11	0.84	2.04
	6	25	8,291	15.54	5.17	0.84	2.07
	7	25	7,619	16.11	4.83	0.82	2.05
5	1	25	8,414	16.49	5.20	0.84	2.08
	2	25	8,888	16.51	5.23	0.84	2.09
	3	25	8,874	16.56	5.20	0.84	2.08
	4	25	8,833	16.54	5.24	0.84	2.10
	5	25	8,757	16.56	5.24	0.84	2.10
	6	25	8,488	16.55	5.19	0.84	2.08
	7	25	7,947	17.71	4.60	0.81	2.01
8	1	22	8,326	15.17	4.37	0.80	1.95
	2	22	8,799	15.02	4.54	0.82	1.93
	3	22	8,824	15.17	4.44	0.81	1.94
	4	22	8,874	14.99	4.59	0.82	1.95
	5	22	8,787	15.02	4.49	0.81	1.96
	6	22	8,604	15.20	4.35	0.80	1.95
	7	22	4,013	16.13	3.86	0.76	1.89

Classical Item Analyses for *SR* and *BCR* items

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

SR items for further scrutiny were flagged if:

- An item distractor was unselected by all students (i.e., nonfunctional distractor), or selected by a large number of high ability students, with low selection from other ability groupings (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.

BCR items for further scrutiny were flagged if:

- An item did not elicit the full range of rubric scores.
- The ratio of mean item score to maximum score was less than .20 or greater than .90.
- An item-total correlation was less than .10.

Dropping any items needed 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 dropping. If the item represented important content that had not been extensively taught, however, it would be justified to retain the item.

Differential Item Functioning Analyses

Differential item functioning (*DIF*) analyses are primarily designed to detect differential item performance across subgroups of a population while controlling for ability.

For the 2003 *MSA-Reading DIF* analyses, the reference group was either male or Caucasian students, and the focal group was either female or African-American students. Because the 2003 *MSA-Reading* included both the *SAT10* items and the "Maryland-specific" items on each field test form, the total item score on a collection of items was used as the matching variable.

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

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

For further information about the *DIF* procedures used for the 2003 *MSA-Reading*, please see section 3.7.

"Not-Reached" Item Analyses

An important consideration in the *item response theory* (*IRT*) analyses employed for the 2003 *MSA-Reading* was the treatment of missing responses to test items. Specifically, these procedures drew a distinction between items that were considered to be intentionally missing (*omits*) and missing items that occurred at the end of a block of items (*not-reached*). Researchers have suggested that "ignoring not-reached items introduces slight biases into item parameter estimates when not-reached items are present and speed is correlated with ability (Allen, Donoghue, & Schoeps, 2001, p. 232)."

For the 2003 *MSA-Reading*, this analysis was performed for each session within each day of the assessment. "Not-reached" items were treated as missing values for the purposes of calibration. In addition, "omit" items were scored as wrong for *SR* items and scored in the lowest category for *BCR* items. In addition, if the "not-reached" rate for an item exceeds 10% (i.e., an item completion rate of less than 90%), additional discussion was required with the MSDE to decide whether or not the item should be dropped to alleviate test "speededness".

Item Selection for Scoring Purposes and 2004 Operational Forms

The selection of items to be included in the final test forms of the 2003 *MSA-Reading* required a careful consideration based on test blueprints, classical item analyses, and *DIF* analyses, and "not-reached" item analyses. Harcourt suggested the following guidelines to choose items included in the final test forms:

- Avoid the use of the items with *p*-values less than .20 and greater than .90.
- Avoid the use of the *BCR* items with score distributions that do not elicit the full range of rubric scores.
- Avoid the use of items with point-biserial or item-total correlation less than .10.
- Avoid the inclusion of 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 reading experts.

In applying these guidelines, a balance should be made between being too harsh and thus dropping items that may affect the content representativeness of the entire set of field test items and being too lenient and allowing items with poor model fit that might affect resulting measures. In addition, reading specialists from the MSDE reviewed the final test forms of the 2003 *MSA-Reading*.

For the 2004 *MSA-Reading*, four operational test forms were constructed and reviewed by reading specialists from the MSDE. They determined the content validity and equivalency of the test forms for each grade level.

1.9 Linking, Equating, and Scaling Procedures

Linking Procedures

To link different test forms at each grade level, linking steps recommended by the National Psychometric Council were taken into consideration. For the 2003 *MSA-Reading*, items that appeared on each test form were included as potential linking items, but only *SR* items were considered as potential linking items.

First, the following calculation were made (SDE, 2001):

- The mean and standard deviation of the linking pool's item difficulties of each form
- The ratio of the standard deviations between form 1 and the rest of the forms
- The correlation between test form 1 and other test form item difficulties
- The difference between test form 1 and other test form item difficulties for each item in the linking pool
- The mean of the differences calculated above
- The median of the differences
- The interquartile range of the differences
- The robust Z for each item in the linking pool where 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).

Once the above calculations were made, the following guidelines were taken in determining possible sets of linking items to be used for the Rasch equating (SDE, 2001):

- Do not include those items with an absolute value of robust Z exceeding 1.645. In addition, if one difficulty or step from a *SR* item is eliminated from the pool based on robust Z, all other difficulties are also removed.
- Do not eliminate more than 20 percent of the pool linking items.
- Consider that the ratio of the standard deviations of the test form 1 and other test form item difficulties should be in the 90 to 110 percent range.
- It is assumed that the correlation of the test form 1 and other test form item difficulties is greater than .95.

Toward this end, Harcourt provided Rasch item difficulty lots and identified items that were to be deleted based on the robust Z statistics. Figure 1.2 provides the Rasch item difficulty of each item for each of the two forms and the robust Z calculated by the definition. The item difficulty plot between form 1 and form 2 indicates that there exist no extreme outliers. The correlation coefficient of the two test forms, 1.00, also indicates a very strong relation between the item parameter estimates of the two test forms.

Item Number	Form 1	Form 2	1 vs 2	Robust Z
1	-2.33	-2.31	0.02	1.577
2	-1.10	-1.19	-0.09	-.901
3	.15	.11	-0.04	.225
4	-.93	-.94	-0.01	.901
5	.93	.84	-0.09	-.901
6	-1.08	-1.07	0.01	1.351
7	.45	.41	-0.04	.225
8	2.29	2.20	-0.09	-.901
9	-.08	-.17	-0.09	-.901
10	1.00	.90	-0.10	-1.126
11	-.11	-.16	-0.05	.000
12	.24	.17	-0.07	-.450
13	-.15	-.23	-0.08	-.676
14	-.56	-.60	-0.04	.225
15	-.03	-.11	-0.08	-.676
16	-1.80	-1.80	0.00	1.126
17	.00	-.03	-0.03	.450
18	.03	-.06	-0.09	-.901
19	.93	.82	-0.11	-1.351
20	-1.06	-1.04	0.02	1.577
21	-1.43	-1.48	-0.05	.000
22	-.91	-.93	-0.02	.676
23	.65	.61	-0.04	.225
24	-.40	-.48	-0.08	-.676
25	.59	.50	-0.09	-.901

	Form 1	Form 2
Mean	-.188	-.242
SD	1.012	.985

	1 vs 1	1 vs 2
Correlation	1.000	1.000
SD ratio	100%	97%

	1 vs 1	1 vs 2
Mean of Difference	.000	-.053
Median of Difference	.000	-.050
Interquartile Range of Difference	.000	.060

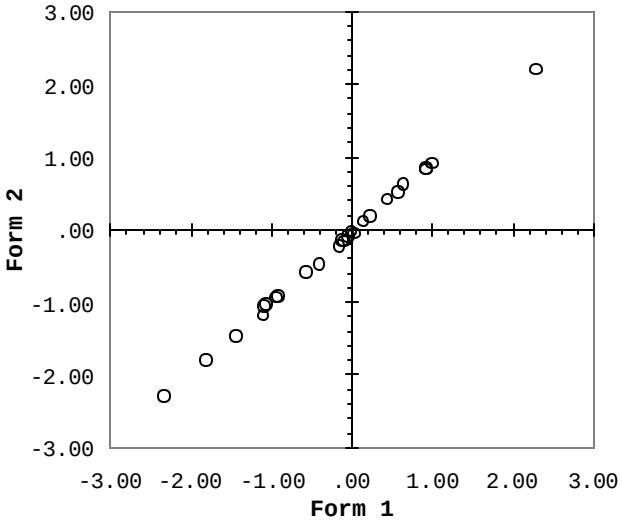


Figure 1.2 Example of Parameters Used to Link Items

Equating Procedures

Equating different test forms ensures that students taking one form of a test are neither advantaged nor disadvantaged when compared to students taking a different form of a test.

For the 2003 *MSA-Reading*, items selected through the linking procedures were used to equate all different test forms of each grade. Because each test form included a subset of unique items, linking items served as anchor items. Thus, whenever a new test form is constructed in the future, the new form will be equal in difficulty to the previous form via linking items. The design to collect data for the 2003 *MSA-Reading* was common item, non-equivalent groups.

In order to obtain parameter estimates for both the unique items on each form and the linking items, the Rasch model (or Partial Credit Model for *BCR* items) was used. For the 2003 *MSA-Reading*, the common items whose calibrations were known were anchored or fixed to their known estimates during calibration of other forms that were to be put on the scale of the first form. In treating these common item parameters as known they were fixed, and the remaining item parameters (for the unique items of each form) were also forced onto the same scale as the anchored (fixed) items.

The final step consisted of obtaining ability score or theta for each raw score point on a form. This was done by iteratively solving the expression:

$$True\ Score = \sum_{i=1}^{I} \sum_{j=0}^{m_i} j \cdot P_{ij}(q)$$

where

$P_{ij}(q)$ = the probability of a correct response for each of the $i = 1, \ldots, I$ items given that the item categories are numbered $0, \ldots, m_i$.

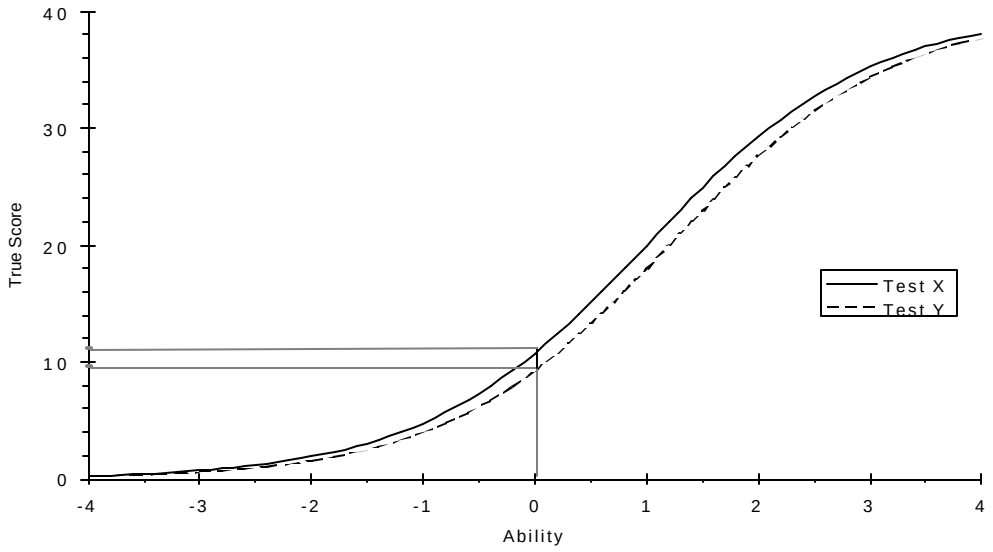


Figure 1.3 True Score Equating

Figure 1.3 illustrates these ideas for two hypothetical test forms, X and Y. In the figure, the true scores on each of the forms are plotted against ability using the true score equation. By drawing a line from the ability (here shown for an ability of 0) to each of the respective curves and moving across to the true score scale, one can find the pairs of true scores that are equated to one another. According to Lord and Wingersky (1984), the procedure applied to true scores can be transferred to observed scores without any major anomalies in the resulting outcomes.

Reporting Scale Scores

In order to facilitate the use and interpretation of the results of the 2003 *MSA-Reading*, scale scores were created based on the information given by both the MSDE and the NPC. The following is the formula to covert each student' ability or theta to scale scores:

$$ReportingAbilityScaleScore = 40 \cdot theta + 400$$

$$ReportingSEM = 40 \cdot SEM$$

where

theta = the IRT ability estimate, and

SEM = the conditional *SEM* of the ability estimate.

Note that the minimum of the scale was set to 0 and the maximum of the scale was set to 800.

1.10 Standard Setting

Crocker and Algina (1986, p. 410) pointed out, "(m)any situations require the setting of cutoff scores before test performance is interpreted. ... The practice of setting cutoff scores is commonly called *standard setting*."

For the 2003 *MSA-Reading*, the *Bookmark procedure* was used to set cutoff scores. CTB/McGraw-Hill and standard setting committees set two cutoff scores for grades 3, 5 and 8, and the following performance categories were created:

- *Advanced*
- *Proficient*
- *Basic*

Integrating *SR* and *BCR* items

During the *Bookmark procedure*, the key material presented to committee members was the ordered item booklet, in which items were ordered by their scale locations as determined by *item response theory* (*IRT*) calibrations (Mitzel, Lewis, Patz, & Green, 2001).

It has been recognized that standard setting methods traditionally work better on one or the other of the item types, but not on both. Thus, large-scale assessment programs have applied one standard setting procedure to the *SR* items and another to the *BCR* items. However, this creates additional steps of resolving different results from different methods, and potentially raises questions around the validity of final cut scores if those methods produce highly different outcomes (Mitzel, Lewis, Patz, & Green, 2001).

These days, *IRT* methodology is applied to large-scale assessment programs that include both *SR* and *BCR* item types to define a single underlying trait. From the perspective of the *Bookmark procedure* based on *IRT*, the method of setting performance standards should also reflect the unity of the underlying content if both *SR* and *BCR* item contents are calibrated to establish a single trait. Thus, scaling these two different item types together allows both item types to be placed into a single ordered item booklet and to be considered jointly by panelists (Mitzel, Lewis, Patz, & Green, 2001).

In dividing items at a cutoff score between mastery and nonmastery, a response probability (RP) of .67 (i.e., 2/3) is applied during standard setting. This implication means that for a given cutoff score, a student with a test score at that point will have a .67 probability of answering an item correctly at that cutoff score, and this is the technical definition of mastery.

SR item is mapped where the examinee has the probability of 2/3 of a correct answer. For *BCR* items with more than two categories, the *Bookmark procedure* follows the NAEP method of creating pseudo-binary items to map the non-zero categories. Consider a *BCR* item with score categories 0, 1, 2, and 3. The first pseudo-binary item is created by keeping the category 0 as zero and recoding the categories 1, 2, and 3 as one. This binary item then is used to place the score category 1 of the *BCR* item. The second pseudo-binary item is created by recoding the categories 0 and 1 as zero and the category 2 and 3 as one. This binary item is then used to map the category 2 of the *BCR* item. Finally, the third pseudo-binary item is created by recoding the categories 0, 1, and 2 as zero and the category 3 as one. This binary item is then used to map the category 3 of the CR item (Huynh, Meyer III, & Barton, 2000).

For the *2003 MSA-Reading* standard setting, a *BCR* item was placed in the booklet at three locations according to scale scores associated with attaining each additional score point. Scoring rubrics were also placed after each *BCR* item to help participants determine the skills and knowledge required to attain a given score point (Mitzel, Lewis, Patz, & Green, 2001).

Further information about the standard setting can be obtained from the MSDE or the *Maryland Standard Setting Technical Report* of CTB/McGraw-Hill (2003, August).

1.11 Score Interpretation

To help provide appropriate interpretation of the 2003 *MSA-Reading* test scores, two types of scores were created: 0-800 scale scores, and performance levels and descriptions.

0-800 Scale Scores

As explained in section 1.9, Linking, Equating, and Scaling, the 2003 *MSA-Reading* produced scale scores that ranged between 0 and 800. Those scale scores have the same meaning within the same grade, but those scores are not comparable across grade levels.

It should be noted that those scale scores have only simple meaning that higher scale scores represent higher performance in reading 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 Levels and Descriptions

As previously explained, performance levels and descriptions provide specific information about students' performance levels and help interpret the 2003 *MSA-Reading* scale scores. They describe what students at a particular level generally know and can be applicable to all students within each grade level. As Table 2.1 shows a range of scale scores at each performance level. For example, grade 3 reading scale scores from 404 to 487 indicate the level of *Proficient*, and students at this level can read grade appropriate text and demonstrate the ability to comprehend literature and informational passages. Further information about the 2003 *MSA-Reading* score interpretation can be obtained from the MSDE.

1.12 Test Validity

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 on-going and independent processes that are essentially independent 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 validity evidence of the 2003 *MSA-Reading*, content-related evidence, evidence of internal structure, and evidence of unidimensionality 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 of the test (Messick, 1989).

The 2003 *MSA-Reading* blueprints provide extensive evidence regarding the alignment between the content of the 2003 *MSA-Reading* and the *VSC*. These blueprints are presented in Appendix C.

Evidence of the Internal Structure of the *MSA-Reading*

The 2003 *MSA-Reading* has three reading processes: *General Reading*, *Literary Reading*, and *Informational Reading*. As can be seen from Tables 4.3 through 4.5, there exist moderately strong intercorrelations among these three processes.

Evidence of Unidimensionality

Measurement implies order and magnitude on a single dimension (Andrich, 1989). Consequently, in the case of scholastic achievement, this requires a linear scale to reflect this idea of measurement. Such a test is considered to be unidimensional (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 2003 *MSA-Reading*, polychoric correlation coefficients were computed with *LISREL 8.5* (Jöreskog & Sörbom, 1993) because they were polytomously scored on reading tests. Principal component analysis was then applied to produce eigenvalues. The first and the second principal component eigenvalues were compared without rotation. Table 1.5 summarizes the results of the first and the second principal component eigenvalues of the 2003 *MSA-Reading*.

The rule of thumb to determine the unidimensionality of a test requires that the eigenvalue of the first component or factor should be at least three times larger than the second one. As can be seen, the size of the eigenvalue of the first component meets the criterion for the unidimensionality. Thus, the assumption of unidimensionality for the 2003 *MSA-Reading* was met.

Table 1.5 The 2003 *MSA-Reading* Eigenvalues between the First and the Second Components

Grade	Form	Number of Items	First Eigenvalue	Second Eigenvalue
3	1	37	11.83	1.59
	2	37	11.53	1.57
	3	37	11.74	1.67
	4	37	12.04	1.53
	5	37	12.50	1.52
	6	37	12.86	1.53
	7	37	11.61	1.65
5	1	37	11.67	1.47
	2	37	11.80	1.48
	3	37	10.98	1.48
	4	37	11.22	1.50
	5	37	11.71	1.50
	6	37	11.65	1.55
	7	37	11.19	1.36
8	1	34	10.94	1.54
	2	34	11.44	1.45
	3	34	11.24	1.59
	4	34	10.86	1.52
	5	34	11.20	1.49
	6	34	10.38	1.49
	7	34	10.22	1.64

1.13 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 2003 *MSA-Reading* item bank continues to be maintained by Harcourt as computer files and paper copies. This enables test items to be readily available to both Harcourt and MSDE staff for reference, test construction, test book design, and printing.

Harcourt maintains a computerized statistical item bank to store supporting and identification information on 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

The item bank Rasch scale statistics were re-calibrated using all of the students' test responses. Thus, the re-calibrated scale would serve as the base scale.

2. CURRENT RESULTS OF THE 2003 *MSA-READING*

This section provides information about the 2003 *MSA-Reading* results for students in grades 3, 5, and 8. Table 2.1 contains information about the cutoff score of each performance level, and Table 2.2 contains the pass rate of each performance level based on the cutoff score.

Table 2.1 The 2003 *MSA-Reading* Cut Scores: Grades 3, 5, and 8

Grade	Cut Score of Performance Level	
	Proficient	Advanced
3	404	488
5	405	455
8	419	463

Table 2.2 The 2003 *MSA-Reading* Pass Rates: Grade 3, 5 and 8

Grade	N	Percentage of Performance Level		
		Basic	Proficient	Advanced
3	64,471	39.97	51.33	8.70
5	67,722	34.02	39.89	26.10
8	67,706	39.18	34.84	25.98

Note: Percentages may not add to 100% due to rounding

3. OVERVIEW OF STATISTICAL SUMMARIES

This section provides general information about statistical and psychometric summaries used for the 2003 *MSA-Reading* program. Actual statistical results described in this section appear in section 4 and Appendices.

3.1 Classical Descriptive Statistics

Table 4.1 contains the classical descriptive statistics of each form for each grade and includes:

- Form number
- Number of items
- Numbers of students[1]
- Means and standard deviations of raw scores
- *Stratified Cronbach Alpha*
- *Standard error of measurement* (*SEM*)

Stratified Cronbach Alpha

The 2003 *MSA-Reading* tests included both *SR* and *BCR* items. Consequently, it was asked to use an adequate reliability coefficient that addressed the important factor, different item type. The following formula depicts the reliability coefficient, *Stratified Cronbach Alpha*:

$$Stratified\ a = 1 - \frac{\sum s_i^2 (1 - r_{ii'})}{s_t^2}$$

where

s_i^2 = variance of score on cluster *i*,

s_t^2 = variance of total score, and

$r_{ii'}$ = reliability coefficient of score on cluster *i*.

These tests were initially considered to be *classically congeneric* (i.e., besides having unequal means and unequal variances in error and observed scores, the test forms also have heterogeneity of true-score variances) where the tasks within the examinations were stratified based on the type of item (i.e., multiple-choice, short answers, extended responses, and extended writing) and by the scoring rubric attached to these items.

Upon examining the variance/covariance matrices, however, it became apparent that in some cases the part covariance of a part was not heterogeneous with respect to other part variances (e.g., the covariance between multiple-choice items and extended responses and between

[1] Note the numbers of students reported in these tables may be lower than the totals reported in the statewide summaries. These analyses were based on the sample of data used to equate the forms of the 2003 *MSA-Reading*.

multiple-choice items and extended writing for grade 3 writing). It was, therefore, determined that although the test forms were *congenerically parallel*, they were not *classically congeneric* (Qualls, 1995). For the 2003 *MSA-Reading*, therefore, the test forms were divided into two strata made up of *SR* and *BCR* items, and the *Stratified Cronbach Alpha* was used as the reliability coefficient.

Standard Error of Measurement (Based on Classical Test Theory)

The *standard error of measurement* (*SEM*) is the standard deviation of errors of measurement that are associated with test scores from a particular group of examinees. In here, a measurement error is the difference between an examinee's actual or obtained score and the theoretical true score counterparts. Consequently, the *SEM* is commonly used in interpreting and reporting individual test scores and score differences on tests (Harvill, 1991).

Classical test theory is based on the following assumptions (Andrich & Luo, 2003):

- Each person *v* has a true score on the construct, usually denoted by the variable T_v
- The best overall indicator of the person's true score is the sum of the scores on the items and is usually denoted by the variable X_v
- This observed score will have an error for each person which is usually denoted by E_v
- These errors are not correlated with the true score
- Across a population of people, the errors sum to 0 and they are normally distributed.

From these assumptions, the following equations can be derived:

$$X_v = T_v + E_v .$$

Therefore,

$$S_x^2 = S_t^2 + S_e^2$$

where

S_x^2 = the variance of the observed score in a population of persons,

S_t^2 = the variance of their true score variance, and

S_e^2 = the error variance.

The reliability of the test can be calculated by the following formula:

$$r_{xx} = \frac{s_t^2}{s_x^2} = \frac{s_x^2 - s_e^2}{s_x^2} .$$

Thus, the *SEM* is calculated by the following formula:

$$S_e = S_x\sqrt{1 - r_{xx}} .$$

For example, consider a student with a score of 90 from a sample of students with a mean score of 60 and variance of 225 on a test with reliability of 0.80. According to the formulas provided above, the obtained score is 90, and its *SEM* is 6.71. Thus, an approximate 68% score band for estimating this students' true score is from 83.29 (90 - 6.71) to 96.71 (90 + 6.71).

Note that this equation is only useful to estimate true score when the test reliability is reasonably high and the obtained score for the examinee is not an extreme deviate from the mean of the appropriate reference group. When we use this equation, consequently, we should be careful with statements so that they do not imply greater precision than is actually involved (Harvill, 1991).

Conditional Standard Error of Measurement (Based on Item Response Theory)

Unlike the *SEM* based on the classical test theory, the *SEM* based on the *item response theory* (*IRT*) is not the same for all persons. For example, if a person gets few or a large number of items correct, the standard error is greater than if the person gets moderate number of items correct. This implies that the standard error of measurement depends on the total score (Andrich & Luo, 2003).

Under the Rasch model, the *SEM* for each person is as follows:

$$s_{\hat{b}} = \frac{1}{\sqrt{\sum_{i=1}^{L} p_{vi}(1-p_{vi})}}$$

where

v = subscript for a person,

i = subscript for an item,

L = length of the test,

$\hat{b}$ = ability estimate, and

p_{vi} = the probability that a person answers an item correctly and defined as follows:

$p_{vi} = \frac{e^{b_v - d_i}}{1 + e^{b_v - d_i}}$ where b_v is person's ability and d_i is item's difficulty.

A confidence band can be found for use in interpreting the ability estimate. For example, an approximate 68% confidence interval for $\hat{b}$ is given by

$\hat{b} \pm SEM$

Note that the standard error for item difficulty is smallest when the probability of passing is close to the probability of failing. That is, when an item is near the threshold level for many persons in the sample, the standard error is small (Embretson & Reise, 2000).

3.2 Scale Score Descriptive Statistics

Table 4.2 provides information about scale score descriptive statistics of each form for each grade and includes:

- Form number
- Number of items
- Numbers of students
- Means and standard deviations of scale scores
- Conditional *standard errors of measurement* (*SEM*) for the proficient and advanced cut scores

In addition, Appendix A provides frequency distributions and histograms of the scale scores of the 2003 *MSA-Reading*.

