

**Review of the Algebra I/Mathematics
for the Technologies II
Field Test**

October 1, 2002

**Division of Accountability
South Carolina Education Oversight Committee**



Alignment Analysis of South Carolina Algebra I/Mathematics for the Technologies II
End of Course Field Test
October 1, 2002

Executive Summary

This report presents the findings from studies of the alignment with standards and the technical qualities of the Algebra I/Mathematics for the Technologies II end of course field test. The studies were conducted as authorized by Section 59-18-320 (A) of the Education Accountability Act, which requires that state assessments and end of course tests be evaluated for their alignment with the state standards, level of difficulty and validity, and for the ability to differentiate levels of achievement.

The Algebra I/Mathematics for the Technologies II end of course test is one of the tests composing the End of Course Examination Program (EOCEP) for benchmark high school courses. The Algebra I/Mathematics for the Technologies II end of course test is based on the SC Algebra I Academic Achievement Standards. Beginning with the 2003-2004 school year, the EOCEP test results will count for 20% of students' final grades for the benchmark courses.

Two studies of the Algebra I/Mathematics for the Technologies II field test were conducted for this report:

1. A committee composed of twelve South Carolina mathematics educators evaluated the alignment of the test items with the achievement standards at a meeting on July 30, 2002. The committee was composed of middle and high school Algebra I and Mathematics for the Technologies II teachers, mathematics supervisors, and college algebra faculty.
2. Three researchers at Michigan State University investigated the technical aspects of the field test and reported their findings to the Education Oversight Committee for this report.

The field test was found to be well aligned with the state Algebra I academic standards and the items reflected a moderate level of cognitive challenge. The field test also exhibited good technical characteristics. A small number (12 of 285) of the items on the field test forms were identified as needing additional technical analysis and possible revision. The test was found to be difficult for the sample of students who participated in the field test in spring 2002, especially for students enrolled in Mathematics for the Technologies II.

Recommendations

The Algebra I/Math for the Technologies II end of course field test appears to be well aligned with the academic standards, to be of high quality and to reflect high standards for performance. Based on these findings, it is recommended that the test be approved with the following recommendations for future improvement:

1. Review the items identified as having potential technical problems and revise as needed.
2. As proposed in the report on standard-setting made to the State Board of Education, monitor the performance of students on the Algebra I/Math for the Technologies II tests administered in the 2002-2003 school year for possible review and revision of the standards set on the field test data.
3. Examine the extent and effects of the use of calculators on the end of course tests administered in 2002-2003 and make needed revisions to the policy on the use of calculators for the test.

Alignment Analysis of South Carolina Algebra I/Mathematics for the Technologies II
End of Course Field Test
October 1, 2002

Background Information and Descriptions of Studies

This report summarizes the results from studies of the Algebra I/Math for the Technologies II End of Course field test administered in spring 2002. The studies were conducted under the auspices of the Education Oversight Committee (EOC) as part of its responsibilities listed in the Education Accountability Act of 1998 (EAA):

Section 59-18-320. (A) After the first statewide field test of the assessment program in each of the four academic areas, and after the field tests of the end of course assessments of benchmark courses, the Education Oversight Committee, established in Section 59-6-10, will review the state assessment program and the course assessments for alignment with the state standards, level of difficulty and validity, and for the ability to differentiate levels of achievement, and will make recommendations for needed changes, if any. The review will be provided to the State Board of Education, the State Department of Education, the Governor, the Senate Education Committee, and the House Education and Public Works Committee as soon as feasible after the field tests. The Department of Education will then report to the Education Oversight Committee no later than one month after receiving the reports on the changes made to the assessments to comply with the recommendations.

The Algebra I/Math for the Technologies II assessment is one of the assessments included in the End of Course Examination Program (EOCEP) for grades nine through twelve. The EOCEP was established in Section 59-18-310(B) of the Education Accountability Act, which states, "The statewide assessment program in the four academic areas shall include grades three through eight, an exit examination which is to be first administered in grade ten, and end of course tests for gateway courses in English/language arts, mathematics, science, and social studies for grades nine through twelve." Section 59-18-320 (C) states that, "After review and approval by the Education Oversight Committee, the end of course assessments of benchmark courses will be administered to all public school students as they complete each benchmark course." Algebra I and Math for the Technologies II were the first benchmark courses identified for test development. End of course tests in English I, Physical Science, and Biology I are also currently under development.

Algebra I and Math for the Technologies II are both based on the Algebra I standards in the SC Academic Achievement Standards for mathematics (see Appendix A). The Algebra standards are taught in a single academic year (or its equivalent on a block schedule) in the Algebra I course. Math for the Technologies II is the second of two one-year courses (or the block schedule equivalent) for Tech Prep students that also address the Algebra standards. The Algebra I course is typically taught to students in the eighth, ninth, or tenth grades, although the course is taken by some advanced seventh grade students. Some students in the eleventh or twelfth grades may also take Algebra I. Math for the Technologies II is typically taken by tenth graders, although students at other high school grades may also take this course.