3.3 Classical and IRT Item Parameters

Appendix B provides both classical and *IRT*-based item parameters and includes:

- Item type (*SR* or *BCR*)
- *p*-value
- Point-biserial correlation: in order for *p*-values of the *BCR* items to be comparable with *p*-values of the *SR* items they were calculated as modified proportions of the maximum obtainable domain scores.
- Rasch difficulty estimate
- Standard error of the Rasch difficulty
- Mean-square infit
- Mean-square outfit

Item sequence numbers represents merely those items that were chosen to be in the final "score form."

Fit Statistics for Rasch Model

Fit statistics are used for evaluating the goodness-of-fit of a model to the data. Fit statistics are calculated by comparing the observed and expected trace lines obtained for an item after parameter estimates are obtained using a particular model. *WINSTEPS* provides two kinds of fit statistics called *mean-squares* that show the size of the randomness or amount of distortion of the measurement system.

Outfit mean-squares are influenced by outliers and are usually easy to diagnose and remedy. *Infit* mean-squares, on the other hand, are influenced by response patterns and are harder to diagnose and remedy. Table 3.1 provides a guideline for evaluating mean-square fit statistics (Linacre & Wright, 2000).

In general, mean-squares near 1.0 indicate little distortion of the measurement system, while values less than 1.0 indicate observations are too predictable (redundancy, model overfit). Values greater than 1.0 indicate unpredictability (unmodeled noise, model underfit).

Table 3.1 Criteria to Evaluate Mean-Square Fit Statistics

Mean-Square	Interpretation
> 2.0	Distorts or degrades the measurement system
1.5 – 2.0	Unproductive for construction of measurement, but not degrading
0.5 – 1.5	Productive for measurement
< 0.5	Unproductive for measurement, but not degrading. May produce misleadingly good reliabilities and separations

3.4 Inter-Rater Reliability

Tables 4.30 through 4.32 contain information about the scoring agreement between rater 1 and rater 2. When the two readers assigned the same score to a student's answer, the scores were in perfect agreement. Scores differed by one score point were adjacent, and scores differed by two or more score points were in discrepancy. For further information about inter-rater agreement, please see section 1.7. For the 2003 *MSA-Reading*, the adjacent agreement rates were above 95%, and perfect agreement rates were above 70% except several items for each grade.

3.5 Correlations among Reading Processes

The 2003 *MSA-Reading* consisted of three reading processes (strands): *General Reading*, *Literary Reading*, and *Informational Reading*. Tables 4.3 through 4.5 contain correlation coefficients among these reading processes. Generally, they show moderately strong intercorrelations among them.

3.6 Decision Accuracy and Consistency at the Cut Scores

Tables 4.6 through 4.8 contain the results of analyses performed to estimate the accuracy and consistency of the decisions for passing (proficient) on the 2003 *MSA-Reading*. The analyses make use of the methods outlined and implemented in Livingston and Lewis (1995), Haertel (1996), and Young and Yoon (1998).

The *accuracy* of a decision is the extent to which it would agree with the decisions that would be made if each student could somehow be tested with all possible parallel forms of the assessments. The *consistency* of a decision is the extent to which it would agree with the decisions that would be made if the students had taken a different form of the examination, equal in difficulty and covering the same content as the form they actually took.

Students can be misclassified in one of two ways. Students who were below the proficiency cut score, but were classified (on the basis of the assessment) as being above a cut score, are considered to be *false positives*. Students who were above the proficiency cut score, but were classified as being below a cut score, are considered to be *false negatives*.

For the 2003 *MSA-Reading*, Tables 4.6 through 4.8 include:

- Performance level
- Accuracy classifications
- False positives
- False negatives
- Consistency classifications

The tables illustrate the general rule that decision consistency is less than decision accuracy.

3.7 Differential Item Functioning

This section provides information about *differential item functioning* (*DIF*) analyses used for the 2003 *MSA-Reading*. For the 2003 *MSA-Reading DIF* analyses, the *reference* group was either male or Caucasian students, and the *focal* group was either female or African-American students.

Since the 2003 *MSA-Reading* was a mixed-format examination, comprising of both *BCR* and *SR* items, the *DIF* procedure used consists of Mantel's (1963) extension of the Mantel-Haenszel procedure (the Mantel Chi-square) for the *BCR* items and the Mantel-Haenszel procedure (Mantel & Haenszel, 1959) for the *SR* items.

Brief Constructed Response (BCR) Items

To help interpret the Mantel Chi-square (Mantel $?^2$), the Educational Testing Service (ETS) *DIF* procedure uses the Mantel statistic in conjunction with the *standardized mean difference* (*SMD*).

Mantel Statistic

The Mantel $?^2$ is simply a conditional mean comparison of the ordered response categories for reference and focal groups combined over values of the matching variable score. By "ordered" we mean that a response of 1 on an item is better than 0, 2 is better than 1, and so on. "Conditional," on the other hand, refers to the comparison of members from the two groups who received the same score on the matching variable, i.e., the total test score in our analysis.

Table 3.2 shows a $2 \times T \times K$ contingency table, where T is the number of response categories and K is the number of levels of the matching variable. The values, $y_1, y_2, \ldots, y_T$ are the T scores that can be gained on the item. The values, n_{Ftk} and n_{Rtk}, represent the numbers of focal and reference groups who are at the k^{th} level of the matching variable and gain an item score of y_t. The "+" indicates total number over a particular index (Zwick, Donoghue, & Grima, 1993).

Table 3.2 $2 \times T$ Contingency Table at the k^{th} level *

Group	Item Score				Total
	y_1	y_2		y_T	
Reference	n_{R1k}	n_{R2k}	…	n_{RTk}	n_{R+k}
Focal	n_{F1k}	n_{F2k}	…	n_{FTk}	n_{F+k}
Total	n_{+1k}	n_{+2k}	…	n_{+Tk}	n_{++k}

*Zwick, et al. (1993)

The Mantel statistics is defined as the following formula:

$$\text{Mantel } ?^2 = \frac{(\sum_k F_k - \sum_k E(F_k))^2}{\sum_k Var(F_k)}$$

where

F_k = the sum of scores for the focal group at the k^{th} level of the matching variable and is defined as follows:

$$F_k = \sum_t y_t n_{Ftk},$$

The expectation of F_k under the null hypothesis is

$$E(F_k) = \frac{n_{F+k}}{n_{++k}} \sum_t y_t n_{+tk}.$$

And, the variance of F_k under the null hypothesis is as follows:

$$Var(F_k) = \frac{n_{R+k} n_{F+k}}{n_{++k}^2 (n_{++k} - 1)} \left[(n_{++k} \sum_t y_t^2 n_{+tk}) - (\sum_t y_t n_{+tk})^2 \right].$$

Under H_0, the Mantel statistic has a chi-square distribution with one degree of freedom. In *DIF* applications, rejecting H_0 suggests that the students of the reference and focal groups who are similar in overall test performance tend to differ in their mean performance. In the case of dichotomous items, on the other hand, the statistics is identical to the Mantel-Haenszel (1959) statistic without the continuity correction (Zwick, Donoghue, & Grima, 1993).

Standardized Mean Difference (SMD)

A summary statistic to accompany the Mantel approach is the *standardized mean difference* (*SMD*) between the reference and focal groups proposed by Dorans and Schmitt (1991). This statistic compares the means of the reference and focal groups, adjusting for differences in the distribution of the reference and focal group members across the values of the matching variable.

$$SMD = \sum_k p_{Fk} m_{Fk} - \sum_k p_{Fk} m_{Rk}$$

where

$p_{Fk} = \frac{n_{F+k}}{n_{F++}}$, the proportion of the focal group members who are at the k^{th} level of the matching variable,

$m_{Fk} = \frac{1}{n_{F+k}(\sum_t y_t n_{Ftk})}$, the mean item score of the focal group members at the k^{th} level, and

m_{Rk} = the analogous value for the reference group.

As can be seen from the equation above, the *SMD* is the difference between the unweighted item mean of the focal group and the weighted item mean of the reference group. The weights for the reference group are applied to make the weighted number of the reference group students the same as in the focal group within the same ability. A negative *SMD* value implies that the focal group has a lower mean item score than the reference group, conditional on the matching variable.

DIF classification for BCR items

The *SMD* is divided by the total group item standard deviation to obtain an effect-size value for the *SMD*. This effect-size *SMD* is then examined in conjunction with the Mantel $?^2$ to obtain *DIF* classifications that are depicted in Table 3.3 below.

Table 3.3 *DIF* Classification for BCR Items

Category	Description	Criterion*
AA	No *DIF*	Non-significant Mantel $?^2$ or Significant Mantel $?^2$ and |SMD/SD| = .17
BB	Weak *DIF*	Significant Mantel $?^2$ and .17 < |SMD/SD| = .25
CC	Strong *DIF*	Significant Mantel $?^2$ and .25 < |SMD/SD|

* SD is the total group standard deviation of the item score in its original metric

Selected Response (SR) Items

For the *SR* items, the Mantel-Haenszel Chi-square (M-H $?^2$) in conjunction with the M-H odds ratio that is transferred to what ETS calls, the delta scale (D).

The Odds Ratio

The odds of a correct response (proportion passing divided by proportion failing) are *P/Q* or *P/(1-P)*. The odds ratio, on the other hand, is simply the odds of a correct response of the reference group divided by the odds of a correct response of the focal group.

For a given item, the odds ratio is defined as follows:

$$a_{M-H} = \frac{P_r / Q_r}{P_f / Qf}.$$

And, the corresponding null hypothesis is that the odds of getting the item correct are equal for the two groups. Thus, the odds ratio is equal to 1:

$$H_0: a_{M-H} = \frac{P_r / Q_r}{P_f / Qf} = 1.$$

The Delta Scale

In order to make the odds ratio symmetrical around zero with its range being in the interval $-\infty$ to $+\infty$, the odds ratio is transformed into a log odds ratio as per the following:

$$b_{M-H} = \ln(a_{\text{M-H}}).$$

The simple natural logarithm transformation of this odds ratio is symmetrical about zero in which zero has the interpretation of equal odds. This *DIF* measure is a signed index where a positive value signifies *DIF* in favor of the reference group while a negative value indicates *DIF* in favor of the focal group. b_{M-H} also has the advantage of being transformed linearly to other interval scale metrics (Camilli & Shepard, 1994). This fact is utilized by ETS in creating their delta scale (D), which is defined as follows:

$$\text{D} = -2.35 \cdot b_{M-H}.$$

DIF classification for SR items

The ETS examines the M-H $?^2$ in conjunction with the delta scale (D) to obtain *DIF* classifications depicted in Table 3.4 below.

Table 3.4 *DIF* Classification for SR Items

Category	Description	Criterion
A	No *DIF*	Non-significant M-H $?^2$ or |D| < 1.0
B	Weak *DIF*	Significant M-H $?^2$ and |D| < 1.5 or Non-significant M-H $?^2$ and |D| =1.0
C	Strong *DIF*	Significant M-H $?^2$ and |D| = 1.5

3.8 Equating and Scaling

Tables 4.9 through 4.29 contain information about raw score to scale score conversion tables for the 2003 *MSA-Reading*. Conditional standard errors for the scale scores are also included.

The Rasch and Partial Credit IRT Models

The most basic expression of the Rasch model is in the *item characteristic curve* (ICC). It shows the probability of a correct response to an item as a function of the ability level. The probability of a correct response is bounded by 1 (certainty of a correct response) and 0 (certainty of an incorrect response). The ability is, in theory, unbounded. In practice, the ability scale ranges from - 4 to + 4 logits for heterogeneous ability groups.

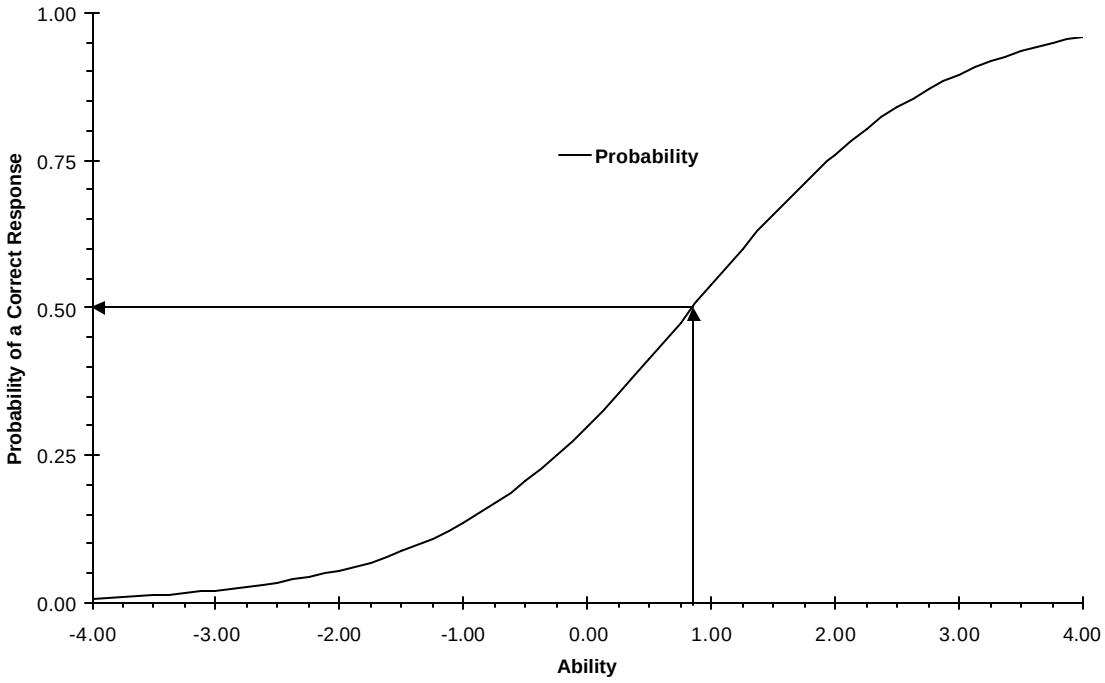
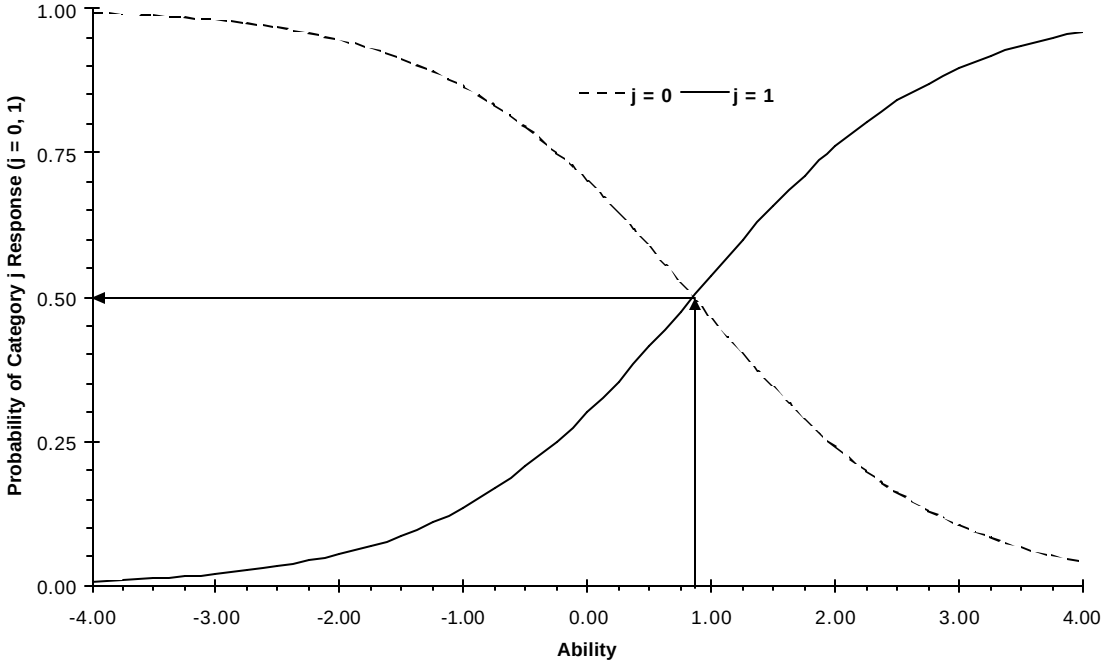


Figure 3.1 Item Characteristic Curve

As an example, consider Figure 3.1 which depicts a item that falls at approximately 0.85 on the ability (horizontal) scale. When a person answers an item at the same level as their ability, then that person has a probability of roughly 50% of answering the item correctly. Another way of expressing this is that if we have a group of 100 people, all of whom have an ability of 0.85, we would expect about 50% of them to answer the item correctly. A person whose ability was above 0.85 would a higher probability of getting the item right, while a person whose ability is below 0.85 would have a lower probability of getting the item right. This makes intuitive sense and is the basic formulation of Rasch measurement for test items having only 2 possible categories (i.e., wrong or right).



Figugure 3.2 Category Response Curves for a One-Step Item

Figure 3.2 extends this formulation to show the probabilities of obtaining a wrong answer or a right answer. The curve on the left ($j = 0$) shows the probability of getting a score of "0" while the curve on the right ($j = 1$) shows the probability of getting a score of "1". The point at which the two curves cross indicates the transition point on the ability scale where the most likely response changes from a "0" to a "1". Here, the probability of answering the item correctly is 50%.

The key step in the formulation, and the point at which the Rasch dichotomous model merges with the PCM, requires us to assume an additional response category. Suppose that, rather than scoring items as completely wrong or completely right, we add a category representing answers that, though not totally correct, are still clearly not totally incorrect. These relationships are shown in Figure 3.3.

The left-most curve ($j = 0$) in Figure 3.3 represents the probability for all examinees getting a score of "0" (completely incorrect) on the item, given their ability. Those of very low ability (e.g., below - 2) are very likely to be in this category and, in fact, are more likely to be in this category than the other two. Those receiving a "1" (partial credit) tend to fall in the middle range of abilities (the middle curve, $j = 1$). The final, right-most curve ($j = 2$) represents the probability for those receiving scores of "2" (completely correct). Very high-ability people are clearly more likely to be in this category than in any other, but there are still some of average and low ability that can get full credit for the item.

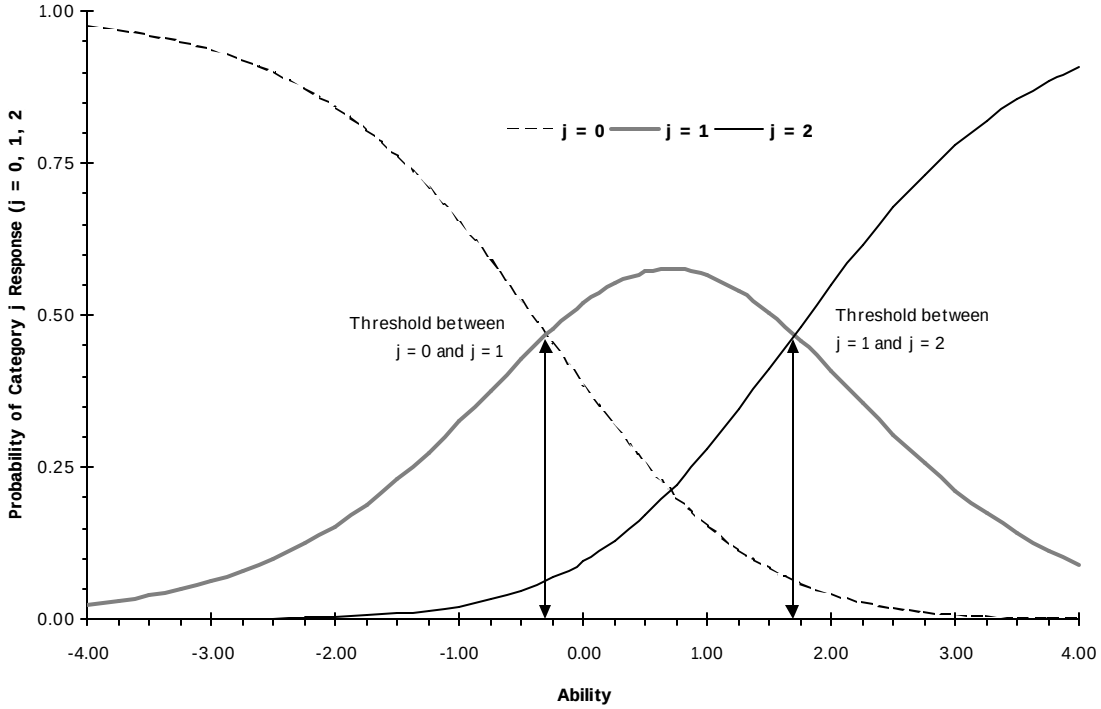


Figure 3.3 Category Response Curves for a Two-Step Item

Although the actual computations are quite complex, the points at which lines cross each other have a similar interpretation as for the dichotomous case. Consider the point at which the $j = 0$ line crosses the $j = 1$ line, indicated by the left arrow. For abilities to the left of (or less than) this point, the probability is greatest for a "0" response. To the right of (or above) this point, and up to the point at which the $j = 1$ and $j = 2$ lines cross (marked by the right arrow), the most likely response is a "1". For abilities to the right of this point, the most likely response is a "2".

Note that the probability of scoring a "1" response ($j = 1$) declines in both directions as ability decreases to the low extreme or increases to the high extreme. These points then may be thought of as the difficulties of crossing the *thresholds* between categories.

An important implication of the formulation can be summarized as: If the commonly used Rasch model applied to dichotomously (right/wrong) scored items can be thought of as simply a special case of the PCM, then the act of scaling multiple-choice items together with polytomous items, whether they have three or more response categories, is a straightforward process of applying the measurement model. The quality of the scaling then can be assessed in terms of known procedures.

One important property of the PCM is its ability to separate the estimation of item/task parameters from the person parameters. With the PCM, as with the Rasch model, the total score given by the sum of the categories in which a person responds is a sufficient statistic for estimating person ability (i.e., no additional information need be estimated). The total number of

responses across examinees in a particular category is a sufficient statistic for estimating the step difficulty for that category. Thus with PCM, the same total score will yield the same ability estimate for different examinees.

The PCM is a direct extension of the dichotomous one-parameter IRT model developed by Rasch (Rasch, 1980). For an item/task involving m_i score categories, one general expression for the probability of scoring x on item/task i is given by

$$P_{xi} = \exp \sum_{j=0}^{x} (q - D_{ij}) / \sum_{k=0}^{m_i} \left[\exp \sum_{j=0}^{k} (q - D_{ij}) \right]$$

where

$x = 0, 1, ..., m_i$, and by definition, $\sum_{j=0}^{0} (q - D_{ij}) = 0$.

The above equation gives the probability of scoring x on the i-th test item as a function of ability (q) and the difficulty of the m_i steps of the task (Masters, 1982).

According to this model, the probability of an examinee scoring in a particular category (step) is the sum of the logit (log-odds) differences between q and D_{ij} of all the completed steps, divided by the sum of the differences of all the steps of a task. Thissen and Steinberg (1986) refers to this model as a divide-by-total model. The parameters estimated by this model are (1) an ability estimate for each person (or ability estimate at each raw score level) and (2) m_i threshold (difficulty) estimates for each task with $m_i + 1$ score categories.

4. THE 2003 *MSA-READING* STATISTICAL SUMMARY

Table 4.1 The 2003 *MSA-Reading* Classical Descriptive Statistics: Grades 3, 5, and 8*

Grade	Form	# Items	n	Mean	SD	Reliability	SEM
3	1	37	8,298	24.78	8.22	.893	2.689
	2	37	8,615	25.09	8.11	.891	2.678
	3	37	8,562	26.01	8.04	.886	2.715
	4	37	8,466	24.12	7.94	.892	2.609
	5	37	8,597	25.40	8.50	.900	2.688
	6	37	8,291	24.88	8.18	.901	2.574
	7	37	7,619	25.73	7.87	.887	2.646
5	1	37	8,414	26.65	8.26	.893	2.702
	2	37	8,888	27.50	7.99	.890	2.650
	3	37	8,874	28.14	7.74	.883	2.647
	4	37	8,833	28.46	7.92	.885	2.686
	5	37	8,757	27.50	7.93	.892	2.606
	6	37	8,488	26.91	7.88	.890	2.614
	7	37	7,947	28.18	7.77	.882	2.669
8	1	34	8,326	26.10	7.73	.884	2.633
	2	34	8,799	26.31	7.91	.890	2.623
	3	34	8,824	26.90	7.42	.885	2.516
	4	34	8,874	26.00	7.33	.880	2.539
	5	34	8,787	25.52	7.29	.882	2.504
	6	34	8,604	26.94	7.00	.873	2.495
	7	34	4,013	26.52	7.14	.870	2.574

* Note: The analyses were based on the sample of data used to equate the forms of the 2003 *MSA-Reading*

Table 4.2 The 2003 *MSA-Reading* Scale Score Descriptive Statistics: Grades 3, 5, and 8

Grade	Form	# Items	n	Mean	SD	Conditional SEM at Cut-Points	
						Prof.	Adv.
3	1	37	9,596	415.84	48.23	12	16
	2	37	9,341	418.17	47.72	12	16
	3	37	9,368	418.27	48.31	12	20
	4	37	9,252	418.85	49.73	12	20
	5	37	9,282	418.78	50.35	12	16
	6	37	9,343	417.74	53.14	12	20
	7	37	8,289	423.76	47.58	12	16
5	1	37	10,050	421.58	47.98	12	16
	2	37	9,867	423.48	48.74	12	16
	3	37	9,802	423.81	48.07	12	16
	4	37	9,739	423.75	47.04	12	16
	5	37	9,696	414.68	50.16	12	16
	6	37	9,700	422.82	50.06	12	16
	7	37	8,868	430.26	46.71	12	12
8	1	34	9,763	429.08	54.13	12	16
	2	34	9,682	429.01	54.69	12	16
	3	34	9,716	429.77	51.71	12	16
	4	34	9,685	428.37	51.75	12	16
	5	34	9,669	428.38	52.51	12	16
	6	34	9,658	429.37	51.93	12	16
	7	34	9,533	439.92	50.73	12	16

Table 4.3 The 2003 *MSA-Reading* Strand Correlations: Grade 3*

Form	n	1	2	3
Form 1				
1. General Reading	8,298	1.00		
2. Literary Reading	8,298	0.66	1.00	
3. Information Reading	8,298	0.68	0.73	1.00
Form 2				
1. General Reading	8,615	1.00		
2. Literary Reading	8,615	0.67	1.00	
3. Information Reading	8,615	0.66	0.71	1.00
Form 3				
1. General Reading	8,562	1.00		
2. Literary Reading	8,562	0.66	1.00	
3. Information Reading	8,562	0.67	0.71	1.00
Form 4				
1. General Reading	8,466	1.00		
2. Literary Reading	8,466	0.68	1.00	
3. Information Reading	8,466	0.68	0.73	1.00
Form 5				
1. General Reading	8,597	1.00		
2. Literary Reading	8,597	0.69	1.00	
3. Information Reading	8,597	0.69	0.75	1.00
Form 6				
1. General Reading	8,291	1.00		
2. Literary Reading	8,291	0.69	1.00	
3. Information Reading	8,291	0.71	0.75	1.00
Form 7				
1. General Reading	7,619	1.00		
2. Literary Reading	7,619	0.64	1.00	
3. Information Reading	7,619	0.68	0.72	1.00

*Note: The analyses were based on the sample of data used to equate the forms of the 2003 *MSA-Reading*

Table 4.4 The 2003 *MSA-Reading* Strand Correlations: Grade 5*

Form	n	1	2	3
Form 1				
1. General Reading	8,414	1.00		
2. Literary Reading	8,414	0.69	1.00	
3. Information Reading	8,414	0.66	0.69	1.00
Form 2				
1. General Reading	8,888	1.00		
2. Literary Reading	8,888	0.68	1.00	
3. Information Reading	8,888	0.67	0.68	1.00
Form 3				
1. General Reading	8,874	1.00		
2. Literary Reading	8,874	0.66	1.00	
3. Information Reading	8,874	0.65	0.66	1.00
Form 4				
1. General Reading	8,833	1.00		
2. Literary Reading	8,833	0.68	1.00	
3. Information Reading	8,833	0.66	0.66	1.00
Form 5				
1. General Reading	8,757	1.00		
2. Literary Reading	8,757	0.68	1.00	
3. Information Reading	8,757	0.67	0.68	1.00
Form 6				
1. General Reading	8,488	1.00		
2. Literary Reading	8,488	0.68	1.00	
3. Information Reading	8,488	0.69	0.69	1.00
Form 7				
1. General Reading	7,947	1.00		
2. Literary Reading	7,947	0.69	1.00	
3. Information Reading	7,947	0.64	0.69	1.00

*Note: The analyses were based on the sample of data used to equate the forms of the 2003 *MSA-Reading*

Table 4.5 The 2003 *MSA-Reading* Strand Correlations: Grade 8*

Form	n	1	2	3
Form 1				
1. General Reading	8,326	1.00		
2. Literary Reading	8,326	0.70	1.00	
3. Information Reading	8,326	0.73	0.71	1.00
Form 2				
1. General Reading	8,799	1.00		
2. Literary Reading	8,799	0.68	1.00	
3. Information Reading	8,799	0.77	0.73	1.00
Form 3				
1. General Reading	8,824	1.00		
2. Literary Reading	8,824	0.66	1.00	
3. Information Reading	8,824	0.74	0.71	1.00
Form 4				
1. General Reading	8,874	1.00		
2. Literary Reading	8,874	0.68	1.00	
3. Information Reading	8,874	0.74	0.69	1.00
Form 5				
1. General Reading	8,787	1.00		
2. Literary Reading	8,787	0.68	1.00	
3. Information Reading	8,787	0.75	0.69	1.00
Form 6				
1. General Reading	8,604	1.00		
2. Literary Reading	8,604	0.66	1.00	
3. Information Reading	8,604	0.73	0.68	1.00
Form 7				
1. General Reading	4,013	1.00		
2. Literary Reading	4,013	0.68	1.00	
3. Information Reading	4,013	0.73	0.67	1.00

*Note: The analyses were based on the sample of data used to equate the forms of the 2003 *MSA-Reading*

Table 4.6 The 2003 *MSA-Reading* Decision Accuracy and Consistency Indices: Grade 3

Form	Performance Cut*	Accuracy	False Positive	False Negative	Consistency
1	*B : PA*	.90	.04	.06	.86
	BP : A	.96	.02	.01	.95
2	*B : PA*	.90	.05	.06	.86
	BP : A	.96	.03	.01	.94
3	*B : PA*	.90	.05	.06	.85
	BP : A	.96	.02	.02	.94
4	*B : PA*	.90	.05	.06	.86
	BP : A	.96	.03	.01	.94
5	*B : PA*	.90	.05	.05	.86
	BP : A	.96	.03	.01	.94
6	*B : PA*	.90	.04	.05	.86
	BP : A	.96	.02	.01	.95
7	*B : PA*	.90	.05	.06	.85
	BP : A	.96	.03	.01	.94

*Note: B:PA denotes the cut between Basic and Proficient, while BP:A denotes the cut between Proficient and Advanced

Table 4.7 The 2003 *MSA-Reading* Decision Accuracy and Consistency Indices: Grade 5

Form	Performance Cut*	Accuracy	False Positive	False Negative	Consistency
1	*B : PA*	.90	.05	.05	.87
	BP : A	.91	.05	.04	.88
2	*B : PA*	.90	.05	.05	.86
	BP : A	.91	.05	.04	.88
3	*B : PA*	.90	.05	.05	.86
	BP : A	.91	.05	.04	.87
4	*B : PA*	.90	.05	.05	.86
	BP : A	.91	.06	.04	.87
5	*B : PA*	.90	.05	.05	.86
	BP : A	.91	.05	.04	.88
6	*B : PA*	.90	.04	.05	.86
	BP : A	.91	.05	.04	.87
7	*B : PA*	.90	.05	.05	.86
	BP : A	.91	.05	.04	.87

*Note: B:PA denotes the cut between Basic and Proficient, while BP:A denotes the cut between Proficient and Advanced

Table 4.8 The 2003 *MSA-Reading* Decision Accuracy and Consistency Indices: Grade 8

Form	Performance Cut*	Accuracy	False Positive	False Negative	Consistency
1	*B : PA*	.91	.04	.05	.87
	BP : A	.88	.06	.05	.84
2	*B : PA*	.91	.04	.05	.88
	BP : A	.89	.06	.05	.84
3	*B : PA*	.91	.04	.05	.87
	BP : A	.88	.06	.06	.84
4	*B : PA*	.91	.05	.05	.87
	BP : A	.88	.06	.06	.83
5	*B : PA*	.91	.04	.05	.87
	BP : A	.88	.06	.06	.84
6	*B : PA*	.90	.04	.05	.86
	BP : A	.88	.06	.06	.83
7	*B : PA*	.90	.04	.06	.86
	BP : A	.88	.07	.06	.83

*Note: B:PA denotes the cut between basic and Proficient, while BP:A denotes the cut between Proficient and Advanced

Table 4.9 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 3 Form 1

	Form 1			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	191	56	135	247
1	220	40	180	260
2	252	28	224	280
3	271	24	247	295
4	286	20	266	306
5	298	20	278	318
6	308	16	292	324
7	317	16	301	333
8	325	16	309	341
9	332	16	316	348
10	340	16	324	356
11	346	16	330	362
12	352	12	340	364
13	358	12	346	370
14	363	12	351	375
15	369	12	357	381
16	374	12	362	386
17	379	12	367	391
18	384	12	372	396
19	389	12	377	401
20	394	12	382	406
21	399	12	387	411
22	404	12	392	416
23	409	12	397	421
24	414	12	402	426
25	419	12	407	431
26	424	12	412	436
27	429	12	417	441
28	434	12	422	446
29	440	12	428	452
30	445	12	433	457
31	451	12	439	463
32	457	12	445	469
33	464	16	448	480
34	470	16	454	486
35	478	16	462	494
36	485	16	469	501
37	494	16	478	510
38	503	16	487	519
39	513	20	493	533
40	525	20	505	545
41	538	24	514	562
42	555	24	531	579
43	576	32	544	608
44	610	40	570	650
45	640	56	584	696

Table 4.10 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 3 Form 2

	Form 2			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	198	56	142	254
1	226	40	186	266
2	256	28	228	284
3	275	24	251	299
4	289	20	269	309
5	300	20	280	320
6	310	16	294	326
7	318	16	302	334
8	326	16	310	342
9	333	16	317	349
10	340	16	324	356
11	346	12	334	358
12	352	12	340	364
13	358	12	346	370
14	363	12	351	375
15	368	12	356	380
16	374	12	362	386
17	379	12	367	391
18	384	12	372	396
19	389	12	377	401
20	394	12	382	406
21	398	12	386	410
22	403	12	391	415
23	408	12	396	420
24	413	12	401	425
25	418	12	406	430
26	423	12	411	435
27	428	12	416	440
28	434	12	422	446
29	439	12	427	451
30	444	12	432	456
31	450	12	438	462
32	456	12	444	468
33	463	16	447	479
34	470	16	454	486
35	477	16	461	493
36	485	16	469	501
37	493	16	477	509
38	502	20	482	522
39	513	20	493	533
40	525	20	505	545
41	539	24	515	563
42	556	24	532	580
43	578	32	546	610
44	612	40	572	652
45	643	56	587	699

Table 4.11 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 3 Form 3

	Form 3			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	198	56	142	254
1	227	40	187	267
2	257	28	229	285
3	275	24	251	299
4	289	20	269	309
5	300	20	280	320
6	309	16	293	325
7	317	16	301	333
8	324	16	308	340
9	331	16	315	347
10	338	12	326	350
11	344	12	332	356
12	349	12	337	361
13	354	12	342	366
14	360	12	348	372
15	364	12	352	376
16	370	12	358	382
17	374	12	362	386
18	379	12	367	391
19	384	12	372	396
20	388	12	376	400
21	392	12	380	404
22	397	12	385	409
23	402	12	390	414
24	406	12	394	418
25	411	12	399	423
26	416	12	404	428
27	421	12	409	433
28	426	12	414	438
29	432	12	420	444
30	437	12	425	449
31	443	12	431	455
32	450	16	434	466
33	456	16	440	472
34	464	16	448	480
35	471	16	455	487
36	480	16	464	496
37	490	20	470	510
38	501	20	481	521
39	514	20	494	534
40	529	24	505	553
41	548	28	520	576
42	575	32	543	607
43	614	40	574	654
44	667	48	619	715
45	706	60	646	766

Table 4.12 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 3 Form 4

	Form 4			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	198	56	142	254
1	228	40	188	268
2	258	28	230	286
3	277	24	253	301
4	291	20	271	311
5	302	20	282	322
6	312	16	296	328
7	321	16	305	337
8	329	16	313	345
9	336	16	320	352
10	343	16	327	359
11	349	12	337	361
12	355	12	343	367
13	361	12	349	373
14	366	12	354	378
15	372	12	360	384
16	377	12	365	389
17	382	12	370	394
18	387	12	375	399
19	392	12	380	404
20	397	12	385	409
21	402	12	390	414
22	407	12	395	419
23	412	12	400	424
24	417	12	405	429
25	422	12	410	434
26	428	12	416	440
27	433	12	421	445
28	439	12	427	451
29	445	12	433	457
30	451	16	435	467
31	458	16	442	474
32	465	16	449	481
33	473	16	457	489
34	481	16	465	497
35	491	20	471	511
36	502	20	482	522
37	515	24	491	539
38	532	24	508	556
39	553	32	521	585
40	585	40	545	625
41	639	48	591	687
42	697	44	653	741
43	741	40	701	781
44	786	44	742	800
45	800	60	762	800

*Note: The highest obtainable scale score is 800

Table 4.13 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 3 Form 5

	Form 5			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	201	56	145	257
1	230	40	190	270
2	260	28	232	288
3	278	24	254	302
4	292	20	272	312
5	302	20	282	322
6	312	16	296	328
7	320	16	304	336
8	327	16	311	343
9	334	16	318	350
10	340	12	328	352
11	346	12	334	358
12	352	12	340	364
13	357	12	345	369
14	362	12	350	374
15	367	12	355	379
16	372	12	360	384
17	377	12	365	389
18	382	12	370	394
19	386	12	374	398
20	391	12	379	403
21	396	12	384	408
22	401	12	389	413
23	406	12	394	418
24	410	12	398	422
25	415	12	403	427
26	420	12	408	432
27	426	12	414	438
28	431	12	419	443
29	436	12	424	448
30	442	12	430	454
31	448	12	436	460
32	454	16	438	470
33	460	16	444	476
34	467	16	451	483
35	475	16	459	491
36	483	16	467	499
37	492	16	476	508
38	502	20	482	522
39	513	20	493	533
40	526	20	506	546
41	541	24	517	565
42	560	28	532	588
43	584	32	552	616
44	621	44	577	665
45	653	56	597	709

Table 4.14 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 3 Form 6

	Form 6			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	188	56	132	244
1	218	40	178	258
2	249	28	221	277
3	269	24	245	293
4	284	20	264	304
5	295	20	275	315
6	306	16	290	322
7	314	16	298	330
8	322	16	306	338
9	330	16	314	346
10	336	16	320	352
11	343	16	327	359
12	349	12	337	361
13	355	12	343	367
14	360	12	348	372
15	366	12	354	378
16	371	12	359	383
17	376	12	364	388
18	382	12	370	394
19	387	12	375	399
20	392	12	380	404
21	397	12	385	409
22	402	12	390	414
23	407	12	395	419
24	412	12	400	424
25	418	12	406	430
26	423	12	411	435
27	428	12	416	440
28	434	12	422	446
29	440	12	428	452
30	446	16	430	462
31	453	16	437	469
32	460	16	444	476
33	468	16	452	484
34	476	16	460	492
35	485	20	465	505
36	496	20	476	516
37	508	20	488	528
38	524	24	500	548
39	543	28	515	571
40	568	32	536	600
41	600	36	564	636
42	638	40	598	678
43	682	44	638	726
44	742	52	690	794
45	788	64	724	800

*Note: The highest obtainable scale score is 800

Table 4.15 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 3 Form 7

	Form 7			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	195	56	139	251
1	224	40	184	264
2	254	28	226	282
3	273	24	249	297
4	287	20	267	307
5	298	20	278	318
6	308	16	292	324
7	317	16	301	333
8	325	16	309	341
9	332	16	316	348
10	339	16	323	355
11	346	12	334	358
12	352	12	340	364
13	357	12	345	369
14	363	12	351	375
15	368	12	356	380
16	374	12	362	386
17	379	12	367	391
18	384	12	372	396
19	389	12	377	401
20	394	12	382	406
21	399	12	387	411
22	404	12	392	416
23	409	12	397	421
24	414	12	402	426
25	419	12	407	431
26	424	12	412	436
27	430	12	418	442
28	435	12	423	447
29	441	12	429	453
30	446	12	434	458
31	452	12	440	464
32	459	16	443	475
33	466	16	450	482
34	473	16	457	489
35	480	16	464	496
36	489	16	473	505
37	498	16	482	514
38	508	20	488	528
39	519	20	499	539
40	532	20	512	552
41	547	24	523	571
42	564	28	536	592
43	587	32	555	619
44	621	40	581	661
45	652	56	596	708

Table 4.16 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 5 Form 1

	Form 1			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	208	56	152	264
1	236	40	196	276
2	266	28	238	294
3	284	24	260	308
4	297	20	277	317
5	307	16	291	323
6	316	16	300	332
7	324	16	308	340
8	331	16	315	347
9	338	16	322	354
10	344	12	332	356
11	350	12	338	362
12	355	12	343	367
13	360	12	348	372
14	365	12	353	377
15	370	12	358	382
16	375	12	363	387
17	379	12	367	391
18	384	12	372	396
19	388	12	376	400
20	393	12	381	405
21	397	12	385	409
22	402	12	390	414
23	406	12	394	418
24	411	12	399	423
25	415	12	403	427
26	420	12	408	432
27	425	12	413	437
28	430	12	418	442
29	435	12	423	447
30	440	12	428	452
31	446	12	434	458
32	452	12	440	464
33	458	16	442	474
34	464	16	448	480
35	472	16	456	488
36	479	16	463	495
37	488	16	472	504
38	498	20	478	518
39	509	20	489	529
40	523	24	499	547
41	539	24	515	563
42	558	28	530	586
43	584	32	552	616
44	622	44	578	666
45	656	60	596	716

Table 4.17 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 5 Form 2

	Form 2			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	195	56	139	251
1	224	40	184	264
2	254	28	226	282
3	272	24	248	296
4	286	20	266	306
5	297	20	277	317
6	306	16	290	322
7	314	16	298	330
8	322	16	306	338
9	329	16	313	345
10	335	12	323	347
11	341	12	329	353
12	347	12	335	359
13	352	12	340	364
14	357	12	345	369
15	362	12	350	374
16	367	12	355	379
17	372	12	360	384
18	377	12	365	389
19	382	12	370	394
20	386	12	374	398
21	391	12	379	403
22	396	12	384	408
23	400	12	388	412
24	405	12	393	417
25	410	12	398	422
26	415	12	403	427
27	420	12	408	432
28	425	12	413	437
29	430	12	418	442
30	436	12	424	448
31	442	12	430	454
32	448	16	432	464
33	454	16	438	470
34	461	16	445	477
35	469	16	453	485
36	477	16	461	493
37	486	16	470	502
38	496	20	476	516
39	507	20	487	527
40	520	24	496	544
41	536	24	512	560
42	556	28	528	584
43	584	36	548	620
44	627	44	583	671
45	662	60	602	722

Table 4.18 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 5 Form 3

	Form 3			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	189	56	133	245
1	218	40	178	258
2	249	28	221	277
3	268	24	244	292
4	282	20	262	302
5	293	20	273	313
6	302	16	286	318
7	311	16	295	327
8	319	16	303	335
9	326	16	310	342
10	332	16	316	348
11	338	12	326	350
12	344	12	332	356
13	350	12	338	362
14	355	12	343	367
15	360	12	348	372
16	365	12	353	377
17	370	12	358	382
18	374	12	362	386
19	379	12	367	391
20	384	12	372	396
21	389	12	377	401
22	393	12	381	405
23	398	12	386	410
24	402	12	390	414
25	407	12	395	419
26	412	12	400	424
27	417	12	405	429
28	422	12	410	434
29	427	12	415	439
30	432	12	420	444
31	438	12	426	450
32	444	12	432	456
33	450	16	434	466
34	457	16	441	473
35	464	16	448	480
36	472	16	456	488
37	480	16	464	496
38	490	20	470	510
39	501	20	481	521
40	514	24	490	538
41	531	24	507	555
42	552	28	524	580
43	582	36	546	618
44	624	44	580	668
45	660	60	600	720

Table 4.19 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 5 Form 4

	Form 4			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	188	56	132	244
1	218	40	178	258
2	250	28	222	278
3	269	24	245	293
4	283	20	263	303
5	295	20	275	315
6	304	16	288	320
7	313	16	297	329
8	320	16	304	336
9	327	16	311	343
10	334	12	322	346
11	340	12	328	352
12	345	12	333	357
13	350	12	338	362
14	356	12	344	368
15	360	12	348	372
16	365	12	353	377
17	370	12	358	382
18	374	12	362	386
19	379	12	367	391
20	384	12	372	396
21	388	12	376	400
22	392	12	380	404
23	397	12	385	409
24	402	12	390	414
25	406	12	394	418
26	411	12	399	423
27	416	12	404	428
28	420	12	408	432
29	425	12	413	437
30	430	12	418	442
31	436	12	424	448
32	441	12	429	453
33	447	12	435	459
34	454	16	438	470
35	460	16	444	476
36	468	16	452	484
37	476	16	460	492
38	485	16	469	501
39	495	20	475	515
40	507	20	487	527
41	521	24	497	545
42	539	28	511	567
43	564	32	532	596
44	602	44	558	646
45	635	60	575	695

Table 4.20 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 5 Form 5

	Form 5			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	182	56	126	238
1	212	40	172	252
2	244	28	216	272
3	264	24	240	288
4	279	20	259	299
5	291	20	271	311
6	302	16	286	318
7	311	16	295	327
8	319	16	303	335
9	326	16	310	342
10	333	16	317	349
11	340	12	328	352
12	346	12	334	358
13	351	12	339	363
14	357	12	345	369
15	362	12	350	374
16	367	12	355	379
17	372	12	360	384
18	377	12	365	389
19	382	12	370	394
20	386	12	374	398
21	391	12	379	403
22	396	12	384	408
23	401	12	389	413
24	406	12	394	418
25	410	12	398	422
26	415	12	403	427
27	420	12	408	432
28	425	12	413	437
29	431	12	419	443
30	436	12	424	448
31	442	12	430	454
32	448	16	432	464
33	455	16	439	471
34	462	16	446	478
35	470	16	454	486
36	478	16	462	494
37	488	20	468	508
38	499	20	479	519
39	512	24	488	536
40	528	24	504	552
41	548	28	520	576
42	570	28	542	598
43	598	32	566	630
44	636	44	592	680
45	668	56	612	724

Table 4.21 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 5 Form 6

	Form 6			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	198	56	142	254
1	227	40	187	267
2	257	28	229	285
3	275	24	251	299
4	288	20	268	308
5	300	20	280	320
6	309	16	293	325
7	317	16	301	333
8	324	16	308	340
9	331	16	315	347
10	338	12	326	350
11	343	12	331	355
12	349	12	337	361
13	354	12	342	366
14	360	12	348	372
15	365	12	353	377
16	370	12	358	382
17	374	12	362	386
18	379	12	367	391
19	384	12	372	396
20	389	12	377	401
21	394	12	382	406
22	398	12	386	410
23	403	12	391	415
24	408	12	396	420
25	413	12	401	425
26	418	12	406	430
27	423	12	411	435
28	428	12	416	440
29	434	12	422	446
30	439	12	427	451
31	446	16	430	462
32	452	16	436	468
33	459	16	443	475
34	466	16	450	482
35	475	16	459	491
36	484	20	464	504
37	496	20	476	516
38	508	24	484	532
39	525	24	501	549
40	545	28	517	573
41	572	32	540	604
42	605	36	569	641
43	642	36	606	678
44	688	44	644	732
45	724	60	664	784

Table 4.22 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 5 Form 7

	Form 7			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	205	56	149	261
1	234	40	194	274
2	263	28	235	291
3	281	24	257	305
4	294	20	274	314
5	305	16	289	321
6	314	16	298	330
7	322	16	306	338
8	329	16	313	345
9	336	16	320	352
10	342	12	330	354
11	348	12	336	360
12	353	12	341	365
13	358	12	346	370
14	364	12	352	376
15	368	12	356	380
16	373	12	361	385
17	378	12	366	390
18	383	12	371	395
19	387	12	375	399
20	392	12	380	404
21	396	12	384	408
22	401	12	389	413
23	405	12	393	417
24	410	12	398	422
25	414	12	402	426
26	419	12	407	431
27	424	12	412	436
28	429	12	417	441
29	434	12	422	446
30	439	12	427	451
31	445	12	433	457
32	450	12	438	462
33	456	12	444	468
34	463	16	447	479
35	470	16	454	486
36	477	16	461	493
37	486	16	470	502
38	495	16	479	511
39	505	20	485	525
40	517	20	497	537
41	532	24	508	556
42	550	28	522	578
43	574	32	542	606
44	612	44	568	656
45	645	60	585	705

Table 4.23 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 8 Form 1

	Form 1			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	187	56	131	243
1	218	40	178	258
2	251	28	223	279
3	272	24	248	296
4	287	20	267	307
5	300	20	280	320
6	310	20	290	330
7	320	16	304	336
8	328	16	312	344
9	335	16	319	351
10	342	16	326	358
11	348	12	336	360
12	354	12	342	366
13	360	12	348	372
14	365	12	353	377
15	370	12	358	382
16	376	12	364	388
17	380	12	368	392
18	385	12	373	397
19	390	12	378	402
20	394	12	382	406
21	399	12	387	411
22	403	12	391	415
23	408	12	396	420
24	412	12	400	424
25	417	12	405	429
26	421	12	409	433
27	426	12	414	438
28	431	12	419	443
29	436	12	424	448
30	441	12	429	453
31	446	12	434	458
32	452	12	440	464
33	457	12	445	469
34	464	16	448	480
35	470	16	454	486
36	477	16	461	493
37	485	16	469	501
38	494	16	478	510
39	503	20	483	523
40	514	20	494	534
41	528	24	504	552
42	544	24	520	568
43	564	28	536	592
44	597	40	557	637
45	628	56	572	684

Table 4.24 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 8 Form 2

	Form 2			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	186	56	130	242
1	216	40	176	256
2	248	28	220	276
3	269	24	245	293
4	284	20	264	304
5	296	20	276	316
6	307	20	287	327
7	316	16	300	332
8	324	16	308	340
9	332	16	316	348
10	339	16	323	355
11	345	12	333	357
12	351	12	339	363
13	356	12	344	368
14	362	12	350	374
15	367	12	355	379
16	372	12	360	384
17	377	12	365	389
18	382	12	370	394
19	386	12	374	398
20	391	12	379	403
21	395	12	383	407
22	400	12	388	412
23	404	12	392	416
24	409	12	397	421
25	413	12	401	425
26	418	12	406	430
27	422	12	410	434
28	427	12	415	439
29	432	12	420	444
30	437	12	425	449
31	442	12	430	454
32	448	12	436	460
33	454	12	442	466
34	460	16	444	476
35	466	16	450	482
36	473	16	457	489
37	481	16	465	497
38	490	16	474	506
39	499	20	479	519
40	510	20	490	530
41	524	24	500	548
42	540	24	516	564
43	560	28	532	588
44	593	40	553	633
45	624	56	568	680

Table 4.25 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 8 Form 3

	Form 3			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	186	56	130	242
1	216	40	176	256
2	248	28	220	276
3	267	24	243	291
4	282	20	262	302
5	294	20	274	314
6	304	16	288	320
7	314	16	298	330
8	322	16	306	338
9	329	16	313	345
10	336	16	320	352
11	342	12	330	354
12	348	12	336	360
13	353	12	341	365
14	358	12	346	370
15	364	12	352	376
16	369	12	357	381
17	374	12	362	386
18	378	12	366	390
19	383	12	371	395
20	388	12	376	400
21	392	12	380	404
22	397	12	385	409
23	402	12	390	414
24	406	12	394	418
25	411	12	399	423
26	416	12	404	428
27	420	12	408	432
28	425	12	413	437
29	430	12	418	442
30	436	12	424	448
31	441	12	429	453
32	447	12	435	459
33	453	12	441	465
34	459	16	443	475
35	466	16	450	482
36	473	16	457	489
37	481	16	465	497
38	490	16	474	506
39	499	20	479	519
40	510	20	490	530
41	523	20	503	543
42	538	24	514	562
43	558	28	530	586
44	591	40	551	631
45	621	56	565	677

Table 4.26 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 8 Form 4

	Form 4			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	176	56	120	232
1	208	40	168	248
2	242	32	210	274
3	263	24	239	287
4	280	24	256	304
5	293	20	273	313
6	304	20	284	324
7	314	16	298	330
8	323	16	307	339
9	331	16	315	347
10	338	16	322	354
11	344	16	328	360
12	350	12	338	362
13	356	12	344	368
14	362	12	350	374
15	367	12	355	379
16	372	12	360	384
17	377	12	365	389
18	382	12	370	394
19	386	12	374	398
20	391	12	379	403
21	396	12	384	408
22	400	12	388	412
23	405	12	393	417
24	409	12	397	421
25	414	12	402	426
26	419	12	407	431
27	424	12	412	436
28	428	12	416	440
29	434	12	422	446
30	439	12	427	451
31	444	12	432	456
32	450	12	438	462
33	456	16	440	472
34	463	16	447	479
35	470	16	454	486
36	478	16	462	494
37	486	16	470	502
38	496	20	476	516
39	508	20	488	528
40	521	24	497	545
41	537	24	513	561
42	556	28	528	584
43	581	32	549	613
44	617	44	573	661
45	649	56	593	705

Table 4.27 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 8 Form 5

	Form 5			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	190	56	134	246
1	220	40	180	260
2	252	28	224	280
3	272	24	248	296
4	287	20	267	307
5	299	20	279	319
6	309	16	293	325
7	318	16	302	334
8	326	16	310	342
9	333	16	317	349
10	340	16	324	356
11	346	12	334	358
12	352	12	340	364
13	357	12	345	369
14	363	12	351	375
15	368	12	356	380
16	373	12	361	385
17	378	12	366	390
18	382	12	370	394
19	387	12	375	399
20	392	12	380	404
21	397	12	385	409
22	402	12	390	414
23	406	12	394	418
24	411	12	399	423
25	416	12	404	428
26	421	12	409	433
27	426	12	414	438
28	431	12	419	443
29	436	12	424	448
30	442	12	430	454
31	448	12	436	460
32	454	16	438	470
33	461	16	445	477
34	468	16	452	484
35	476	16	460	492
36	484	16	468	500
37	494	20	474	514
38	504	20	484	524
39	516	20	496	536
40	529	24	505	553
41	544	24	520	568
42	563	28	535	591
43	586	32	554	618
44	620	40	580	660
45	651	56	595	707

Table 4.28 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 8 Form 6

	Form 6			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	184	56	128	240
1	214	40	174	254
2	247	28	219	275
3	267	24	243	291
4	282	20	262	302
5	294	20	274	314
6	305	16	289	321
7	314	16	298	330
8	322	16	306	338
9	329	16	313	345
10	336	16	320	352
11	342	12	330	354
12	348	12	336	360
13	354	12	342	366
14	359	12	347	371
15	364	12	352	376
16	369	12	357	381
17	374	12	362	386
18	378	12	366	390
19	383	12	371	395
20	388	12	376	400
21	392	12	380	404
22	397	12	385	409
23	402	12	390	414
24	406	12	394	418
25	411	12	399	423
26	415	12	403	427
27	420	12	408	432
28	425	12	413	437
29	430	12	418	442
30	436	12	424	448
31	441	12	429	453
32	447	12	435	459
33	453	16	437	469
34	460	16	444	476
35	467	16	451	483
36	475	16	459	491
37	484	16	468	500
38	494	20	474	514
39	505	20	485	525
40	519	24	495	543
41	535	24	511	559
42	556	28	528	584
43	584	36	548	620
44	626	44	582	670
45	662	60	602	722

Table 4.29 The 2003 *MSA-Reading* Raw Score to Scale Score Conversion Table: Grade 8 Form 7

	Form 7			
Raw Score	Scale Score (SS)	Standard Error (SEM)	SS - 1SEM	SS + 1SEM
0	194	56	138	250
1	224	40	184	264
2	256	28	228	284
3	276	24	252	300
4	291	20	271	311
5	303	20	283	323
6	313	16	297	329
7	322	16	306	338
8	330	16	314	346
9	338	16	322	354
10	344	16	328	360
11	351	12	339	363
12	356	12	344	368
13	362	12	350	374
14	368	12	356	380
15	373	12	361	385
16	378	12	366	390
17	383	12	371	395
18	388	12	376	400
19	392	12	380	404
20	397	12	385	409
21	402	12	390	414
22	406	12	394	418
23	411	12	399	423
24	415	12	403	427
25	420	12	408	432
26	425	12	413	437
27	430	12	418	442
28	435	12	423	447
29	440	12	428	452
30	445	12	433	457
31	450	12	438	462
32	456	12	444	468
33	462	16	446	478
34	469	16	453	485
35	476	16	460	492
36	484	16	468	500
37	492	16	476	508
38	502	20	482	522
39	512	20	492	532
40	525	20	505	545
41	540	24	516	564
42	558	28	530	586
43	582	32	550	614
44	617	40	577	657
45	649	56	593	705

Table 4. 30 The 2003 *MSA-Reading* Score Difference between Rater 1 and Rater 2: Grade 3

Form	Item	Perfect		Adjacent		Discrepancy		Total	
		n	%	N	%	n	%	n	%
1	1	7057	74.9	2293	24.3	77	0.8	9427	100.0
	2	6525	70.7	2597	28.1	113	1.2	9235	100.0
	3	6287	69.1	2685	29.5	120	1.3	9092	100.0
	4	7452	79.5	1878	20.0	39	0.4	9369	100.0
	5	6485	71.2	2537	27.9	84	0.9	9106	100.0
	6	6364	70.4	2499	27.7	171	1.9	9034	100.0
2	1	6718	73.6	2358	25.8	51	0.6	9127	100.0
	2	6300	69.7	2624	29.0	111	1.2	9035	100.0
	3	6645	75.3	2072	23.5	110	1.2	8827	100.0
	4	7657	85.5	1261	14.1	34	0.4	8952	100.0
	5	6960	77.2	2012	22.3	43	0.5	9015	100.0
	6	6976	78.7	1834	20.7	52	0.6	8862	100.0
3	1	6763	73.4	2348	25.5	102	1.1	9213	100.0
	2	6790	74.6	2251	24.7	56	0.6	9097	100.0
	3	6094	68.0	2741	30.6	121	1.4	8956	100.0
	4	7228	80.2	1671	18.5	118	1.3	9017	100.0
	5	7282	80.1	1773	19.5	37	0.4	9092	100.0
	6	6579	75.1	2144	24.5	38	0.4	8761	100.0
4	1	7269	80.6	1654	18.3	91	1.0	9014	100.0
	2	7651	85.3	1283	14.3	31	0.3	8965	100.0
	3	6696	76.8	1963	22.5	56	0.6	8715	100.0
	4	6928	77.2	2033	22.7	8	0.1	8969	100.0
	5	8141	90.2	862	9.6	21	0.2	9024	100.0
	6	6689	75.5	2007	22.7	164	1.9	8860	100.0
5	1	6746	74.7	2208	24.4	81	0.9	9035	100.0
	2	6198	71.0	2436	27.9	90	1.0	8724	100.0
	3	5282	61.1	3153	36.5	208	2.4	8643	100.0
	4	6757	76.1	2042	23.0	78	0.9	8877	100.0
	5	5954	66.9	2826	31.8	119	1.3	8899	100.0
	6	7289	82.2	1547	17.5	27	0.3	8863	100.0
6	1	7747	85.1	1359	14.9	0	0.0	9106	100.0
	2	7553	88.4	976	11.4	13	0.2	8542	100.0
	3	6675	76.3	1974	22.6	103	1.2	8752	100.0
	4	7237	85.3	1225	14.4	27	0.3	8489	100.0
	5	6609	75.2	2147	24.4	34	0.4	8790	100.0
	6	6934	79.5	1734	19.9	52	0.6	8720	100.0
7	1	6025	74.2	2060	25.4	35	0.4	8120	100.0
	2	5690	71.2	2165	27.1	139	1.7	7994	100.0
	3	5584	71.0	2202	28.0	80	1.0	7866	100.0
	4	6371	78.3	1716	21.1	48	0.6	8135	100.0
	5	6094	76.6	1815	22.8	45	0.6	7954	100.0
	6	6146	77.1	1737	21.8	84	1.1	7967	100.0

Table 4.31 The 2003 *MSA-Reading* Score Difference between Rater 1 and Rater 2: Grade 5

Form	Item	Perfect		Adjacent		Discrepancy		Total	
		n	%	n	%	n	%	n	%
1	1	7021	70.9	2754	27.8	124	1.3	9899	100.0
	2	6553	66.4	3177	32.2	141	1.4	9871	100.0
	3	6517	66.4	3201	32.6	99	1.0	9817	100.0
	4	7530	75.9	2260	22.8	127	1.3	9917	100.0
	5	7536	76.5	2278	23.1	35	0.4	9849	100.0
	6	6340	65.1	3213	33.0	191	2.0	9744	100.0
2	1	6678	68.5	2973	30.5	101	1.0	9752	100.0
	2	7575	78.2	2089	21.6	25	0.3	9689	100.0
	3	6830	70.9	2765	28.7	32	0.3	9627	100.0
	4	6887	72.5	2541	26.8	68	0.7	9496	100.0
	5	6536	67.9	3001	31.2	83	0.9	9620	100.0
	6	7019	73.0	2512	26.1	90	0.9	9621	100.0
3	1	7535	77.8	2141	22.1	10	0.1	9686	100.0
	2	7301	76.0	2274	23.7	29	0.3	9604	100.0
	3	7855	83.0	1529	16.2	83	0.9	9467	100.0
	4	6733	70.1	2758	28.7	113	1.2	9604	100.0
	5	5897	61.4	3567	37.1	145	1.5	9609	100.0
	6	6975	72.3	2581	26.8	92	1.0	9648	100.0
4	1	5900	61.2	3613	37.5	131	1.4	9644	100.0
	2	6526	67.6	3045	31.5	84	0.9	9655	100.0
	3	7329	76.0	2230	23.1	79	0.8	9638	100.0
	4	6919	73.2	2505	26.5	32	0.3	9456	100.0
	5	7496	78.8	1994	21.0	17	0.2	9507	100.0
	6	5995	64.2	3076	32.9	268	2.9	9339	100.0
5	1	5509	57.3	3806	39.6	298	3.1	9613	100.0
	2	6415	67.6	2851	30.0	227	2.4	9493	100.0
	3	4967	52.2	4305	45.2	247	2.6	9519	100.0
	4	6649	69.8	2817	29.6	54	0.6	9520	100.0
	5	6503	67.8	3007	31.4	76	0.8	9586	100.0
	6	6622	71.9	2523	27.4	67	0.7	9212	100.0
6	1	6689	70.2	2799	29.4	38	0.4	9526	100.0
	2	6974	72.8	2568	26.8	32	0.3	9574	100.0
	3	7616	80.2	1867	19.7	17	0.2	9500	100.0
	4	7319	76.4	2228	23.3	27	0.3	9574	100.0
	5	7536	80.2	1832	19.5	28	0.3	9396	100.0
	6	6625	70.6	2690	28.7	65	0.7	9380	100.0
7	1	5717	65.5	2872	32.9	145	1.7	8734	100.0
	2	5791	66.6	2681	30.9	218	2.5	8690	100.0
	3	6093	71.0	2435	28.4	49	0.6	8577	100.0
	4	6461	73.5	2213	25.2	114	1.3	8788	100.0
	5	5497	63.0	3097	35.5	137	1.6	8731	100.0
	6	5440	63.1	2945	34.2	234	2.7	8619	100.0

Table 4.32 The 2003 *MSA-Reading* Score Difference between Rater 1 and Rater 2: Grade 8

Form	Item	Perfect		Adjacent		Discrepancy		Total	
		n	%	n	%	n	%	n	%
1	1	5839	62.8	3302	35.5	156	1.7	9297	100.0
	2	5627	61.2	3338	36.3	226	2.5	9191	100.0
	3	5273	58.8	3415	38.1	287	3.2	8975	100.0
	4	6166	66.4	3033	32.7	85	0.9	9284	100.0
	5	5932	62.9	3292	34.9	203	2.2	9427	100.0
	6	5096	55.3	3847	41.7	280	3.0	9223	100.0
2	1	6289	66.9	3023	32.2	84	0.9	9396	100.0
	2	6399	69.9	2705	29.6	45	0.5	9149	100.0
	3	6203	66.9	2964	32.0	100	1.1	9267	100.0
	4	5932	65.2	3062	33.6	111	1.2	9105	100.0
	5	5437	59.7	3502	38.5	164	1.8	9103	100.0
	6	5513	62.8	3079	35.1	185	2.1	8777	100.0
3	1	6423	68.0	2844	30.1	181	1.9	9448	100.0
	2	5403	56.8	3929	41.3	172	1.8	9504	100.0
	3	6526	70.0	2556	27.4	242	2.6	9324	100.0
	4	6358	67.0	2993	31.5	145	1.5	9496	100.0
	5	6203	65.9	2976	31.6	228	2.4	9407	100.0
	6	6917	74.7	2108	22.8	234	2.5	9259	100.0
4	1	5749	61.4	3403	36.4	207	2.2	9359	100.0
	2	6769	73.3	2425	26.2	46	0.5	9240	100.0
	3	5980	63.8	3242	34.6	147	1.6	9369	100.0
	4	6679	71.6	2629	28.2	21	0.2	9329	100.0
	5	6016	64.5	3125	33.5	193	2.1	9334	100.0
	6	6710	72.0	2558	27.5	47	0.5	9315	100.0
5	1	6243	67.6	2910	31.5	87	0.9	9240	100.0
	2	6265	68.3	2799	30.5	105	1.1	9169	100.0
	3	6458	69.7	2666	28.8	135	1.5	9259	100.0
	4	5621	60.6	3450	37.2	208	2.2	9279	100.0
	5	6023	64.3	3265	34.9	80	0.9	9368	100.0
	6	5380	58.5	3586	39.0	228	2.5	9194	100.0
6	1	6546	69.2	2803	29.6	107	1.1	9456	100.0
	2	5146	54.9	3789	40.5	432	4.6	9367	100.0
	3	4397	47.1	4263	45.6	683	7.3	9343	100.0
	4	6481	70.2	2706	29.3	49	0.5	9236	100.0
	5	7302	77.9	2057	21.9	16	0.2	9375	100.0
	6	5770	61.9	3428	36.8	127	1.4	9325	100.0
7	1	5988	65.2	3068	33.4	125	1.4	9181	100.0
	2	5837	64.8	2947	32.7	220	2.4	9004	100.0
	3	5433	61.3	3151	35.6	279	3.1	8863	100.0
	4	6208	67.4	2941	31.9	62	0.7	9211	100.0
	5	6561	70.3	2730	29.3	37	0.4	9328	100.0
	6	5762	62.8	3275	35.7	131	1.4	9168	100.0

REFERENCES

AERA, APA, & NCME (1999). *Standards for educational and psychological testing*. Washington, D.C.: Author.

Allen, N. L., Donoghue, J. R., & Schoeps, T. L. (2001). *The NAEP 1998 technical report* (Technical Report). Washington, DC: National Center for Educational Statistics.

Andrich, A. (1988). *Rasch models for measurement*. Newbury Park, CA: SAGE Publications, Inc.

Andrich, A. (1989). Distinctions between assumptions and requirements in measurement in the social sciences. In J. A. Keats, R. Taft, R. A. Heath, & H. H. Lovibond (Eds.) *Mathematical and theoretical systems.* North-Holland: Elsevier Science Publisher B.V.

Andrich, A., & Luo, G. (2003). *Modern measurement and analysis in social science*. Murdoch University, Perth, Western Australia.

Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items.* Thousand Oaks, CA: SAGE Publications.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: Holt Rinehart Wilson.

CTB/McGraw-Hill (2003, August). *The Maryland standard setting technical report*. (Technical Report). Monterey, CA: CTB/McGraw-Hill.

Dorans, N. J., & Schmitt, A. P. (1991). *Constructed-response and differential item functioning: A pragmatic approach* (ETS Research Report No. 91-49). Princeton, NJ: Educational Testing Service.

Embretson, S., & Reise, S. (2000). *Item response theory for psychologists.* New Jersey : Lawrence Erlbaum Associates, Publishers.

Haertel, E. H. (1996). *Estimating the decision consistency from a single administration of a performance assessment battery. A report on the National Board of Professional Teaching Standards McGEN Assessment*. Palo Alto, CA: Stanford University.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: SAGE Publications, Inc.

Harvill, L. M. (1991). Standard error of measurement. *Educational Measurement: Issues and Practice,* 10, 181-189.

Huynh, H., Meyer III, J. P., & Barton, K. (2000). *Technical documentation for the 1999 Palmetto achievement challenge tests of English language arts and mathematics, grades three through eight* (Technical Report). Columbia: South Carolina Department of Education.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8 & PRELIS 2: User's reference guide*. Chicago: Scientific Software International.

Linacre, J. M., & Wright, B. D. (2000). *A user's guide to WINSTEPS: Rasch-model computer program.* Chicago, IL: MESA Press.

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement, 32,* 179-197.

Lord, F. M., & Wingersky, M. S. (1984). Comparison of IRT true-score and equipercentile observed-score "equatings." *Applied Psychological Measurement, 8,* 452-461.

Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association, 58,* 690-700.

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

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrica, 47,* 149-174.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18,* 5-11.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards* (pp. 249-282). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Qualls, A. L. (1995). Estimating the reliability of a test containing multiple item formats, *Applied Measurement in Education,* 8, 111-120.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests.* Chicago, IL: University of Chicago Press.

Ryan, J. P. (1983). Introduction to latent trait analysis and item response theory. In W. E. Hathaway (Ed.), *Testing in the schools. New directions for testing and measurement, 19,* San Francisco: Jossey-Bass.

South Carolina Department of Education. (2001). *Technical documentation for the 2000 Palmetto achievement challenge tests of English language arts and mathematics* (Technical Report). Columbia: South Carolina Department of Education.

Suen, H. K. (1990). *Principles of test theories*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrica, 51,* 567-577.

Young, M. J., & Yoon, B. (1998, April). *Estimating the consistency and accuracy of classifications in a standards-referenced assessment.* (CSE Technical Report 475). Center for the Study of Evaluation, Standards, and Student Testing. Los Angeles, CA: University of California, Los Angeles.

Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement, 30,* 233-251.

APPENDIX A: THE 2003 *MSA-READING* SCALE SCORE HISTOGRAMS

Grade=3 Form=1

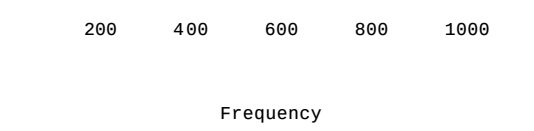
Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
188	*	26	26	0.27	0.27
203		0	26	0.00	0.27
218		2	28	0.02	0.29
233		0	28	0.00	0.29
248		3	31	0.03	0.32
263		0	31	0.00	0.32
278		7	38	0.07	0.40
293	*	25	63	0.26	0.66
308	*	33	96	0.34	1.00
323	***	127	223	1.32	2.32
338	****	214	437	2.23	4.55
353	**********	618	1055	6.44	10.99
368	**********	898	1953	9.36	20.35
383	**********	1128	3081	11.75	32.11
398	**********	1167	4248	12.16	44.27
413	**********	1163	5411	12.12	56.39
428	**********	1139	6550	11.87	68.26
443	**********	659	7209	6.87	75.13
458	**********	945	8154	9.85	84.97
473	**********	558	8712	5.81	90.79
488	*********	425	9137	4.43	95.22
503	***	172	9309	1.79	97.01
518	**	121	9430	1.26	98.27
533	***	125	9555	1.30	99.57
548		0	9555	0.00	99.57
563	*	31	9586	0.32	99.90
578		7	9593	0.07	99.97
593		0	9593	0.00	99.97
608		3	9596	0.03	100.00

400 800

Frequency

Grade=3 Form=2

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
203	*	23	23	0.25	0.25
218		0	23	0.00	0.25
233		1	24	0.01	0.26
248		0	24	0.00	0.26
263		4	28	0.04	0.30
278		3	31	0.03	0.33
293		12	43	0.13	0.46
308	*	36	79	0.39	0.85
323	****	105	184	1.12	1.97
338	*********	230	414	2.46	4.43
353	**********************	552	966	5.91	10.34
368	********************************	818	1784	8.76	19.10
383	*************************************	967	2751	10.35	29.45
398	**	1085	3836	11.62	41.07
413	**	1150	4986	12.31	53.38
428	**	1129	6115	12.09	65.46
443	***************************	711	6826	7.61	73.08
458	*************************************	997	7823	10.67	83.75
473	**********************	601	8424	6.43	90.18
488	*****************	457	8881	4.89	95.08
503	*******	169	9050	1.81	96.88
518	******	139	9189	1.49	98.37
533	*****	113	9302	1.21	99.58
548		0	9302	0.00	99.58
563	*	26	9328	0.28	99.86
578		10	9338	0.11	99.97
593		0	9338	0.00	99.97
608		3	9341	0.03	100.00



Grade=3 Form=3

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
203	*	27	27	0.29	0.29
218		0	27	0.00	0.29
233		0	27	0.00	0.29
248		0	27	0.00	0.29
263		1	28	0.01	0.30
278		3	31	0.03	0.33
293		9	40	0.10	0.43
308	*	26	66	0.28	0.70
323	*	73	139	0.78	1.48
338	******	315	454	3.36	4.85
353	******	295	749	3.15	8.00
368	******************	974	1723	10.40	18.39
383	******************	973	2696	10.39	28.78
398	**********************	1121	3817	11.97	40.75
413	************************	1188	5005	12.68	53.43
428	**********************	1087	6092	11.60	65.03
443	***************	764	6856	8.16	73.19
458	**********************	1052	7908	11.23	84.42
473	******	313	8221	3.34	87.76
488	**********	554	8775	5.91	93.67
503	****	224	8999	2.39	96.06
518	***	166	9165	1.77	97.83
533	**	115	9280	1.23	99.06
548	*	57	9337	0.61	99.67
563		0	9337	0.00	99.67
578		22	9359	0.23	99.90
593		0	9359	0.00	99.90
608		7	9366	0.07	99.98
623		0	9366	0.00	99.98
638		0	9366	0.00	99.98
653		0	9366	0.00	99.98
668		2	9368	0.02	100.00

--------+-------+-------+

400 800 1200

Frequency

Grade=3 Form=4

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
203	*	28	28	0.30	0.30
218		0	28	0.00	0.30
233		3	31	0.03	0.34
248		0	31	0.00	0.34
263		2	33	0.02	0.36
278		2	35	0.02	0.38
293		4	39	0.04	0.42
308	**	49	88	0.53	0.95
323	****	110	198	1.19	2.14
338	**********	268	466	2.90	5.04
353	**********	383	849	4.14	9.18
368	**********	849	1698	9.18	18.35
383	**********	996	2694	10.77	29.12
398	**********	1082	3776	11.69	40.81
413	**********	1138	4914	12.30	53.11
428	**********	1065	5979	11.51	64.62
443	**********	757	6736	8.18	72.81
458	**********	721	7457	7.79	80.60
473	**********	656	8113	7.09	87.69
488	**********	562	8675	6.07	93.76
503	*********	232	8907	2.51	96.27
518	******	153	9060	1.65	97.92
533	****	107	9167	1.16	99.08
548	**	57	9224	0.62	99.70
563		0	9224	0.00	99.70
578		0	9224	0.00	99.70
593	*	21	9245	0.23	99.92
608		0	9245	0.00	99.92
623		0	9245	0.00	99.92
638		7	9252	0.08	100.00

200 400 600 800 1000

Frequency

Grade=3 Form=5

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
203	*	22	22	0.24	0.24
218		0	22	0.00	0.24
233		0	22	0.00	0.24
248		0	22	0.00	0.24
263		4	26	0.04	0.28
278		5	31	0.05	0.33
293		10	41	0.11	0.44
308	**	48	89	0.52	0.96
323	*****	130	219	1.40	2.36
338	**********	246	465	2.65	5.01
353	**********	565	1030	6.09	11.10
368	**********	758	1788	8.17	19.26
383	**********	890	2678	9.59	28.85
398	**********	1008	3686	10.86	39.71
413	**********	1089	4775	11.73	51.44
428	**********	1075	5850	11.58	63.03
443	**********	1025	6875	11.04	74.07
458	**********	671	7546	7.23	81.30
473	**********	601	8147	6.47	87.77
488	**********	526	8673	5.67	93.44
503	*********	215	8888	2.32	95.76
518	*******	177	9065	1.91	97.66
533	****	109	9174	1.17	98.84
548	**	52	9226	0.56	99.40
563	*	30	9256	0.32	99.72
578	*	17	9273	0.18	99.90
593		0	9273	0.00	99.90
608		0	9273	0.00	99.90
623		7	9280	0.08	99.98
638		0	9280	0.00	99.98
653		2	9282	0.02	100.00

200 400 600 800 1000

Frequency

Grade=3 Form=6

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
	|				
188	|*	31	31	0.33	0.33
203	|	0	31	0.00	0.33
218	|	3	34	0.03	0.36
233	|	0	34	0.00	0.36
248	|	3	37	0.03	0.40
263	|	3	40	0.03	0.43
278	|	9	49	0.10	0.52
293	|*	16	65	0.17	0.70
308	|***	77	142	0.82	1.52
323	|**	58	200	0.62	2.14
338	|***************	397	597	4.25	6.39
353	|******************	415	1012	4.44	10.83
368	|******************************	863	1875	9.24	20.07
383	|**	1048	2923	11.22	31.29
398	|**	1144	4067	12.24	43.53
413	|**	1050	5117	11.24	54.77
428	|***************************************	1026	6143	10.98	65.75
443	|***********************************	721	6864	7.72	73.47
458	|**********************************	661	7525	7.07	80.54
473	|**********************************	665	8190	7.12	87.66
488	|**********************	305	8495	3.26	90.92
503	|*****************************	497	8992	5.32	96.24
518	|******	161	9153	1.72	97.97
533	|	0	9153	0.00	97.97
548	|*****	113	9266	1.21	99.18
563	|**	54	9320	0.58	99.75
578	|	0	9320	0.00	99.75
593	|	0	9320	0.00	99.75
608	|*	20	9340	0.21	99.97
623	|	0	9340	0.00	99.97
638	|	3	9343	0.03	100.00

--------+-------+-------+-------+-------+------

200 400 600 800 1000

Frequency

Grade=3 Form=7

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
195		11	11	0.13	0.13
210		0	11	0.00	0.13
225		1	12	0.01	0.14
240		0	12	0.00	0.14
255		1	13	0.01	0.16
270		5	18	0.06	0.22
285		10	28	0.12	0.34
300		11	39	0.13	0.47
315	**	53	92	0.64	1.11
330	****	100	192	1.21	2.32
345	**********	291	483	3.51	5.83
360	**********	370	853	4.46	10.29
375	**********	673	1526	8.12	18.41
390	**********	913	2439	11.01	29.42
405	**********	996	3435	12.02	41.44
420	**********	1039	4474	12.53	53.98
435	**********	1048	5522	12.64	66.62
450	**********	677	6199	8.17	74.79
465	**********	683	6882	8.24	83.03
480	**********	537	7419	6.48	89.50
495	**********	432	7851	5.21	94.72
510	******	157	8008	1.89	96.61
525	*********	228	8236	2.75	99.36
540	*	31	8267	0.37	99.73
555		0	8267	0.00	99.73
570	*	13	8280	0.16	99.89
585		7	8287	0.08	99.98
600		0	8287	0.00	99.98
615		2	8289	0.02	100.00

200 400 600 800 1000

Frequency

Grade=5 Form=1

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
210		15	15	0.15	0.15
225		0	15	0.00	0.15
240		3	18	0.03	0.18
255		0	18	0.00	0.18
270		3	21	0.03	0.21
285		5	26	0.05	0.26
300	*	36	62	0.36	0.62
315	*	49	111	0.49	1.10
330	***	136	247	1.35	2.46
345	********	407	654	4.05	6.51
360	************	655	1309	6.52	13.02
375	***************	766	2075	7.62	20.65
390	************************	1249	3324	12.43	33.07
405	**********************	1105	4429	11.00	44.07
420	**********************	1145	5574	11.39	55.46
435	************************	1193	6767	11.87	67.33
450	*****************	856	7623	8.52	75.85
465	**********************	1095	8718	10.90	86.75
480	*******	352	9070	3.50	90.25
495	**********	544	9614	5.41	95.66
510	***	168	9782	1.67	97.33
525	***	134	9916	1.33	98.67
540	*	72	9988	0.72	99.38
555	*	42	10030	0.42	99.80
570		0	10030	0.00	99.80
585		15	10045	0.15	99.95
600		0	10045	0.00	99.95
615		4	10049	0.04	99.99
630		0	10049	0.00	99.99
645		0	10049	0.00	99.99
660		1	10050	0.01	100.00

--------+-------+-------+-

400 800 1200

Frequency

Grade=5 Form=2

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
198		18	18	0.18	0.18
210		0	18	0.00	0.18
222		1	19	0.01	0.19
234		0	19	0.00	0.19
246		0	19	0.00	0.19
258		2	21	0.02	0.21
270		6	27	0.06	0.27
282		7	34	0.07	0.34
294		19	53	0.19	0.54
306	*	27	80	0.27	0.81
318	**	99	179	1.00	1.81
330	***	155	334	1.57	3.39
342	*****	273	607	2.77	6.15
354	*******	333	940	3.37	9.53
366	********	410	1350	4.16	13.68
378	**********	743	2093	7.53	21.21
390	**********	568	2661	5.76	26.97
402	**********	1008	3669	10.22	37.18
414	**********	724	4393	7.34	44.52
426	**********	1249	5642	12.66	57.18
438	**********	873	6515	8.85	66.03
450	**********	956	7471	9.69	75.72
462	********	424	7895	4.30	80.01
474	**********	816	8711	8.27	88.28
486	*******	374	9085	3.79	92.07
498	******	299	9384	3.03	95.10
510	****	211	9595	2.14	97.24
522	***	145	9740	1.47	98.71
534	**	75	9815	0.76	99.47
546		0	9815	0.00	99.47
558	*	37	9852	0.37	99.85
570		0	9852	0.00	99.85
582		15	9867	0.15	100.00

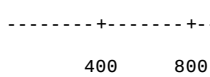
--------+-------+-------+-

400 800 1200

Frequency

Grade=5 Form=3

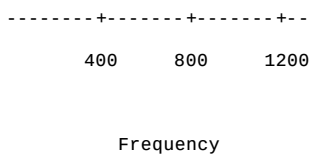
Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
195		22	22	0.22	0.22
210		0	22	0.00	0.22
225		1	23	0.01	0.23
240		0	23	0.00	0.23
255		3	26	0.03	0.27
270		3	29	0.03	0.30
285		3	32	0.03	0.33
300	*	28	60	0.29	0.61
315	*	69	129	0.70	1.32
330	**	101	230	1.03	2.35
345	******	301	531	3.07	5.42
360	**********	532	1063	5.43	10.84
375	**********	727	1790	7.42	18.26
390	**********	915	2705	9.33	27.60
405	**********	1520	4225	15.51	43.10
420	**********	1268	5493	12.94	56.04
435	**********	855	6348	8.72	64.76
450	**********	1348	7696	13.75	78.51
465	**********	800	8496	8.16	86.68
480	*******	349	8845	3.56	90.24
495	**********	546	9391	5.57	95.81
510	****	187	9578	1.91	97.71
525	**	124	9702	1.27	98.98
540		0	9702	0.00	98.98
555	*	68	9770	0.69	99.67
570		0	9770	0.00	99.67
585	*	28	9798	0.29	99.96
600		0	9798	0.00	99.96
615		0	9798	0.00	99.96
630		3	9801	0.03	99.99
645		0	9801	0.00	99.99
660		1	9802	0.01	100.00



Frequency

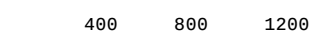
Grade=5 Form=4

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
195		16	16	0.16	0.16
210		0	16	0.00	0.16
225		0	16	0.00	0.16
240		0	16	0.00	0.16
255		4	20	0.04	0.21
270		2	22	0.02	0.23
285		9	31	0.09	0.32
300	*	32	63	0.33	0.65
315	**	81	144	0.83	1.48
330	***	129	273	1.32	2.80
345	******	324	597	3.33	6.13
360	*********	474	1071	4.87	11.00
375	************	668	1739	6.86	17.86
390	**********************	1153	2892	11.84	29.70
405	*********************	1081	3973	11.10	40.79
420	**********************	1208	5181	12.40	53.20
435	***************************	1346	6527	13.82	67.02
450	*****************	913	7440	9.37	76.39
465	*****************	834	8274	8.56	84.96
480	**************	710	8984	7.29	92.25
495	******	291	9275	2.99	95.24
510	****	219	9494	2.25	97.48
525	***	131	9625	1.35	98.83
540	*	71	9696	0.73	99.56
555		0	9696	0.00	99.56
570	*	35	9731	0.36	99.92
585		0	9731	0.00	99.92
600		6	9737	0.06	99.98
615		0	9737	0.00	99.98
630		2	9739	0.02	100.00



Grade=5 Form=5

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
188		24	24	0.25	0.25
203		0	24	0.00	0.25
218		1	25	0.01	0.26
233		0	25	0.00	0.26
248		1	26	0.01	0.27
263		1	27	0.01	0.28
278		3	30	0.03	0.31
293		7	37	0.07	0.38
308	*	38	75	0.39	0.77
323	**	98	173	1.01	1.78
338	****	182	355	1.88	3.66
353	*********	465	820	4.80	8.46
368	**********	658	1478	6.79	15.24
383	**********	846	2324	8.73	23.97
398	**********	979	3303	10.10	34.07
413	**********	1078	4381	11.12	45.18
428	**********	1277	5658	13.17	58.35
443	**********	1260	6918	13.00	71.35
458	**********	834	7752	8.60	79.95
473	**********	798	8550	8.23	88.18
488	*******	334	8884	3.44	91.63
503	******	311	9195	3.21	94.83
518	****	206	9401	2.12	96.96
533	***	143	9544	1.47	98.43
548	**	91	9635	0.94	99.37
563		0	9635	0.00	99.37
578	*	34	9669	0.35	99.72
593		20	9689	0.21	99.93
608		0	9689	0.00	99.93
623		0	9689	0.00	99.93
638		7	9696	0.07	100.00



Grade=5 Form=6

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
203		18	18	0.19	0.19
218		0	18	0.00	0.19
233		1	19	0.01	0.20
248		0	19	0.00	0.20
263		5	24	0.05	0.25
278		8	32	0.08	0.33
293		8	40	0.08	0.41
308	*	49	89	0.51	0.92
323	**	113	202	1.16	2.08
338	*******	352	554	3.63	5.71
353	******	292	846	3.01	8.72
368	******************	877	1723	9.04	17.76
383	*****************	852	2575	8.78	26.55
398	********************	1021	3596	10.53	37.07
413	**********************	1161	4757	11.97	49.04
428	**********************	1199	5956	12.36	61.40
443	*****************	830	6786	8.56	69.96
458	*****************	861	7647	8.88	78.84
473	*****************	806	8453	8.31	87.14
488	*******	372	8825	3.84	90.98
503	************	575	9400	5.93	96.91
518		0	9400	0.00	96.91
533	***	163	9563	1.68	98.59
548	**	96	9659	0.99	99.58
563		0	9659	0.00	99.58
578	*	29	9688	0.30	99.88
593		0	9688	0.00	99.88
608		11	9699	0.11	99.99
623		0	9699	0.00	99.99
638		1	9700	0.01	100.00

--------+-------+-------+

400 800 1200

Frequency

Grade=5 Form=7

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
210		16	16	0.18	0.18
225		0	16	0.00	0.18
240		1	17	0.01	0.19
255		0	17	0.00	0.19
270		6	23	0.07	0.26
285		2	25	0.02	0.28
300		23	48	0.26	0.54
315	*	42	90	0.47	1.01
330	**	105	195	1.18	2.20
345	***	141	336	1.59	3.79
360	*******	325	661	3.66	7.45
375	**********	560	1221	6.31	13.77
390	******************	962	2183	10.85	24.62
405	******************	881	3064	9.93	34.55
420	*********************	1055	4119	11.90	46.45
435	************************	1191	5310	13.43	59.88
450	*************************	1263	6573	14.24	74.12
465	**********	782	7355	8.82	82.94
480	**********	675	8030	7.61	90.55
495	******	281	8311	3.17	93.72
510	*******	360	8671	4.06	97.78
525	**	110	8781	1.24	99.02
540		0	8781	0.00	99.02
555	*	53	8834	0.60	99.62
570		23	8857	0.26	99.88
585		0	8857	0.00	99.88
600		0	8857	0.00	99.88
615		10	8867	0.11	99.99
630		0	8867	0.00	99.99
645		1	8868	0.01	100.00

--------+-------+-------+-

400 800 1200

Frequency

Grade=8 Form=1

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
	|				
188	|**	81	81	0.83	0.83
203	|	0	81	0.00	0.83
218	|	1	82	0.01	0.84
233	|	0	82	0.00	0.84
248	|	3	85	0.03	0.87
263	|	0	85	0.00	0.87
278	|	10	95	0.10	0.97
293	|	15	110	0.15	1.13
308	|*	50	160	0.51	1.64
323	|**	105	265	1.08	2.71
338	|***	157	422	1.61	4.32
353	|*****	241	663	2.47	6.79
368	|**********	540	1203	5.53	12.32
383	|**********	690	1893	7.07	19.39
398	|************************	1237	3130	12.67	32.06
413	|********************	985	4115	10.09	42.15
428	|********************	1062	5177	10.88	53.03
443	|**********************	1091	6268	11.17	64.20
458	|************************	1187	7455	12.16	76.36
473	|**********	793	8248	8.12	84.48
488	|**********	682	8930	6.99	91.47
503	|*****	262	9192	2.68	94.15
518	|*****	244	9436	2.50	96.65
533	|***	154	9590	1.58	98.23
548	|**	85	9675	0.87	99.10
563	|*	55	9730	0.56	99.66
578	|	0	9730	0.00	99.66
593	|*	27	9757	0.28	99.94
608	|	0	9757	0.00	99.94
623	|	6	9763	0.06	100.00

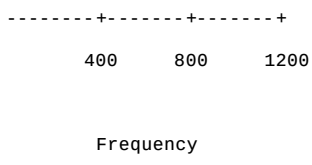
--------+-------+-------+-

400 800 1200

Frequency

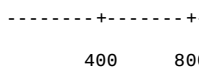
Grade=8 Form=2

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
188	**	89	89	0.92	0.92
203		0	89	0.00	0.92
218		5	94	0.05	0.97
233		0	94	0.00	0.97
248		3	97	0.03	1.00
263		5	102	0.05	1.05
278		9	111	0.09	1.15
293		17	128	0.18	1.32
308		24	152	0.25	1.57
323	*	66	218	0.68	2.25
338	**	116	334	1.20	3.45
353	*******	343	677	3.54	6.99
368	**********	502	1179	5.18	12.18
383	***************	751	1930	7.76	19.93
398	************************	1210	3140	12.50	32.43
413	********************	1038	4178	10.72	43.15
428	**********************	1139	5317	11.76	54.92
443	**********************	1162	6479	12.00	66.92
458	**********	792	7271	8.18	75.10
473	***************	757	8028	7.82	82.92
488	**************	677	8705	6.99	89.91
503	******	282	8987	2.91	92.82
518	*********	425	9412	4.39	97.21
533		0	9412	0.00	97.21
548	***	133	9545	1.37	98.59
563	**	89	9634	0.92	99.50
578		0	9634	0.00	99.50
593	*	37	9671	0.38	99.89
608		0	9671	0.00	99.89
623		11	9682	0.11	100.00



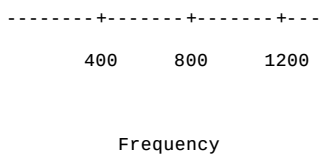
Grade=8 Form=3

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
188	*	65	65	0.67	0.67
203		0	65	0.00	0.67
218		1	66	0.01	0.68
233		0	66	0.00	0.68
248		1	67	0.01	0.69
263		7	74	0.07	0.76
278		8	82	0.08	0.84
293		12	94	0.12	0.97
308	*	48	142	0.49	1.46
323	**	97	239	1.00	2.46
338	****	195	434	2.01	4.47
353	*******	366	800	3.77	8.23
368	**********	492	1292	5.06	13.30
383	************	666	1958	6.85	20.15
398	**********	785	2743	8.08	28.23
413	**********	974	3717	10.02	38.26
428	**********	1194	4911	12.29	50.55
443	**********	1395	6306	14.36	64.90
458	**********	913	7219	9.40	74.30
473	**********	907	8126	9.34	83.64
488	**********	775	8901	7.98	91.61
503	******	277	9178	2.85	94.46
518	*******	364	9542	3.75	98.21
533	**	85	9627	0.87	99.08
548		0	9627	0.00	99.08
563	*	60	9687	0.62	99.70
578		0	9687	0.00	99.70
593	*	25	9712	0.26	99.96
608		0	9712	0.00	99.96
623		4	9716	0.04	100.00



Grade=8 Form=4

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
180	**	82	82	0.85	0.85
195		0	82	0.00	0.85
210		4	86	0.04	0.89
225		0	86	0.00	0.89
240		1	87	0.01	0.90
255		0	87	0.00	0.90
270		7	94	0.07	0.97
285		9	103	0.09	1.06
300	*	30	133	0.31	1.37
315	*	29	162	0.30	1.67
330	**	88	250	0.91	2.58
345	*****	274	524	2.83	5.41
360	*********	431	955	4.45	9.86
375	**********	623	1578	6.43	16.29
390	**********	786	2364	8.12	24.41
405	**********	1022	3386	10.55	34.96
420	**********	1139	4525	11.76	46.72
435	**********	1241	5766	12.81	59.54
450	**********	1367	7133	14.11	73.65
465	**********	849	7982	8.77	82.42
480	**********	820	8802	8.47	90.88
495	*******	328	9130	3.39	94.27
510	*****	245	9375	2.53	96.80
525	***	169	9544	1.74	98.54
540	*	73	9617	0.75	99.30
555	*	42	9659	0.43	99.73
570		0	9659	0.00	99.73
585		19	9678	0.20	99.93
600		0	9678	0.00	99.93
615		7	9685	0.07	100.00



Grade=8 Form=5

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
195	**	77	77	0.80	0.80
210		0	77	0.00	0.80
225		3	80	0.03	0.83
240		0	80	0.00	0.83
255		4	84	0.04	0.87
270		7	91	0.07	0.94
285		13	104	0.13	1.08
300		16	120	0.17	1.24
315	*	71	191	0.73	1.98
330	**	115	306	1.19	3.16
345	******	304	610	3.14	6.31
360	******	305	915	3.15	9.46
375	**********	798	1713	8.25	17.72
390	**********	793	2506	8.20	25.92
405	**********	1036	3542	10.71	36.63
420	**********	1221	4763	12.63	49.26
435	**********	1308	6071	13.53	62.79
450	**********	895	6966	9.26	72.04
465	**********	865	7831	8.95	80.99
480	**********	766	8597	7.92	88.91
495	******	324	8921	3.35	92.26
510	**********	496	9417	5.13	97.39
525	**	114	9531	1.18	98.57
540	**	81	9612	0.84	99.41
555		0	9612	0.00	99.41
570	*	27	9639	0.28	99.69
585		23	9662	0.24	99.93
600		0	9662	0.00	99.93
615		7	9669	0.07	100.00

400 800 1200

Frequency

Grade=8 Form=6

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
180	**	78	78	0.81	0.81
195		0	78	0.00	0.81
210		2	80	0.02	0.83
225		0	80	0.00	0.83
240		2	82	0.02	0.85
255		0	82	0.00	0.85
270		4	86	0.04	0.89
285		13	99	0.13	1.03
300	*	33	132	0.34	1.37
315	*	53	185	0.55	1.92
330	**	108	293	1.12	3.03
345	***	171	464	1.77	4.80
360	********	390	854	4.04	8.84
375	**********	539	1393	5.58	14.42
390	**********	1014	2407	10.50	24.92
405	**********	1062	3469	11.00	35.92
420	**********	1125	4594	11.65	47.57
435	**********	1316	5910	13.63	61.19
450	**********	922	6832	9.55	70.74
465	**********	950	7782	9.84	80.58
480	**********	825	8607	8.54	89.12
495	********	379	8986	3.92	93.04
510	*****	263	9249	2.72	95.77
525	****	176	9425	1.82	97.59
540	***	146	9571	1.51	99.10
555	*	56	9627	0.58	99.68
570		0	9627	0.00	99.68
585		23	9650	0.24	99.92
600		0	9650	0.00	99.92
615		0	9650	0.00	99.92
630		7	9657	0.07	99.99
645		0	9657	0.00	99.99
660		1	9658	0.01	100.00

--------+-------+-------+--

400 800 1200

Frequency

Grade=8 Form=7

Scale Score		Freq	Cum. Freq	Percent	Cum. Percent
195	*	63	63	0.66	0.66
210		0	63	0.00	0.66
225		1	64	0.01	0.67
240		0	64	0.00	0.67
255		1	65	0.01	0.68
270		1	66	0.01	0.69
285		9	75	0.09	0.79
300		7	82	0.07	0.86
315	*	32	114	0.34	1.20
330	*	29	143	0.30	1.50
345	***	173	316	1.81	3.31
360	****	191	507	2.00	5.32
375	********	398	905	4.17	9.49
390	**********	858	1763	9.00	18.49
405	**********	895	2658	9.39	27.88
420	**********	1076	3734	11.29	39.17
435	**********	1183	4917	12.41	51.58
450	**********	1327	6244	13.92	65.50
465	**********	959	7203	10.06	75.56
480	**********	863	8066	9.05	84.61
495	**********	715	8781	7.50	92.11
510	*****	260	9041	2.73	94.84
525	****	216	9257	2.27	97.10
540	***	148	9405	1.55	98.66
555	*	74	9479	0.78	99.43
570		0	9479	0.00	99.43
585	*	32	9511	0.34	99.77
600		0	9511	0.00	99.77
615		19	9530	0.20	99.97
630		0	9530	0.00	99.97
645		3	9533	0.03	100.00

APPENDIX B: THE 2003 *MSA-READING* CLASSICAL AND IRT ITEM PARAMETERS

Table B.1 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 3 Form 1

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
2	*SR*	0.91	0.20	-2.33	0.04	1.08	1.52
5	*SR*	0.79	0.44	-1.10	0.03	0.95	0.87
6	*SR*	0.58	0.41	0.15	0.03	1.06	1.05
9	*SR*	0.77	0.39	-0.93	0.03	1.02	0.96
11	*SR*	0.43	0.36	0.93	0.03	1.12	1.15
15	*SR*	0.78	0.33	-1.08	0.03	1.06	1.22
18	*SR*	0.50	0.40	0.45	0.03	1.07	1.09
20	*SR*	0.19	0.32	2.29	0.03	1.06	1.26
23	*SR*	0.62	0.43	-0.08	0.03	1.02	1.01
30	*SR*	0.41	0.44	1.00	0.03	1.01	1.06
31	*SR*	0.63	0.31	-0.11	0.03	1.16	1.19
32	*SR*	0.56	0.45	0.24	0.02	1.00	0.98
34	*SR*	0.62	0.52	-0.15	0.03	0.92	0.85
36	*SR*	0.71	0.47	-0.56	0.03	0.95	0.87
41	*SR*	0.61	0.56	-0.03	0.03	0.86	0.81
44	*SR*	0.87	0.42	-1.80	0.04	0.89	0.75
49	*SR*	0.61	0.37	0.00	0.03	1.10	1.18
55	*SR*	0.60	0.53	0.03	0.03	0.91	0.86
56	*SR*	0.43	0.25	0.93	0.03	1.25	1.41
57	*SR*	0.78	0.42	-1.06	0.03	0.98	1.04
58	*SR*	0.82	0.52	-1.43	0.03	0.84	0.63
59	*SR*	0.75	0.55	-0.91	0.03	0.84	0.73
61	*SR*	0.47	0.52	0.65	0.03	0.94	0.93
69	*SR*	0.58	0.47	-0.40	0.03	0.98	0.95
70	*SR*	0.42	0.47	0.59	0.03	0.99	1.00
72	*BCR*	0.35	0.60	1.74	0.02	0.93	0.92
73	*SR*	0.50	0.45	0.55	0.02	1.01	1.02
76	*SR*	0.63	0.44	-0.15	0.03	1.00	1.02
77	*SR*	0.90	0.38	-2.25	0.04	0.91	0.88
78	*BCR*	0.33	0.58	1.68	0.02	1.08	1.07
79	*SR*	0.48	0.47	0.62	0.03	1.00	1.00
81	*BCR*	0.43	0.51	0.80	0.02	1.01	1.02
82	*SR*	0.68	0.37	-0.39	0.03	1.07	1.12
84	*BCR*	0.43	0.64	0.90	0.02	0.92	0.91
85	*SR*	0.43	0.51	0.90	0.03	0.92	0.94
88	*SR*	0.55	0.43	0.27	0.03	1.03	1.04
90	*SR*	0.59	0.48	0.03	0.03	0.97	0.93

Table B.2 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 3 Form 2

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
2	*SR*	0.91	0.18	-2.27	0.04	1.09	1.53
5	*SR*	0.80	0.44	-1.10	0.03	0.92	0.81
6	*SR*	0.58	0.40	0.15	0.02	1.07	1.05
9	*SR*	0.77	0.41	-0.93	0.03	1.01	0.98
11	*SR*	0.44	0.37	0.93	0.02	1.10	1.14
15	*SR*	0.78	0.31	-1.08	0.03	1.11	1.29
18	*SR*	0.51	0.41	0.45	0.02	1.05	1.06
20	*SR*	0.20	0.33	2.29	0.03	1.05	1.25
23	*SR*	0.64	0.44	-0.08	0.02	0.99	0.96
30	*SR*	0.42	0.46	1.00	0.02	0.98	1.04
31	*SR*	0.64	0.31	-0.11	0.03	1.15	1.19
32	*SR*	0.57	0.45	0.24	0.02	1.00	0.97
34	*SR*	0.63	0.53	-0.15	0.03	0.90	0.82
36	*SR*	0.71	0.47	-0.56	0.03	0.95	0.88
41	*SR*	0.63	0.57	-0.03	0.02	0.84	0.78
44	*SR*	0.87	0.43	-1.76	0.03	0.89	0.73
49	*SR*	0.61	0.39	0.00	0.02	1.06	1.08
55	*SR*	0.61	0.51	0.03	0.02	0.92	0.87
56	*SR*	0.44	0.24	0.93	0.02	1.25	1.41
57	*SR*	0.77	0.41	-1.06	0.03	1.01	1.12
58	*SR*	0.83	0.53	-1.44	0.03	0.83	0.64
59	*SR*	0.75	0.56	-0.91	0.03	0.84	0.71
61	*SR*	0.47	0.49	0.65	0.02	0.98	0.97
69	*SR*	0.59	0.46	-0.43	0.03	0.98	0.95
70	*SR*	0.43	0.45	0.59	0.03	1.01	1.03
72	*BCR*	0.51	0.69	0.53	0.02	0.83	0.83
75	*BCR*	0.40	0.59	1.24	0.02	0.93	0.93
76	*SR*	0.63	0.48	-0.13	0.03	0.97	0.92
77	*SR*	0.47	0.39	0.71	0.02	1.07	1.11
79	*SR*	0.74	0.48	-0.80	0.03	0.93	0.87
80	*SR*	0.43	0.40	0.90	0.02	1.06	1.10
81	*BCR*	0.38	0.50	1.45	0.02	1.01	1.02
82	*SR*	0.48	0.28	0.69	0.02	1.20	1.28
83	*SR*	0.52	0.37	0.46	0.02	1.10	1.12
85	*SR*	0.57	0.53	0.23	0.02	0.91	0.88
87	*BCR*	0.41	0.55	1.18	0.02	1.01	1.02
89	*SR*	0.57	0.43	0.15	0.02	1.02	1.08

Table B.3 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 3 Form 3

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
2	*SR*	0.91	0.18	-2.23	0.04	1.09	1.71
5	*SR*	0.80	0.44	-1.10	0.03	0.91	0.81
6	*SR*	0.57	0.40	0.15	0.02	1.05	1.03
9	*SR*	0.78	0.39	-0.93	0.03	0.99	0.91
11	*SR*	0.44	0.39	0.93	0.02	1.08	1.12
15	*SR*	0.78	0.32	-1.08	0.03	1.10	1.34
18	*SR*	0.52	0.41	0.45	0.02	1.04	1.06
20	*SR*	0.21	0.33	2.29	0.03	1.10	1.26
23	*SR*	0.62	0.45	-0.08	0.02	1.00	0.99
30	*SR*	0.43	0.46	1.00	0.03	0.98	1.03
31	*SR*	0.64	0.31	-0.11	0.03	1.13	1.18
32	*SR*	0.58	0.45	0.24	0.02	0.99	0.95
34	*SR*	0.63	0.51	-0.15	0.03	0.91	0.83
36	*SR*	0.71	0.47	-0.56	0.03	0.95	0.85
41	*SR*	0.62	0.58	-0.03	0.02	0.83	0.77
44	*SR*	0.88	0.43	-1.83	0.04	0.89	0.73
49	*SR*	0.60	0.39	0.00	0.02	1.07	1.10
55	*SR*	0.60	0.52	0.03	0.02	0.91	0.85
56	*SR*	0.44	0.25	0.93	0.02	1.24	1.39
57	*SR*	0.78	0.42	-1.06	0.03	0.97	1.10
58	*SR*	0.83	0.53	-1.47	0.03	0.83	0.68
59	*SR*	0.75	0.57	-0.91	0.03	0.85	0.73
61	*SR*	0.46	0.51	0.65	0.02	0.95	0.96
69	*SR*	0.59	0.47	-0.44	0.03	0.97	0.92
70	*SR*	0.42	0.46	0.59	0.03	1.01	1.06
71	*SR*	0.72	0.50	-0.59	0.03	0.91	0.85
72	*BCR*	0.47	0.47	1.03	0.02	1.30	1.43
74	*SR*	0.71	0.39	-0.59	0.03	1.03	1.04
75	*BCR*	0.45	0.62	0.77	0.02	0.92	0.90
76	*SR*	0.47	0.38	0.73	0.02	1.07	1.13
80	*SR*	0.82	0.36	-1.49	0.03	0.99	1.10
82	*SR*	0.62	0.41	-0.11	0.03	1.04	1.07
83	*SR*	0.62	0.38	-0.09	0.03	1.08	1.15
84	*BCR*	0.46	0.52	1.81	0.02	1.11	1.20
87	*BCR*	0.28	0.58	2.89	0.02	0.93	0.93
88	*SR*	0.69	0.52	-0.59	0.03	0.90	0.85
90	*SR*	0.78	0.52	-1.26	0.03	0.86	0.67

Table B.4 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 3 Form 4

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
2	*SR*	0.91	0.18	-2.32	0.04	1.09	1.77
5	*SR*	0.79	0.46	-1.10	0.03	0.94	0.81
6	*SR*	0.58	0.40	0.15	0.03	1.08	1.06
9	*SR*	0.77	0.40	-0.93	0.03	1.01	0.99
11	*SR*	0.43	0.39	0.93	0.02	1.09	1.13
15	*SR*	0.78	0.33	-1.08	0.03	1.08	1.33
18	*SR*	0.51	0.42	0.45	0.03	1.05	1.06
20	*SR*	0.20	0.36	2.29	0.03	1.04	1.20
23	*SR*	0.63	0.46	-0.08	0.03	0.99	0.98
30	*SR*	0.43	0.46	1.00	0.03	1.00	1.05
31	*SR*	0.64	0.31	-0.11	0.03	1.16	1.20
32	*SR*	0.58	0.46	0.24	0.02	1.01	0.97
34	*SR*	0.62	0.54	-0.15	0.03	0.91	0.82
36	*SR*	0.71	0.47	-0.56	0.03	0.97	0.90
41	*SR*	0.63	0.58	-0.03	0.03	0.84	0.76
44	*SR*	0.87	0.44	-1.81	0.04	0.88	0.70
49	*SR*	0.60	0.38	0.00	0.03	1.11	1.18
55	*SR*	0.61	0.53	0.03	0.03	0.90	0.84
56	*SR*	0.44	0.27	0.93	0.03	1.24	1.38
57	*SR*	0.78	0.42	-1.06	0.03	0.97	1.07
58	*SR*	0.82	0.53	-1.41	0.03	0.84	0.65
59	*SR*	0.76	0.57	-0.91	0.03	0.83	0.70
61	*SR*	0.47	0.51	0.65	0.03	0.95	0.95
69	*SR*	0.59	0.45	-0.46	0.03	1.01	1.00
70	*SR*	0.43	0.45	0.59	0.03	1.03	1.06
73	*SR*	0.71	0.43	-0.56	0.03	1.00	0.98
74	*SR*	0.54	0.42	0.39	0.02	1.05	1.08
75	*BCR*	0.21	0.58	4.42	0.02	0.90	0.89
78	*BCR*	0.23	0.60	3.75	0.02	0.92	0.91
79	*SR*	0.57	0.53	0.15	0.03	0.92	0.93
80	*SR*	0.65	0.46	-0.30	0.03	0.99	0.98
81	*BCR*	0.37	0.46	3.02	0.03	1.02	1.03
82	*SR*	0.62	0.52	-0.03	0.03	0.93	0.88
84	*BCR*	0.47	0.53	2.03	0.02	1.02	1.04
86	*SR*	0.39	0.39	1.13	0.03	1.09	1.14
88	*SR*	0.51	0.32	0.49	0.02	1.20	1.30
90	*SR*	0.72	0.55	-0.66	0.03	0.86	0.74

Table B.5 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 3 Form 5

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
2	*SR*	0.91	0.19	-2.29	0.04	1.10	1.61
5	*SR*	0.80	0.44	-1.10	0.03	0.91	0.81
6	*SR*	0.58	0.41	0.15	0.02	1.08	1.08
9	*SR*	0.77	0.39	-0.93	0.03	1.01	0.97
11	*SR*	0.43	0.39	0.93	0.02	1.10	1.14
15	*SR*	0.78	0.31	-1.08	0.03	1.09	1.31
18	*SR*	0.50	0.41	0.45	0.03	1.08	1.12
20	*SR*	0.20	0.34	2.29	0.03	1.05	1.20
23	*SR*	0.63	0.46	-0.08	0.03	1.00	0.98
30	*SR*	0.43	0.48	1.00	0.03	0.98	1.03
31	*SR*	0.63	0.31	-0.11	0.03	1.19	1.23
32	*SR*	0.58	0.45	0.24	0.02	1.01	0.98
34	*SR*	0.62	0.53	-0.15	0.03	0.91	0.84
36	*SR*	0.71	0.46	-0.56	0.03	0.97	0.89
41	*SR*	0.63	0.59	-0.03	0.03	0.83	0.75
44	*SR*	0.87	0.43	-1.82	0.04	0.89	0.73
49	*SR*	0.61	0.38	0.00	0.03	1.10	1.15
55	*SR*	0.60	0.54	0.03	0.03	0.92	0.87
56	*SR*	0.43	0.26	0.93	0.02	1.27	1.44
57	*SR*	0.78	0.41	-1.06	0.03	0.99	1.17
58	*SR*	0.83	0.53	-1.49	0.03	0.83	0.62
59	*SR*	0.76	0.56	-0.91	0.03	0.82	0.69
61	*SR*	0.46	0.50	0.65	0.03	0.98	0.98
69	*SR*	0.59	0.45	-0.47	0.03	1.00	1.02
70	*SR*	0.43	0.45	0.59	0.03	1.04	1.06
71	*SR*	0.66	0.48	-0.27	0.03	0.97	0.93
72	*BCR*	0.34	0.60	1.83	0.02	0.93	0.92
73	*SR*	0.84	0.46	-1.53	0.03	0.88	0.78
74	*SR*	0.55	0.57	0.30	0.02	0.87	0.82
75	*BCR*	0.42	0.61	0.97	0.02	1.08	1.06
79	*SR*	0.68	0.47	-0.50	0.03	0.97	1.02
83	*SR*	0.71	0.47	-0.60	0.03	0.96	0.95
84	*BCR*	0.37	0.51	1.79	0.02	1.08	1.07
85	*SR*	0.35	0.43	1.37	0.03	1.02	1.09
86	*SR*	0.55	0.44	0.29	0.02	1.04	1.06
87	*BCR*	0.52	0.68	0.83	0.02	0.83	0.83
89	*SR*	0.58	0.49	0.09	0.03	0.97	0.93

Table B.6 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 3 Form 6

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
2	*SR*	0.90	0.20	-2.23	0.04	1.11	1.83
5	*SR*	0.79	0.46	-1.10	0.03	0.93	0.82
6	*SR*	0.58	0.42	0.15	0.03	1.08	1.05
9	*SR*	0.78	0.41	-0.93	0.03	0.99	0.93
11	*SR*	0.44	0.39	0.93	0.03	1.12	1.17
15	*SR*	0.77	0.33	-1.08	0.03	1.12	1.35
18	*SR*	0.50	0.41	0.45	0.03	1.10	1.14
20	*SR*	0.20	0.33	2.29	0.03	1.07	1.28
23	*SR*	0.63	0.45	-0.08	0.03	1.02	1.02
30	*SR*	0.43	0.47	1.00	0.03	1.02	1.08
31	*SR*	0.64	0.32	-0.11	0.03	1.17	1.22
32	*SR*	0.57	0.46	0.24	0.03	1.03	1.01
34	*SR*	0.62	0.55	-0.15	0.03	0.91	0.83
36	*SR*	0.70	0.48	-0.56	0.03	0.97	0.88
41	*SR*	0.62	0.57	-0.03	0.03	0.87	0.82
44	*SR*	0.86	0.44	-1.73	0.03	0.89	0.77
49	*SR*	0.61	0.39	0.00	0.03	1.12	1.17
55	*SR*	0.61	0.55	0.03	0.03	0.90	0.84
56	*SR*	0.44	0.24	0.93	0.03	1.32	1.52
57	*SR*	0.78	0.41	-1.06	0.03	0.98	1.25
58	*SR*	0.82	0.54	-1.44	0.03	0.83	0.64
59	*SR*	0.75	0.58	-0.91	0.03	0.84	0.70
61	*SR*	0.48	0.50	0.65	0.03	1.00	0.99
69	*SR*	0.59	0.49	-0.45	0.03	0.99	0.98
70	*SR*	0.43	0.47	0.59	0.03	1.04	1.06
71	*SR*	0.48	0.49	0.70	0.03	0.99	0.99
72	*BCR*	0.49	0.61	0.96	0.02	0.90	0.89
73	*SR*	0.77	0.47	-0.98	0.03	0.95	0.85
74	*SR*	0.79	0.46	-1.11	0.03	0.93	0.89
75	*BCR*	0.17	0.56	4.00	0.02	0.95	0.92
79	*SR*	0.81	0.46	-1.45	0.03	0.92	0.86
82	*SR*	0.65	0.55	-0.24	0.03	0.90	0.86
83	*SR*	0.49	0.40	0.62	0.03	1.11	1.18
84	*BCR*	0.31	0.57	3.56	0.02	1.01	1.00
87	*BCR*	0.40	0.62	1.81	0.02	0.90	0.89
88	*SR*	0.54	0.57	0.33	0.03	0.89	0.83
89	*SR*	0.70	0.47	-0.63	0.03	0.98	0.97

Table B.7 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 3 Form 7

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
2	*SR*	0.90	0.21	-2.10	0.04	1.08	1.52
5	*SR*	0.80	0.45	-1.10	0.03	0.96	0.82
6	*SR*	0.57	0.41	0.15	0.03	1.08	1.06
9	*SR*	0.79	0.37	-0.93	0.03	1.03	0.98
11	*SR*	0.43	0.39	0.93	0.03	1.07	1.11
15	*SR*	0.79	0.31	-1.08	0.03	1.10	1.31
18	*SR*	0.52	0.42	0.45	0.03	1.05	1.06
20	*SR*	0.20	0.32	2.29	0.03	1.02	1.19
23	*SR*	0.64	0.42	-0.08	0.03	1.04	1.05
30	*SR*	0.43	0.45	1.00	0.03	0.99	1.04
31	*SR*	0.63	0.33	-0.11	0.03	1.17	1.24
32	*SR*	0.58	0.46	0.24	0.03	1.00	0.96
34	*SR*	0.63	0.53	-0.15	0.03	0.93	0.85
36	*SR*	0.71	0.47	-0.56	0.03	1.00	0.91
51	*SR*	0.60	0.59	-0.03	0.03	0.87	0.80
54	*SR*	0.82	0.49	-1.44	0.04	0.86	0.67
59	*SR*	0.65	0.35	0.00	0.03	1.08	1.15
65	*SR*	0.66	0.52	0.03	0.03	0.87	0.82
76	*SR*	0.50	0.26	0.93	0.03	1.24	1.38
77	*SR*	0.84	0.31	-1.06	0.03	0.92	1.05
78	*SR*	0.89	0.46	-1.86	0.04	0.85	0.62
79	*SR*	0.80	0.52	-0.91	0.03	0.78	0.64
81	*SR*	0.53	0.49	0.65	0.03	0.96	0.95
89	*SR*	0.78	0.45	-0.98	0.03	0.95	0.91
90	*SR*	0.58	0.48	0.59	0.03	0.98	0.97
42	*BCR*	0.45	0.52	1.00	0.02	1.04	1.04
43	*SR*	0.46	0.46	0.84	0.03	0.99	1.00
46	*SR*	0.66	0.44	-0.16	0.03	1.01	1.05
47	*SR*	0.89	0.41	-1.99	0.04	0.90	0.93
48	*BCR*	0.32	0.60	1.93	0.02	1.03	1.02
49	*SR*	0.48	0.45	0.75	0.03	1.01	1.03
66	*BCR*	0.41	0.51	1.26	0.02	1.03	1.04
67	*SR*	0.72	0.36	-0.47	0.03	1.08	1.12
69	*BCR*	0.35	0.60	1.74	0.02	0.91	0.90
70	*SR*	0.47	0.50	0.84	0.03	0.93	0.95
73	*SR*	0.58	0.42	0.20	0.03	1.04	1.07
75	*SR*	0.62	0.46	0.00	0.03	1.00	0.99

Table B.8 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 5 Form 1

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.63	0.51	0.08	0.03	0.92	0.88
4	*SR*	0.55	0.44	0.51	0.02	1.00	1.00
5	*SR*	0.54	0.41	0.56	0.02	1.03	1.03
6	*SR*	0.58	0.55	0.37	0.02	0.88	0.82
7	*SR*	0.63	0.48	0.08	0.03	0.96	0.92
9	*SR*	0.87	0.31	-1.52	0.03	1.02	1.05
10	*SR*	0.80	0.50	-0.96	0.03	0.89	0.75
11	*SR*	0.82	0.28	-1.11	0.03	1.08	1.35
13	*SR*	0.82	0.29	-1.14	0.03	1.07	1.51
14	*SR*	0.79	0.33	-0.91	0.03	1.07	1.17
16	*SR*	0.75	0.46	-0.71	0.03	0.98	1.00
17	*SR*	0.71	0.51	-0.47	0.03	0.95	0.87
18	*SR*	0.80	0.54	-1.22	0.03	0.88	0.72
19	*SR*	0.68	0.44	-0.37	0.03	1.05	1.11
20	*SR*	0.66	0.42	-0.25	0.03	1.08	1.18
23	*SR*	0.55	0.38	0.49	0.02	1.07	1.08
25	*SR*	0.68	0.47	-0.19	0.03	0.95	0.88
28	*SR*	0.60	0.30	0.28	0.02	1.16	1.23
31	*SR*	0.56	0.47	0.45	0.02	0.96	0.95
33	*SR*	0.75	0.48	-0.65	0.03	0.93	0.87
34	*SR*	0.37	0.41	1.44	0.03	1.02	1.05
37	*SR*	0.64	0.42	-0.02	0.03	1.05	1.14
41	*SR*	0.66	0.47	-0.38	0.03	1.01	0.97
45	*SR*	0.45	0.44	0.69	0.03	1.05	1.10
49	*SR*	0.61	0.47	-0.92	0.03	0.88	0.75
51	*BCR*	0.52	0.58	0.66	0.02	1.00	1.00
52	*SR*	0.56	0.32	0.48	0.02	1.14	1.25
53	*SR*	0.69	0.51	-0.23	0.03	0.90	0.82
54	*SR*	0.83	0.53	-1.17	0.03	0.83	0.60
55	*BCR*	0.41	0.64	1.59	0.02	0.90	0.90
57	*SR*	0.80	0.50	-0.95	0.03	0.89	0.71
61	*SR*	0.65	0.42	-0.03	0.03	1.01	1.00
63	*SR*	0.49	0.45	0.80	0.02	0.99	0.97
65	*SR*	0.56	0.25	0.47	0.02	1.22	1.30
67	*BCR*	0.47	0.64	1.08	0.02	0.87	0.87
69	*BCR*	0.29	0.49	2.50	0.02	1.05	1.05
70	*SR*	0.52	0.38	0.65	0.02	1.16	1.23

Table B.9 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 5 Form 2

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.64	0.50	0.08	0.02	0.92	0.90
4	*SR*	0.56	0.44	0.51	0.02	1.00	1.01
5	*SR*	0.54	0.42	0.56	0.02	1.02	1.02
6	*SR*	0.58	0.54	0.37	0.02	0.89	0.84
7	*SR*	0.64	0.49	0.08	0.02	0.95	0.90
9	*SR*	0.86	0.33	-1.44	0.03	1.02	1.04
10	*SR*	0.79	0.51	-0.96	0.03	0.90	0.77
11	*SR*	0.82	0.26	-1.11	0.03	1.11	1.41
13	*SR*	0.81	0.30	-1.14	0.03	1.13	1.51
14	*SR*	0.79	0.33	-0.91	0.03	1.08	1.24
16	*SR*	0.75	0.47	-0.71	0.03	0.97	0.96
17	*SR*	0.71	0.51	-0.47	0.03	0.94	0.86
18	*SR*	0.80	0.56	-1.22	0.03	0.83	0.69
19	*SR*	0.68	0.45	-0.37	0.03	1.05	1.13
20	*SR*	0.65	0.44	-0.25	0.03	1.08	1.13
23	*SR*	0.57	0.37	0.49	0.02	1.07	1.09
25	*SR*	0.69	0.46	-0.19	0.03	0.95	0.86
28	*SR*	0.59	0.32	0.33	0.02	1.15	1.19
31	*SR*	0.56	0.48	0.45	0.02	0.96	0.93
33	*SR*	0.76	0.48	-0.65	0.03	0.90	0.87
34	*SR*	0.38	0.41	1.44	0.02	1.02	1.08
37	*SR*	0.64	0.42	-0.03	0.03	1.05	1.11
41	*SR*	0.65	0.47	-0.38	0.03	1.02	1.00
45	*SR*	0.44	0.46	0.69	0.03	1.03	1.04
49	*SR*	0.61	0.48	-0.93	0.03	0.90	0.76
53	*SR*	0.91	0.36	-1.98	0.04	0.95	0.85
55	*BCR*	0.50	0.61	1.13	0.02	0.88	0.87
56	*BCR*	0.46	0.53	1.27	0.02	0.98	0.97
58	*SR*	0.85	0.42	-1.47	0.03	0.94	0.85
59	*SR*	0.63	0.44	0.06	0.02	1.01	0.99
60	*SR*	0.71	0.46	-0.41	0.03	0.97	0.93
62	*SR*	0.76	0.49	-0.66	0.03	0.91	0.81
63	*SR*	0.75	0.25	-0.57	0.03	1.15	1.50
64	*BCR*	0.35	0.57	1.60	0.02	1.08	1.09
67	*BCR*	0.31	0.53	1.84	0.02	1.05	1.09
68	*SR*	0.80	0.47	-0.97	0.03	0.93	0.80
70	*SR*	0.74	0.37	-0.71	0.03	1.04	1.08

Table B.10 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 5 Form 3

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.65	0.50	0.08	0.02	0.90	0.92
4	*SR*	0.54	0.45	0.51	0.02	0.98	0.98
5	*SR*	0.54	0.41	0.56	0.02	1.02	1.00
6	*SR*	0.58	0.54	0.37	0.02	0.88	0.84
7	*SR*	0.64	0.45	0.08	0.02	0.97	0.93
9	*SR*	0.86	0.33	-1.44	0.03	1.01	0.99
10	*SR*	0.80	0.51	-0.96	0.03	0.90	0.75
11	*SR*	0.83	0.26	-1.11	0.03	1.06	1.36
13	*SR*	0.81	0.30	-1.14	0.03	1.10	1.39
14	*SR*	0.78	0.34	-0.91	0.03	1.06	1.22
16	*SR*	0.76	0.45	-0.71	0.03	0.95	0.98
17	*SR*	0.71	0.50	-0.47	0.03	0.93	0.87
18	*SR*	0.81	0.55	-1.22	0.03	0.84	0.67
19	*SR*	0.68	0.44	-0.37	0.03	1.03	1.12
20	*SR*	0.65	0.44	-0.25	0.03	1.04	1.10
23	*SR*	0.57	0.39	0.49	0.02	1.04	1.04
25	*SR*	0.68	0.47	-0.19	0.03	0.95	0.89
28	*SR*	0.59	0.35	0.31	0.02	1.14	1.19
31	*SR*	0.56	0.48	0.45	0.02	0.95	0.93
33	*SR*	0.75	0.48	-0.65	0.03	0.92	0.84
34	*SR*	0.37	0.49	1.44	0.02	1.02	1.08
37	*SR*	0.64	0.41	-0.07	0.03	1.04	1.13
41	*SR*	0.67	0.48	-0.48	0.03	0.98	0.95
45	*SR*	0.45	0.45	0.69	0.03	1.03	1.06
49	*SR*	0.62	0.48	-0.92	0.03	0.90	0.74
51	*BCR*	0.48	0.51	1.08	0.02	0.97	0.97
54	*SR*	0.75	0.40	-0.60	0.03	1.00	0.95
55	*BCR*	0.40	0.58	1.82	0.02	0.94	0.94
57	*SR*	0.51	0.42	0.69	0.02	1.01	1.02
59	*SR*	0.57	0.24	0.40	0.02	1.22	1.31
60	*SR*	0.78	0.43	-0.83	0.03	0.96	0.91
61	*SR*	0.90	0.32	-1.82	0.04	0.97	1.01
65	*SR*	0.55	0.41	0.52	0.02	1.02	1.05
66	*SR*	0.83	0.34	-1.14	0.03	1.02	1.03
67	*BCR*	0.64	0.54	-0.30	0.02	1.01	1.01
68	*SR*	0.42	0.32	1.18	0.02	1.13	1.22
69	*BCR*	0.58	0.58	0.26	0.02	0.94	0.94

Table B.11 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 5 Form 4

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.64	0.51	0.08	0.02	0.90	0.87
4	*SR*	0.55	0.45	0.51	0.02	0.98	0.97
5	*SR*	0.54	0.42	0.56	0.02	1.00	0.99
6	*SR*	0.59	0.55	0.37	0.02	0.86	0.82
7	*SR*	0.65	0.48	0.08	0.02	0.93	0.87
9	*SR*	0.86	0.31	-1.44	0.03	1.03	1.06
10	*SR*	0.79	0.50	-0.96	0.03	0.93	0.78
11	*SR*	0.82	0.27	-1.11	0.03	1.10	1.33
13	*SR*	0.81	0.32	-1.14	0.03	1.09	1.33
14	*SR*	0.78	0.35	-0.91	0.03	1.08	1.21
16	*SR*	0.76	0.47	-0.71	0.03	0.95	0.93
17	*SR*	0.71	0.51	-0.47	0.03	0.94	0.87
18	*SR*	0.80	0.55	-1.22	0.03	0.87	0.75
19	*SR*	0.67	0.45	-0.37	0.03	1.04	1.08
20	*SR*	0.65	0.44	-0.25	0.03	1.05	1.09
23	*SR*	0.56	0.38	0.49	0.02	1.05	1.07
25	*SR*	0.68	0.46	-0.19	0.03	0.95	0.88
28	*SR*	0.58	0.29	0.35	0.02	1.15	1.20
31	*SR*	0.56	0.46	0.45	0.02	0.96	0.96
33	*SR*	0.77	0.48	-0.65	0.03	0.89	0.83
34	*SR*	0.38	0.41	1.44	0.02	1.02	1.06
37	*SR*	0.64	0.41	-0.03	0.03	1.05	1.09
41	*SR*	0.66	0.47	-0.42	0.03	1.00	0.98
45	*SR*	0.46	0.47	0.69	0.03	1.00	1.01
49	*SR*	0.63	0.48	-0.97	0.03	0.89	0.75
51	*BCR*	0.59	0.58	0.27	0.02	0.97	0.97
53	*SR*	0.79	0.48	-0.87	0.03	0.91	0.81
55	*BCR*	0.68	0.56	-0.81	0.02	0.95	0.95
57	*SR*	0.71	0.40	-0.36	0.03	1.01	1.05
58	*SR*	0.69	0.36	-0.24	0.03	1.06	1.10
60	*SR*	0.91	0.40	-2.02	0.04	0.90	0.81
61	*SR*	0.65	0.29	0.02	0.02	1.14	1.15
63	*SR*	0.67	0.26	-0.12	0.03	1.17	1.31
64	*BCR*	0.43	0.53	1.42	0.02	0.98	0.99
65	*SR*	0.79	0.37	-0.86	0.03	1.01	1.06
66	*SR*	0.57	0.34	0.38	0.02	1.10	1.20
69	*BCR*	0.34	0.57	1.84	0.02	1.01	1.01

Table B.12 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 5 Form 5

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.64	0.51	0.08	0.02	0.92	0.89
4	*SR*	0.56	0.46	0.51	0.02	0.99	0.98
5	*SR*	0.54	0.42	0.56	0.02	1.03	1.03
6	*SR*	0.59	0.55	0.37	0.02	0.88	0.82
7	*SR*	0.65	0.48	0.08	0.02	0.95	0.90
9	*SR*	0.87	0.33	-1.50	0.03	1.01	1.09
10	*SR*	0.80	0.50	-0.96	0.03	0.89	0.76
11	*SR*	0.82	0.27	-1.11	0.03	1.10	1.50
13	*SR*	0.82	0.31	-1.14	0.03	1.06	1.47
14	*SR*	0.78	0.34	-0.91	0.03	1.10	1.23
16	*SR*	0.76	0.45	-0.71	0.03	0.98	0.98
17	*SR*	0.72	0.53	-0.47	0.03	0.90	0.81
18	*SR*	0.80	0.53	-1.22	0.03	0.85	0.73
19	*SR*	0.68	0.44	-0.37	0.03	1.06	1.16
20	*SR*	0.65	0.43	-0.25	0.03	1.08	1.19
23	*SR*	0.56	0.38	0.49	0.02	1.07	1.10
25	*SR*	0.69	0.47	-0.19	0.03	0.95	0.88
28	*SR*	0.58	0.32	0.39	0.02	1.15	1.21
31	*SR*	0.56	0.48	0.45	0.02	0.96	0.93
33	*SR*	0.76	0.49	-0.65	0.03	0.91	0.87
34	*SR*	0.37	0.41	1.44	0.03	1.02	1.11
37	*SR*	0.65	0.43	-0.11	0.03	1.04	1.10
41	*SR*	0.66	0.48	-0.44	0.03	1.00	1.01
45	*SR*	0.45	0.45	0.69	0.03	1.05	1.08
49	*SR*	0.62	0.48	-0.97	0.03	0.89	0.77
51	*BCR*	0.46	0.53	1.28	0.02	1.02	1.02
54	*SR*	0.74	0.42	-0.50	0.03	1.00	0.99
56	*BCR*	0.51	0.53	0.71	0.02	0.98	0.98
57	*SR*	0.64	0.30	0.06	0.03	1.17	1.38
58	*SR*	0.72	0.48	-0.44	0.03	0.95	0.88
59	*SR*	0.57	0.50	0.41	0.02	0.95	0.91
61	*SR*	0.79	0.45	-0.83	0.03	0.94	0.86
63	*SR*	0.75	0.50	-0.62	0.03	0.89	0.78
65	*SR*	0.54	0.40	0.55	0.02	1.05	1.07
67	*BCR*	0.46	0.49	1.11	0.02	1.03	1.03
68	*SR*	0.63	0.48	0.12	0.02	0.95	0.94
69	*BCR*	0.42	0.47	1.06	0.02	1.02	1.03

Table B.13 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 5 Form 6

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.64	0.50	0.08	0.03	0.93	0.90
4	*SR*	0.55	0.45	0.51	0.02	1.00	1.00
5	*SR*	0.54	0.43	0.56	0.02	1.02	1.02
6	*SR*	0.58	0.54	0.37	0.02	0.89	0.84
7	*SR*	0.64	0.49	0.08	0.03	0.95	0.89
9	*SR*	0.87	0.31	-1.51	0.03	1.03	1.13
10	*SR*	0.80	0.50	-0.96	0.03	0.88	0.75
11	*SR*	0.82	0.27	-1.11	0.03	1.09	1.42
13	*SR*	0.82	0.30	-1.14	0.03	1.10	1.40
14	*SR*	0.78	0.33	-0.91	0.03	1.11	1.31
16	*SR*	0.76	0.46	-0.71	0.03	0.97	0.95
17	*SR*	0.72	0.51	-0.47	0.03	0.93	0.83
18	*SR*	0.80	0.55	-1.22	0.03	0.85	0.71
19	*SR*	0.69	0.43	-0.37	0.03	1.05	1.11
20	*SR*	0.66	0.43	-0.25	0.03	1.07	1.19
23	*SR*	0.56	0.38	0.49	0.02	1.07	1.09
25	*SR*	0.68	0.48	-0.19	0.03	0.97	0.88
28	*SR*	0.57	0.32	0.40	0.02	1.16	1.24
31	*SR*	0.56	0.47	0.45	0.02	0.97	0.96
33	*SR*	0.76	0.50	-0.65	0.03	0.89	0.79
34	*SR*	0.37	0.41	1.44	0.03	1.03	1.09
37	*SR*	0.64	0.43	-0.04	0.03	1.05	1.10
41	*SR*	0.67	0.46	-0.48	0.03	1.02	0.98
45	*SR*	0.45	0.45	0.69	0.03	1.04	1.09
49	*SR*	0.62	0.48	-0.99	0.04	0.89	0.73
52	*SR*	0.78	0.36	-0.81	0.03	1.04	1.11
55	*BCR*	0.41	0.46	2.13	0.02	1.05	1.06
56	*BCR*	0.46	0.62	1.80	0.02	0.90	0.88
57	*SR*	0.67	0.48	-0.11	0.03	0.95	0.92
59	*SR*	0.86	0.38	-1.51	0.03	0.98	0.88
60	*SR*	0.63	0.40	0.08	0.03	1.06	1.09
62	*SR*	0.84	0.40	-1.31	0.03	0.96	0.98
63	*SR*	0.56	0.44	0.47	0.02	1.01	1.01
66	*SR*	0.73	0.29	-0.49	0.03	1.14	1.56
67	*BCR*	0.33	0.65	2.13	0.02	0.87	0.86
69	*BCR*	0.26	0.55	2.91	0.02	0.96	0.96
70	*SR*	0.87	0.42	-1.94	0.04	0.88	0.64

Table B.14 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 5 Form 7

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.66	0.50	0.08	0.03	0.92	0.90
4	*SR*	0.55	0.46	0.51	0.03	0.99	0.98
5	*SR*	0.55	0.41	0.56	0.03	1.03	1.04
6	*SR*	0.58	0.55	0.37	0.03	0.89	0.83
7	*SR*	0.65	0.48	0.08	0.03	0.94	0.90
9	*SR*	0.88	0.34	-1.49	0.04	1.00	0.95
10	*SR*	0.82	0.48	-0.96	0.03	0.89	0.73
11	*SR*	0.83	0.26	-1.11	0.03	1.11	1.36
13	*SR*	0.85	0.27	-1.14	0.03	1.02	1.43
14	*SR*	0.80	0.28	-0.91	0.03	1.13	1.38
16	*SR*	0.80	0.39	-0.71	0.03	0.95	0.96
17	*SR*	0.76	0.44	-0.47	0.03	0.94	0.85
18	*SR*	0.87	0.47	-1.22	0.03	0.78	0.64
19	*SR*	0.75	0.34	-0.37	0.03	1.03	1.09
20	*SR*	0.72	0.33	-0.25	0.03	1.07	1.16
23	*SR*	0.56	0.39	0.49	0.03	1.06	1.07
25	*SR*	0.69	0.47	-0.19	0.03	0.98	0.92
38	*SR*	0.63	0.25	0.24	0.03	1.20	1.27
46	*SR*	0.59	0.49	0.45	0.03	0.94	0.93
48	*SR*	0.78	0.47	-0.65	0.03	0.89	0.79
49	*SR*	0.38	0.41	1.44	0.03	1.01	1.05
42	*SR*	0.75	0.35	-0.45	0.03	1.05	1.19
61	*SR*	0.79	0.38	-0.74	0.03	1.01	1.00
65	*SR*	0.59	0.42	0.69	0.03	1.01	1.01
69	*SR*	0.90	0.40	-1.82	0.04	0.91	0.82
26	*BCR*	0.52	0.56	0.65	0.02	1.02	1.02
27	*SR*	0.56	0.34	0.59	0.03	1.11	1.19
28	*SR*	0.71	0.52	-0.26	0.03	0.89	0.81
29	*SR*	0.84	0.53	-1.18	0.03	0.84	0.61
30	*BCR*	0.43	0.66	1.45	0.02	0.86	0.86
32	*SR*	0.78	0.54	-0.77	0.03	0.86	0.70
51	*SR*	0.65	0.42	0.09	0.03	1.01	0.99
53	*SR*	0.51	0.45	0.83	0.03	0.98	0.97
55	*SR*	0.58	0.28	0.45	0.03	1.17	1.25
57	*BCR*	0.49	0.64	0.90	0.02	0.90	0.90
59	*BCR*	0.32	0.51	2.33	0.02	1.04	1.04
60	*SR*	0.53	0.29	0.74	0.03	1.18	1.25

Table B.15 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 8 Form 1

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.67	0.47	0.07	0.03	0.98	0.96
6	*SR*	0.50	0.32	0.96	0.02	1.15	1.21
8	*SR*	0.59	0.46	0.51	0.03	0.99	0.97
9	*SR*	0.89	0.40	-1.57	0.04	0.95	0.80
22	*SR*	0.97	0.29	-3.60	0.09	0.96	1.23
23	*SR*	0.57	0.26	0.64	0.02	1.23	1.33
25	*SR*	0.81	0.26	-0.80	0.03	1.13	1.43
26	*SR*	0.61	0.38	0.39	0.03	1.09	1.12
29	*SR*	0.72	0.46	-0.19	0.03	0.97	0.96
31	*SR*	0.65	0.46	0.20	0.03	0.98	0.99
32	*SR*	0.50	0.23	0.98	0.02	1.26	1.41
33	*SR*	0.63	0.42	0.29	0.03	0.98	0.94
35	*SR*	0.76	0.33	-0.46	0.03	1.04	1.09
37	*SR*	0.72	0.43	-0.24	0.03	1.01	1.03
38	*SR*	0.76	0.46	-0.49	0.03	0.96	0.92
40	*SR*	0.70	0.52	-0.12	0.03	0.91	0.85
41	*SR*	0.82	0.32	-1.02	0.03	1.09	1.35
44	*SR*	0.69	0.53	-0.16	0.03	0.92	0.88
46	*SR*	0.68	0.53	-0.22	0.03	0.93	0.86
48	*SR*	0.64	0.57	-0.05	0.03	0.84	0.73
49	*SR*	0.65	0.48	-0.16	0.03	1.01	1.00
50	*SR*	0.65	0.51	-0.18	0.03	0.96	0.92
51	*BCR*	0.46	0.70	1.32	0.02	0.82	0.82
52	*SR*	0.57	0.49	0.58	0.02	0.95	0.92
56	*BCR*	0.45	0.73	1.35	0.02	0.79	0.78
57	*SR*	0.62	0.49	0.36	0.03	0.96	0.94
59	*SR*	0.41	0.26	1.40	0.03	1.23	1.37
60	*SR*	0.46	0.47	1.16	0.02	0.98	0.99
61	*SR*	0.78	0.29	-0.62	0.03	1.10	1.30
62	*SR*	0.48	0.56	1.06	0.02	0.86	0.84
64	*BCR*	0.57	0.61	0.19	0.02	0.93	0.93
66	*SR*	0.88	0.35	-1.57	0.04	0.97	1.13
67	*SR*	0.76	0.33	-0.45	0.03	1.08	1.36
68	*BCR*	0.51	0.60	0.67	0.02	0.96	0.95

Table B.16 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 8 Form 2

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.67	0.47	0.07	0.03	0.96	0.93
6	*SR*	0.51	0.34	0.86	0.02	1.13	1.21
8	*SR*	0.59	0.46	0.51	0.02	0.98	0.97
9	*SR*	0.88	0.44	-1.57	0.04	0.94	0.77
22	*SR*	0.97	0.36	-3.46	0.08	0.97	0.83
23	*SR*	0.57	0.26	0.64	0.02	1.22	1.34
25	*SR*	0.80	0.29	-0.80	0.03	1.13	1.41
26	*SR*	0.61	0.39	0.39	0.02	1.07	1.08
29	*SR*	0.70	0.47	-0.19	0.03	0.97	0.99
31	*SR*	0.63	0.45	0.20	0.02	1.01	1.01
32	*SR*	0.48	0.27	0.98	0.02	1.23	1.38
33	*SR*	0.63	0.43	0.29	0.02	0.97	0.94
35	*SR*	0.76	0.35	-0.46	0.03	1.00	1.03
37	*SR*	0.72	0.45	-0.24	0.03	0.98	0.99
38	*SR*	0.75	0.48	-0.49	0.03	0.95	0.92
40	*SR*	0.69	0.53	-0.12	0.03	0.91	0.84
41	*SR*	0.81	0.38	-1.01	0.03	1.05	1.25
44	*SR*	0.68	0.54	-0.16	0.03	0.92	0.89
46	*SR*	0.68	0.53	-0.31	0.03	0.94	0.88
48	*SR*	0.63	0.57	-0.10	0.03	0.86	0.74
49	*SR*	0.64	0.47	-0.16	0.03	1.02	1.05
50	*SR*	0.63	0.50	-0.18	0.03	1.00	1.00
52	*SR*	0.78	0.44	-0.70	0.03	0.97	0.96
53	*SR*	0.53	0.40	0.74	0.02	1.06	1.09
55	*BCR*	0.55	0.64	0.31	0.02	0.91	0.91
56	*BCR*	0.55	0.57	0.38	0.02	1.04	1.04
58	*SR*	0.72	0.42	-0.28	0.03	1.02	1.02
59	*SR*	0.64	0.39	0.19	0.03	1.08	1.12
61	*SR*	0.63	0.42	0.26	0.02	1.04	1.05
62	*SR*	0.53	0.47	0.78	0.02	0.97	0.95
64	*BCR*	0.52	0.66	0.61	0.02	0.90	0.90
66	*SR*	0.81	0.47	-0.95	0.03	0.93	0.84
68	*BCR*	0.44	0.70	1.26	0.02	0.85	0.85
69	*SR*	0.49	0.45	0.93	0.02	1.00	1.01

Table B.17 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 8 Form 3

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.67	0.47	0.07	0.03	0.98	0.93
6	*SR*	0.51	0.32	0.89	0.02	1.13	1.19
8	*SR*	0.59	0.46	0.51	0.02	0.98	0.96
9	*SR*	0.89	0.41	-1.57	0.04	0.93	0.86
22	*SR*	0.97	0.35	-3.47	0.08	0.96	0.88
23	*SR*	0.57	0.26	0.64	0.02	1.21	1.30
25	*SR*	0.80	0.27	-0.80	0.03	1.16	1.42
26	*SR*	0.61	0.37	0.39	0.02	1.08	1.09
29	*SR*	0.71	0.47	-0.19	0.03	0.96	0.93
31	*SR*	0.64	0.45	0.20	0.02	0.98	0.97
32	*SR*	0.50	0.25	0.98	0.02	1.21	1.34
33	*SR*	0.64	0.43	0.29	0.02	0.95	0.90
35	*SR*	0.76	0.34	-0.46	0.03	1.02	1.05
37	*SR*	0.72	0.44	-0.24	0.03	1.00	0.98
38	*SR*	0.75	0.48	-0.49	0.03	0.96	0.90
40	*SR*	0.69	0.51	-0.12	0.03	0.93	0.87
41	*SR*	0.81	0.36	-0.98	0.03	1.08	1.20
44	*SR*	0.69	0.54	-0.16	0.03	0.90	0.86
46	*SR*	0.69	0.52	-0.34	0.03	0.94	0.86
48	*SR*	0.64	0.56	-0.10	0.03	0.86	0.75
49	*SR*	0.65	0.47	-0.16	0.03	1.01	1.00
50	*SR*	0.65	0.51	-0.18	0.03	0.96	0.92
51	*BCR*	0.48	0.64	1.15	0.02	0.92	0.91
52	*SR*	0.87	0.47	-1.43	0.04	0.91	0.94
53	*SR*	0.75	0.31	-0.44	0.03	1.12	1.52
55	*BCR*	0.66	0.57	-0.40	0.02	1.00	1.00
58	*SR*	0.88	0.48	-1.60	0.04	0.90	0.76
60	*SR*	0.67	0.43	0.06	0.03	1.01	1.03
63	*SR*	0.83	0.48	-1.04	0.03	0.93	0.83
64	*BCR*	0.47	0.56	0.85	0.02	0.97	0.98
66	*SR*	0.62	0.56	0.30	0.02	0.93	0.90
68	*BCR*	0.33	0.63	1.76	0.02	0.89	0.89
69	*SR*	0.60	0.45	0.40	0.02	0.99	0.98
70	*SR*	0.66	0.29	0.05	0.03	1.18	1.44

Table B.18 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 8 Form 4

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.65	0.47	0.07	0.03	0.97	0.93
6	*SR*	0.50	0.32	0.87	0.02	1.13	1.18
8	*SR*	0.58	0.47	0.51	0.02	0.96	0.94
9	*SR*	0.88	0.43	-1.57	0.04	0.93	0.80
22	*SR*	0.97	0.38	-3.51	0.08	0.96	0.85
23	*SR*	0.55	0.30	0.64	0.02	1.16	1.24
25	*SR*	0.80	0.31	-0.80	0.03	1.08	1.27
26	*SR*	0.61	0.39	0.39	0.02	1.05	1.05
29	*SR*	0.71	0.48	-0.19	0.03	0.94	0.91
31	*SR*	0.64	0.46	0.20	0.02	0.97	0.96
32	*SR*	0.49	0.27	0.98	0.02	1.19	1.28
33	*SR*	0.62	0.45	0.29	0.02	0.93	0.88
35	*SR*	0.75	0.36	-0.46	0.03	1.00	1.02
37	*SR*	0.71	0.46	-0.24	0.03	0.98	0.96
38	*SR*	0.75	0.50	-0.49	0.03	0.93	0.86
40	*SR*	0.69	0.53	-0.12	0.03	0.90	0.88
41	*SR*	0.81	0.37	-0.98	0.03	1.06	1.17
44	*SR*	0.68	0.56	-0.16	0.03	0.89	0.84
46	*SR*	0.68	0.54	-0.32	0.03	0.92	0.82
48	*SR*	0.63	0.57	-0.05	0.03	0.84	0.73
49	*SR*	0.64	0.47	-0.16	0.03	1.02	1.01
50	*SR*	0.64	0.51	-0.18	0.03	0.95	0.90
54	*SR*	0.39	0.37	1.44	0.02	1.05	1.09
55	*BCR*	0.44	0.60	1.01	0.02	0.87	0.85
56	*BCR*	0.58	0.62	0.43	0.02	0.94	0.93
58	*SR*	0.48	0.50	1.00	0.02	0.90	0.89
59	*SR*	0.68	0.31	-0.04	0.03	1.13	1.29
60	*SR*	0.62	0.30	0.26	0.02	1.16	1.29
61	*SR*	0.63	0.27	0.26	0.02	1.19	1.34
62	*SR*	0.53	0.54	0.75	0.02	0.88	0.84
63	*SR*	0.69	0.25	-0.08	0.03	1.19	1.37
64	*BCR*	0.53	0.55	0.35	0.02	0.97	0.97
68	*BCR*	0.50	0.60	1.17	0.02	0.98	0.99
69	*SR*	0.85	0.38	-1.34	0.03	0.99	1.02

Table B.19 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 8 Form 5

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.66	0.46	0.07	0.03	0.98	0.96
6	*SR*	0.51	0.33	0.87	0.02	1.13	1.21
8	*SR*	0.59	0.47	0.51	0.02	0.96	0.94
9	*SR*	0.89	0.42	-1.57	0.04	0.90	0.78
22	*SR*	0.97	0.35	-3.64	0.08	0.96	1.41
23	*SR*	0.56	0.27	0.64	0.02	1.21	1.31
25	*SR*	0.80	0.30	-0.80	0.03	1.12	1.29
26	*SR*	0.61	0.37	0.39	0.02	1.08	1.11
29	*SR*	0.71	0.48	-0.19	0.03	0.94	0.91
31	*SR*	0.64	0.45	0.20	0.02	0.99	1.01
32	*SR*	0.49	0.26	0.98	0.02	1.21	1.33
33	*SR*	0.62	0.44	0.29	0.02	0.95	0.92
35	*SR*	0.75	0.35	-0.46	0.03	1.02	1.06
37	*SR*	0.71	0.45	-0.24	0.03	1.00	0.98
38	*SR*	0.75	0.49	-0.49	0.03	0.93	0.88
40	*SR*	0.70	0.51	-0.12	0.03	0.91	0.87
41	*SR*	0.81	0.38	-1.01	0.03	1.06	1.18
44	*SR*	0.68	0.54	-0.16	0.03	0.92	0.91
46	*SR*	0.68	0.53	-0.31	0.03	0.93	0.86
48	*SR*	0.63	0.56	-0.07	0.03	0.85	0.74
49	*SR*	0.64	0.49	-0.16	0.03	0.99	0.99
50	*SR*	0.63	0.51	-0.18	0.03	0.98	0.98
51	*BCR*	0.40	0.58	1.64	0.02	0.95	0.95
52	*SR*	0.86	0.45	-1.36	0.03	0.94	0.80
53	*SR*	0.85	0.40	-1.28	0.03	1.01	1.00
55	*BCR*	0.27	0.52	2.48	0.02	1.00	1.01
57	*SR*	0.66	0.47	0.09	0.03	0.98	0.93
58	*SR*	0.72	0.49	-0.28	0.03	0.95	0.90
61	*SR*	0.64	0.41	0.17	0.02	1.04	1.11
62	*SR*	0.83	0.48	-1.12	0.03	0.92	0.80
63	*SR*	0.41	0.37	1.35	0.02	1.06	1.13
65	*BCR*	0.51	0.59	0.72	0.02	0.97	0.97
68	*BCR*	0.42	0.58	1.46	0.02	0.98	0.98
70	*SR*	0.75	0.35	-0.58	0.03	1.07	1.30

Table B.20 The 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 8 Form 6

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.68	0.44	0.07	0.03	0.97	0.92
6	*SR*	0.51	0.31	0.88	0.02	1.13	1.2
8	*SR*	0.59	0.44	0.51	0.02	0.99	0.98
9	*SR*	0.89	0.42	-1.57	0.04	0.9	0.79
22	*SR*	0.97	0.31	-3.56	0.08	0.99	1.14
23	*SR*	0.57	0.24	0.64	0.02	1.21	1.31
25	*SR*	0.80	0.28	-0.80	0.03	1.13	1.28
26	*SR*	0.62	0.36	0.39	0.02	1.08	1.11
29	*SR*	0.71	0.45	-0.19	0.03	0.97	0.99
31	*SR*	0.64	0.45	0.20	0.03	0.98	0.99
32	*SR*	0.51	0.25	0.98	0.02	1.21	1.31
33	*SR*	0.64	0.42	0.29	0.02	0.95	0.9
35	*SR*	0.76	0.34	-0.46	0.03	1.01	1.02
37	*SR*	0.72	0.44	-0.24	0.03	0.98	0.97
38	*SR*	0.76	0.48	-0.49	0.03	0.94	0.89
40	*SR*	0.70	0.51	-0.12	0.03	0.91	0.84
41	*SR*	0.81	0.37	-0.94	0.03	1.05	1.17
44	*SR*	0.69	0.54	-0.16	0.03	0.91	0.87
46	*SR*	0.68	0.52	-0.23	0.03	0.93	0.85
48	*SR*	0.64	0.56	-0.06	0.03	0.85	0.73
49	*SR*	0.65	0.46	-0.16	0.03	1.03	1.04
50	*SR*	0.65	0.51	-0.18	0.03	0.95	0.9
51	*BCR*	0.55	0.55	0.39	0.02	0.97	0.97
53	*SR*	0.52	0.39	0.85	0.02	1.05	1.08
54	*SR*	0.82	0.31	-0.95	0.03	1.06	1.27
55	*BCR*	0.55	0.59	0.59	0.02	0.95	0.96
57	*SR*	0.74	0.47	-0.34	0.03	0.95	0.91
60	*SR*	0.68	0.41	0.00	0.03	1.02	1.06
62	*SR*	0.74	0.36	-0.37	0.03	1.06	1.07
65	*BCR*	0.39	0.56	2.13	0.02	0.93	0.93
66	*SR*	0.91	0.44	-1.82	0.04	0.9	0.79
67	*SR*	0.63	0.38	0.29	0.02	1.06	1.03
68	*BCR*	0.49	0.55	1.04	0.02	1.03	1.03
69	*SR*	0.74	0.40	-0.42	0.03	1.02	1.06

Table B.21 2003 *MSA-Reading* Classical and IRT Item Parameters: Grade 8 Form 7

Item Number	Item Type	P-Value	Point-Biserial	Rasch Difficulty	SEM	Ms. Infit	Ms. Outfit
3	*SR*	0.68	0.46	0.07	0.04	1.00	1.00
6	*SR*	0.52	0.33	1.02	0.04	1.13	1.19
8	*SR*	0.63	0.46	0.51	0.04	0.98	0.93
9	*SR*	0.89	0.39	-1.57	0.06	1.00	0.84
22	*SR*	0.97	0.26	-3.30	0.12	0.98	1.66
23	*SR*	0.58	0.23	0.64	0.04	1.25	1.40
25	*SR*	0.81	0.26	-0.80	0.05	1.21	1.49
26	*SR*	0.64	0.37	0.39	0.04	1.08	1.14
29	*SR*	0.73	0.44	-0.19	0.04	1.00	1.08
41	*SR*	0.65	0.45	0.20	0.04	1.02	1.05
42	*SR*	0.54	0.25	0.98	0.04	1.22	1.33
43	*SR*	0.64	0.36	0.29	0.04	1.05	1.03
45	*SR*	0.82	0.29	-0.46	0.04	0.92	0.95
47	*SR*	0.76	0.39	-0.24	0.04	0.99	0.97
48	*SR*	0.80	0.44	-0.49	0.04	0.90	0.87
50	*SR*	0.72	0.50	-0.12	0.04	0.95	0.91
61	*SR*	0.84	0.27	-0.91	0.05	1.11	1.39
64	*SR*	0.74	0.49	-0.16	0.04	0.92	0.85
66	*SR*	0.84	0.45	-0.97	0.05	0.94	0.79
68	*SR*	0.78	0.51	-0.50	0.04	0.85	0.69
69	*SR*	0.76	0.39	-0.16	0.04	0.96	1.00
70	*SR*	0.78	0.44	-0.18	0.04	0.87	0.81
31	*BCR*	0.42	0.70	1.66	0.02	0.80	0.80
32	*SR*	0.58	0.52	0.70	0.04	0.91	0.88
36	*BCR*	0.40	0.70	1.87	0.02	0.82	0.81
37	*SR*	0.62	0.51	0.48	0.04	0.93	0.89
39	*SR*	0.43	0.27	1.43	0.04	1.20	1.32
40	*SR*	0.43	0.49	1.48	0.04	0.93	0.93
51	*SR*	0.77	0.33	-0.37	0.04	1.07	1.31
52	*SR*	0.48	0.54	1.22	0.04	0.86	0.83
54	*BCR*	0.56	0.62	0.33	0.03	0.91	0.91
56	*SR*	0.87	0.35	-1.45	0.06	0.97	1.06
57	*SR*	0.73	0.30	-0.26	0.04	1.10	1.36
58	*BCR*	0.44	0.53	1.58	0.03	1.00	1.00

APPENDIX C: THE 2003 *MSA-READING* BLUEPRINTS

Table C.1 The 2003 *MSA-Reading* Blueprint: Grade 3

Reading Standards	*SAT10*	*MSA*
Goal: Students examine, construct and extend the meaning of a variety of self-selected and assigned text (traditional and electronic) by applying a range of reading strategies and analytic techniques		
Expectation: General Reading Process		
Indicators of Learning		
Phonemic Awareness		
1. distinguish beginning, middle, and ending sounds in words	3 *SR*	
Assessment Limits:		
☐ recognizing within words the structural elements required for decoding		
2. distinguish long and short vowel sounds	2 *SR*	
Assessment Limits:		
☐ recognizing and decoding vowel sounds words		
Word Study		
3. recognize compound words, contractions, common abbreviations and common syntax	1 *SR*	
Assessment Limits:		
☐ creating and understanding compound words		
☐ creating and understanding contractions in sentences		
☐ recognizing and appropriately using abbreviations in sentences		
4. use context to determine the meaning of words (semantics)	6 *SR*	
Assessment Limits:		
☐ demonstrating the ability to use context clues to assign meaning to unfamiliar words		

Reading Standards	*SAT10*	*MSA*
Acquisition and Application of New Vocabulary		
5. use prior word knowledge such as prefixes and suffixes to determine the meaning of words	2 *SR*	
Assessment Limits:		
☐ using prefixes and suffixes to assign meaning to words		
6. monitor texts for unknown words using sentence and word context to find meaning	1 *SR*	
Assessment Limits:		
☐ identifying unknown words and using the context to find meaning		
7. use prior knowledge of known words in unknown compound words to predict their meaning		
Assessment Limits:		
☐ predicting meaning using context and known words within a compound word		
Expectation: Informational Reading Process		
Indicators of Learning		
Comprehension of Text		
8. state a purpose for reading and identify who would use the text	1 *SR*	5 *SR*
Assessment Limits:		
☐ identifying appropriate purpose for reading		
☐ identifying who would use the text		
9. relate prior knowledge and experience to literal and inferential information found in text	1 *SR*	4 *SR* 1 *BCR*
Assessment Limits:		
☐ using prior knowledge and experience to gain literal and inferential understanding of the text		

Reading Standards	*SAT10*	*MSA*
10. ask clarifying questions concerning essential textual elements of exposition (e.g., why, how) and demonstrate comprehension by pinpointing answers in text		5 *SR* 4 *BCR*
Assessment Limits:		
☐ identifying answers to clarifying questions in the text		
11. determine author's purpose		1 *SR* 2 *BCR*
Assessment Limits:		
☐ identifying author's purpose based on the text		
12. extract appropriate and significant information from text, including problems and solutions, major points, and identify central ideas in the text	1 *SR*	3 *SR* 2 *BCR*
Assessment Limits:		
☐ identifying significant details in the text		
13. distinguish between cause and effect, and fact and opinion		5 *SR* 2 *BCR*
Assessment Limits:		
☐ Identifying the cause and effect relationships in the text		
☐ identifying facts in the text		
☐ identifying opinions based on the text		
14. compare and contrast information in different texts		1 *SR* 1 *BCR*
Assessment Limits:		
☐ identifying related information in multiple texts and comparing and contrasting this information		
15. restate information from the text	1 *SR*	2 *SR*
Assessment Limits:		
☐ identifying and restating information in the text in written form		

Reading Standards	*SAT10*	*MSA*
Evaluation of Text		
16. explain the connections between illustrations and text and how they support the text		5 *SR*
Assessment Limits:		
☐ using illustrations and text to gain and share information in written form		
17. identify common text features	1 *SR*	3 *SR* 1 *BCR*
Assessment Limits:		
☐ applying knowledge of text features to understanding the text		
18. evaluate the author's use of various techniques to influence the reader's feelings and attitudes		5 *SR* 3 *BCR*
Assessment Limits:		
☐ identifying author's techniques used to influence the readers feelings and attitudes		
19. evaluate the appropriateness of a title		3 *SR* 2 *BCR*
Assessment Limits:		
☐ identifying the title and stating appropriateness based on explicit and implying information		
Expectation: Literary Reading Process		
Indicators of Learning		
Characteristics of Literary Genres		
20. identify the characteristics that define the literary genres of poetry, drama, and prose	3 *SR*	6 *SR* 1 *BCR*
Assessment Limits:		
☐ recognizing poetry, plays or narrative works		

Reading Standards	*SAT10*	*MSA*
Comprehension, of Literary Text		
21. identify the main ideas in fictional words and relate them to prior experience or the experiences of others	1 *SR*	5 *SR* 8 *BCR*
Assessment Limits:		
☐ identifying the main idea		
22. identify the elements of plot, character, and setting in literary works	1 *SR*	23 *SR* 4 *BCR*
Assessment Limits:		
☐ identifying the plot, character and setting		
23. explain the connections between illustrations and text and how they support text		7 *SR* 2 *BCR*
24. summarize stories, plays, and poems		1 *SR* 2 *BCR*
Assessment Limits:		
☐ summarizing significant ideas in a text		
Comparison of Literary Text from Diverse Cultures		
25. identify basic plots of classic myths, folk tales, legends, and fables from around the world and connect them to prior experience or the experiences of others		1 *BCR*
Assessment Limits:		
☐ identifying the plot in texts or across texts that relate to many cultures (such as honesty, or friendship)		

Table C.2 The 2003 *MSA-Reading* Blueprint: Grade 5

Reading Standards	*SAT10*	*MSA*
Goal: Students examine, construct and extend the meaning of a variety of self-selected and assigned text (traditional and electronic) by applying a range of reading strategies and analytic techniques		
Expectation: General Reading Process (State Accountability)		
Indicators of Learning		
Concepts of Print and Structural Features of Text		
1. identify and use common organizational structures such as comparison and contrast, cause and effect, and chronological order to gain meaning from text		
Assessment Limits:		
☐ identifying the following organizational elements (internal text structure) of both fiction and nonfiction:		
- comparison and contrast		
- cause and effect		
- chronological order		
☐ using organizational elements to gain meaning from text		
2. use glossaries, table of contents, chapter headings and subheadings, indexes, and sidebars to locate information in text		
Assessment Limits:		
☐ identifying the following organizational elements (external text structure) of both fiction and nonfiction:		
- glossaries		
- table of contents		
- chapter headings and subheadings		
- indexes		
- sidebars		
☐ using organizational elements to locate information		

Reading Standards	*SAT10*	*MSA*
Acquisition and Application of New Vocabulary		
3. use knowledge of word parts such as prefixes (e.g., un-, pre-, dis-), suffixes (e.g., -ful, -less)		
Assessment Limits:		
☐ using knowledge of prefixes and suffixes to understand unfamiliar vocabulary		
4. use knowledge of word relationships, including antonyms, synonyms, homographs, homophones, and idioms to determine the meaning of words and phrases	10 *SR*	
Assessment Limits:		
☐ using knowledge of the following to determine the meaning of unfamiliar words and phrases:		
- antonyms		
- synonyms		
- homographs		
- homophones		
- idioms		
5. apply such context clues as definition, example, comparison and contrast, cause and effect to discern word meanings	5 *SR*	
Assessment Limits:		
☐ applying the following context clues to discern word meaning:		
- definition		
- example		
- comparison and contrast		
- cause and effect		

Reading Standards	*SAT10*	*MSA*
Expectation: Informational Reading Process (State Accountability)		
Indicators of Learning		
Comprehension of Text		
6. use prior knowledge and ideas presented in texts to make and confirm predictions	1 *SR*	6 *SR* 4 *BCR*
Assessment Limits:		
☐ using prior knowledge and ideas presented to make predictions		
☐ confirming predictions based on text		
7. evaluate new information and hypotheses by testing them again known information and ideas		2 *SR*
Assessment Limits:		
☐ testing new information against known information and ideas		
☐ forming hypotheses		
☐ testing hypotheses against known information and ideas		
8. revise and clarify steps in a set of directions, instructions, or procedures		2 *SR* 3 *BCR*
Assessment Limits:		
☐ revising steps in a set of directions, instructions, or procedures		
☐ clarifying steps in a set of directions, instructions, or procedures		
9. know and use different focusing, monitoring, and assessing reading strategies (e.g., skimming and scanning) to comprehend text		3 *SR*
Assessment Limits:		
☐ knowing the following active reading strategies:		
- mark or highlight		
- connecting text to known information and ideas		
- ask questions		
- predict		
- visualize		
- clarify		
☐ using different active reading strategies to comprehend text		

Reading Standards	*SAT10*	*MSA*
10. evaluate the various functions of language used (to inform, to persuade) to comprehend text		
Assessment Limits:		
☐ identifying words or phrases that inform or persuade		
☐ testing the effectiveness of informational and persuasive language to comprehend text		
11. summarize text in a manner that reflects the main ideas and significant details (, and its underlying meaning)	1 *SR*	1 *SR* 2 *BCR*
Assessment Limits:		
☐ composing a summary that includes the main ideas and the significant details		
12. determine the author's purpose	1 *SR*	4 *SR* 1 *BCR*
Assessment Limits:		
☐ using information in the text to determine whether the author's purpose is to inform or persuade		
13. compare and contrast information in the text with prior knowledge		2 *SR* 1 *BCR*
Assessment Limits:		
☐ comparing and contrasting information with prior knowledge		
14. summarize the steps in text		5 *SR*
Assessment Limits:		
☐ composing a summary of the steps in a process		
15. reorganize information from the text into a different form (charts, drawings, or graphic organizers)		1 *SR*
Assessment Limits:		
☐ reorganizing information from text in the following formats:		
- chart		
- drawing		
- graphic organizer		

Reading Standards	*SAT10*	*MSA*
16. identify additional information needed		2 *SR*
Assessment Limits:		
☐ identifying additional information needed to comprehend text		
Evaluation of Text		
17. explain how the tone is reflected in the author's style		3 *SR* 4 *BCR*
Assessment Limits:		
☐ identifying the stylistic choices that affect the tone of a text		
- word choice		
- sentence structure and length		
- literary devices, such as figurative language, symbols, dialogue, and imagery		
18. distinguish relevant from irrelevant information contained within text and identify possible points of confusion	1 *SR*	1 *SR*
Assessment Limits:		
☐ distinguishing relevant from irrelevant information		
☐ identifying possible points of confusion		
19. distinguish among facts, supported inferences, and opinions in text	1 *SR*	8 *SR* 2 *BCR*
Assessment Limits:		
☐ identifying facts, supported inferences, and opinions		
☐ differentiating facts from supported inferences and opinions		
20. evaluate the usefulness of information		2 *SR* 1 *BCR*
Assessment Limits:		
☐ evaluating how textual information might be used		

Reading Standards	*SAT10*	*MSA*
Expectation: Literary Reading Process (State Accountability)		
Indicators of Learning		
Characteristics of Literary Genres		
21. distinguish the characteristics of fiction and non-fiction		3 *SR* 1 *BCR*
Assessment Limits:		
☐ distinguishing fictional from non-fictional text		
Comprehension of Literary Text		
22. determine the theme whether it is implied or state directly	1 *SR*	4 *SR* 4 *BCR*
Assessment Limits:		
☐ determining the theme as the author's main point		
23. identify the main incidents of a plot, their causes, how they influence future action, and how they are resolved	4 *SR*	19 *SR* 2 *BCR*
Assessment Limits:		
☐ identifying the main incidents of a plot as the following:		
- exposition		
- rising action		
- climax		
- falling action		
- resolution		
☐ identifying or inferring the causes of those incidents		
☐ explaining or predicting how these incidents influence future action		
☐ identifying how the plot is resolved		
24. analyze the influence of setting on the mood and meaning of the text		2 *SR* 2 *BCR*
Assessment Limits:		
☐ identifying mood as the feeling that a literary work gives to the reader		
☐ analyzing the influence of setting on the mood of a text		

Reading Standards	*SAT10*	*MSA*
25. summarize the text and identify the main story elements		7 *SR* 5 *BCR*
Assessment Limits:		
☐ composing a summary that identifies one or more story elements:		
- plot		
- setting		
- characterization		
- theme		
- point of view		
26. evaluate text for elements of realism or fantasy		2 *SR* 1 *BCR*
Assessment Limits:		
☐ distinguishing elements of realism and elements of fantasy		
27. identify and analyze the effects of sound in poetry (e.g., alliteration, assonance, consonance, rhythm, onomatopoeia, and rhyme scheme)		
Assessment Limits:		
☐ identifying sound in poetry created by the following figurative language:		
- alliteration		
- assonance		
- consonance		
- rhythm		
- onomatopoeia		
- rhyme scheme		
☐ analyzing the effects of sound in poetry created by figurative language		
28. evaluate the author's choice of title		4 *SR* 1 *BCR*
Assessment Limits:		
☐ evaluating the author's choice of title		

Reading Standards	*SAT10*	*MSA*
Comparison of Literary Text from Diverse Cultures		
29. compare and contrast tales from diverse cultures by tracing the exploits of one character type and connect them to prior experience or the experiences of others		1 *SR* 2 *BCR*
Assessment Limits:		
☐ identifying character types:		
- heroes and heroines		
- wise versus foolish humans and animals		
- good versus evil characters		
☐ comparing and contrasting character types		
☐ connecting a character's exploits to prior experience or the experiences of others		

Table C.3 The 2003 *MSA-Reading* Blueprint: Grade 8

Reading Standards	*SAT10*	*MSA*
Goal: Students examine, construct and extend the meaning of a variety of self-selected and assigned text (traditional and electronic) by applying a range of reading strategies and analytic techniques		
Expectation: General Reading Process		
Indicators of Learning		
Comprehension of Text		
1. draw inferences, conclusions or generalizations about text and support them with textual evidence and experience	4 *SR*	
Assessment Limits:		
❒ drawing inferences, conclusions, or generalizations based upon information in the text		
❒ supporting inferences, conclusions, or generalizations with expressed and/or implied information from the text or from the reader's own experience		
2. determine the author's purpose and identify and trace the development of an author's argument, viewpoint or perspective in text	1 *SR*	
Assessment Limits:		
❒ using information in the text to determine the author's purpose		
❒ identifying an author's argument, viewpoint, or perspective and citing evidence from the text to show how the author develops it		
3. summarize the text	4 *SR*	
Assessment Limits:		
❒ composing a summary that includes the main ideas and the significant details		
Evaluation of Text		
4. recognize instances of propaganda and persuasive techniques	3 *SR*	
Assessment Limits:		
❒ recognizing logical fallacies, such as red herring, appeals to readers' fear or pity, snob appeal, bandwagon approach, flattery, hasty generalizations, absolute statements, and stereotypes		

Reading Standards	*SAT10*	*MSA*
5. evaluate the usefulness, clarity, and internal consistency of the text's organizational structure		
Assessment Limits:		
☐ determining the usefulness, clarity, and consistency of the text's organizational structure in relation to the author's argument, viewpoint, or perspective		
6. assess the adequacy, accuracy, and appropriateness of an author's details to support claims and assertions, noting instances of bias and stereotyping		
Assessment Limits:		
☐ evaluating the quality of specific details that support the author's position or argument		
☐ identifying details that reflect author's bias and stereotyping		
Acquisition and Application of New Vocabulary		
7. use idioms, analogies, metaphors, and similes to infer the literal and figurative meaning of phrases in literary text	3 *SR*	
Assessment Limits:		
☐ determining the literal meaning of specific phrases in a literary text by analyzing features of language, including idiomatic expressions, analogies, and figures of speech (e.g. metaphors and similes)		
☐ drawing inferences about the figurative meaning of specific phrases in a literary text by analyzing features of language, including idiomatic expressions, analogies, and figures of speech (e.g. metaphors and similes)		
8. use knowledge of Greek, Latin, and Anglo-Saxon roots to understand content specific vocabulary		
Assessment Limits:		
☐ applying knowledge of Greek, Latin, and Anglo-Saxon roots and word families to define unfamiliar, content-specific vocabulary		
Expectation: Informational Reading Process		
Indicators of Learning		
Comprehension of Text		
9. compare and contrast information from different articles or procedures on the same topic	1 *SR*	3 *BCR*
Assessment Limits:		
☐ identifying similarities and differences in information across multiple texts that address the same topic		

Reading Standards	*SAT10*	*MSA*
10. identify and trace the development of an author's argument, viewpoint or perspective in text	1 *SR*	7 *SR* 4 *BCR*
Assessment Limits:		
☐ identifying an author's argument, viewpoint, or perspective		
☐ citing evidence from the text to illustrate the development of an author's argument, viewpoint, or perspective		
11. connect and clarify main ideas and concepts and identify their relationship to other sources, related topics, or prior experiences		5 *SR* 1 *BCR*
Assessment Limits:		
☐ making connections among main ideas and concepts in a text		
☐ determining the relationship between the main ideas and concepts in one text and the sources or related topics of another text or the reader's prior experience		
12. identify how someone would use the text to summarize in a manner that reflects the main ideas, significant details, its underlying meaning and explain the usefulness of the text		4 *SR*
Assessment Limits:		
☐ explaining the usefulness of a text for the reader of society		
13. summarize text in a manner that reflects the main ideas and significant details		5 *SR* 1 *BCR*
Assessment Limits:		
☐ composing a summary that includes the main ideas and the significant details		
14. compare and contrast information with prior knowledge		6 *BCR*
Assessment Limits:		
☐ identifying similarities and differences between information in a text and the readers' prior knowledge		
Evaluation of Text		
15. recognize instances of propaganda and persuasive techniques	2 *SR*	1 *SR* 1 *BCR*
Assessment Limits:		
☐ recognizing logical fallacies, such as red herring, appeals to readers' fear or pity, snob appeal, bandwagon approach, flattery, hasty generalizations, absolute statements, and stereotypes		

Reading Standards	*SAT10*	*MSA*
16. infer word meaning through identification and analysis of analogies and other word relationships		4 *SR*
Assessment Limits:		
☐ applying the characteristics of analogies and other word relationships to infer the meaning of an unfamiliar word		
17. use idioms, analogies and figures of speech to infer the literal and figurative meaning of phrases in literary text		7 *SR* 1 *BCR*
Assessment Limits:		
☐ determining the literal meaning of specific phrases in a literary text by analyzing features of language, including idiomatic expressions, analogies, and figures of speech		
☐ drawing inferences about the figurative meaning of specific phrases in a literary text by analyzing features of language, including idiomatic expressions, analogies, and figures of speech		
18. use knowledge of Greek, Latin, and Anglo-Saxon roots to understand content specific vocabulary		2 *SR*
Assessment Limits:		
☐ applying knowledge of Greek, Latin, and Anglo-Saxon roots and word families to define unfamiliar, content-specific vocabulary		
19. distinguish and explain the “shades of meaning” for related words		5 *SR*
Assessment Limits:		
☐ differentiating among the variations in meaning of related words (e.g. pretty, beautiful, attractive, cute, nice-looking, good-looking)		
20. identifying the connotation and denotation of new words and apply them in writing and speaking	1 *SR*	2 *SR* 1 *BCR*
Assessment Limits:		
☐ determining the intended an implied meaning and associations of unfamiliar words		
☐ applying newly acquired vocabulary in speech and writing		

Reading Standards	*SAT10*	*MSA*
Expectation: Literary Reading Process		
Indicators of Learning		
Characteristics of Literary Genres		
21. identify the author's message and explain how the characteristics of different forms of prose (e.g. short story, novel, essay) shape the meaning of the author's message	1 *SR*	5 *SR* 1 *BCR*
Assessment Limits:		
☐ drawing conclusions about the message of a literary text		
☐ drawing inferences about how the author uses the characteristics of different literary forms to express the message of a literary text		
Comprehension of Literary Text		
22. compare and contrast the ways similar themes are expressed in multiple literary works and explain how the theme represents a view or comment on life		4 *SR* 3 *BCR*
Assessment Limits:		
☐ identifying the ways in which similar themes are expressed across multiple literary works		
☐ explaining how a literary theme comments on life		
23. compare works that express a universal theme (e.g. good and evil), providing evidence to support the ideas		4 *SR* 2 *BCR*
Assessment Limits:		
☐ comparing multiple literary works that express a similar theme		
☐ supporting ideas about how authors use the elements of literature to express universal ideas		
24. identify elements of plot and characterization and analyze how the qualities of the central characters determine resolution of the conflict	1 *SR*	6 *SR* 3 *BCR*
Assessment Limits:		
☐ identifying elements of plot, including exposition, rising action, climax, falling action, resolution (Students will not be asked to label events in a plot)		
☐ identifying elements of characterization, including the character's thoughts, words, actions, and what others say about the character (Students will not be asked to label types of characterization)		
☐ analyzing the relationship between the qualities of the central characters and the resolution of the conflict		

Reading Standards	*SAT10*	*MSA*
25. analyze characterization as delineated through a character's thoughts, words, speech patterns, and actions	1 *SR*	7 *SR* 2 *BCR*
Assessment Limits:		
☐ drawing inferences about a literary character based on that character's thoughts, words, speech patterns, and actions		
☐ supporting inferences about characterization with evidence from the text		
26. explain how literary (e.g. figurative language) {simile, metaphor, hyperbole, personification, allusions, and imagery} create meaning for readers	2 *SR*	11 *SR* 4 *BCR*
Assessment Limits:		
☐ analyzing examples of literary elements (e.g., figurative language,) to determine how readers use them to create meaning from a text		
27. identify and trace the development of an author's argument, viewpoint or perspective in text		2 *SR* 2 *BCR*
Assessment Limits:		
☐ identifying an author's argument, viewpoint, or perspective in a literary work		
☐ supporting inferences about an author's argument, viewpoint, or perspective in a literary work with information from the text		
Comparison of Literary Text from Diverse Cultures		
28. compare and contrast the motivation and reactions of characters from different historical eras and/or cultures who confront similar challenges and situations and connect them to prior experience or the experiences of others		3 *SR* 1 *BCR*
Assessment Limits:		
☐ comparing and contrasting the motivation and behavior of literary characters confronting similar situations		
☐ connecting and contrasting the behavior of literary characters to the reader's prior knowledge and experience or to the prior knowledge or experience of others		