The EOCEP tests are administered to students at the end of the benchmark course. Under State Board of Education Regulation 43-262.4, End of Course Tests, students' test results are to be included in the calculation of their grades for the course. Beginning with the 2003-2004 school year, EOCEP test results will be weighted 20% in the determination of students' final course grades.

Two studies of the Algebra I/Math for the Technologies II field test were conducted to evaluate the alignment and technical qualities of the test:

1. A committee composed of twelve South Carolina mathematics educators evaluated the alignment of the test items with the achievement standards at a meeting on July 30, 2002. The committee was composed of middle and high school Algebra I and Mathematics for the Technologies II teachers, mathematics supervisors, and college algebra faculty (Appendix B). The agenda and working documents for this meeting are listed in Appendix C.
2. Three researchers at Michigan State University investigated the technical aspects of the field test and reported their findings to the Education Oversight Committee (their report is in Appendix D).

The evaluation alignment committee members were provided copies of the standards, the field test items, and descriptors for the various levels of cognitive demands which test items make on students as they attempt to answer them. The reviewers worked in pairs as they examined a set of test items to determine the standards assessed and the cognitive demands made. The technical reviewers were provided a data file containing student responses to the field test items.

Results

The Algebra I/Math for the Technologies II field test was administered in spring 2002 to students enrolled in the courses in a statewide sample of middle- and high-schools. The numbers of students participating in the field test and their grade levels are listed in Table 1.

Table I
Numbers of Students Participating and Grade Levels
Algebra I/Math for the Technologies II Field Test, Spring 2002

Grade Level	Number (%)
7	82 (1.9)
8	1040 (24.3)
9	1826 (42.7)
10	934 (21.8)
11	230 (5.4)
12	101 (2.4)
Not Identified	65 (1.5)
Total	4278 (100.0)

Approximately two-thirds (67.5%) of the students reported they were enrolled in Algebra I, and 30.1% reported they were enrolled in Math for the Technologies II; the remaining 2.4% of students did not report their course enrollments.

The Algebra I/Math for the Technologies End of Course Test is composed of 50 multiple choice selected response items. There were 60 multiple choice items on each of the six field test forms administered in spring 2002. The use of only multiple choice items on the Algebra I/Math for the Technologies II test makes it possible to promptly score and report student scores for use in calculating course grades. The field tests were not timed, and students were permitted to use as much time as they needed to complete the test. Test administrators were asked to allot two to two and one-half hours for testing. In addition to answering the test items, students responded to a questionnaire regarding their use of calculators, the extent to which the field test items matched their algebra instruction, and their impressions about various aspects of the test and its administration.

Students were allowed to use calculators for the field test. Calculators could be used for computing, graphing, or both. Not all students used calculators. Approximately 30% of Algebra I students reported they did not use a calculator for the field test, while approximately 40% of Math for the Technologies II students reported they did not use a calculator.

Alignment Study

The Algebra I/Math for the Technologies II field test consisted of six test forms containing 60 items each. Fifteen of the 60 items were common to all forms, with the remaining 45 items per form unique to each form. The items on each field test form were examined by a pair of standards alignment evaluation committee members. Thus each pair of committee members evaluated the alignment with the standards of 60 test items. Each member of each pair of reviewers read and answered each test item, determined the standard(s) it assessed, and made a judgment about the level of cognitive demands made by the item based on a six-point scale related to Bloom's Taxonomy of Educational Objectives (Appendix E).

The purpose of the alignment determination was to document the extent to which the items on the field test addressed the Algebra I Standards. The committee's documentation provides information on the comprehensiveness of the test's coverage of the standards. The alignment committee's findings can be compared to the intended coverage of the standards listed in the test Blueprint (Appendix F). The test Blueprint specifies the design for the test and lists the ranges of the number of items intended to assess each of the standards. The comparison of the alignment committee's findings to the Algebra I/Math for the Technologies II Blueprint is presented in Table 2.

Table 2
Standards Found To Be Aligned With Test Items and Match To Test Blueprint
EOC Assessment and Standards Alignment Committee
Algebra I/Math for Technology II End of Course Field Test, July 30, 2002

Standard	Number of Items On Test (From Blueprint)	Percent of Items On Test (From Blueprint)	Percent of Items Identified by EOC Alignment Comm.*
I. Understanding Functions	19-21	38-42	42.9
IA. Relationships	4-6	8-12	11.2
IB. Linear & Quadratic Functions & Data Representation	4-6	8-12	10.6
IC. Generalizations, Algebraic Symbols, & Matrices	3-5	6-10	5.9
ID. Algebraic Expressions in Problem Solving Situations	5-7	10-14	15.3
II. Linear Functions	21-23	42-46	41.2
IIA. Representations	4-6	8-12	7.9
IIIB. Interpretations	7-9	14-18	19.4
IIIC. Equations & Inequalities	5-7	10-14	8.2
IIID. Systems of Linear Equations	2-4	4-8	5.6
III. Quadratic and Other Functions	7-9	14-18	15.9
IIIA. Quadratic Functions	4-6	8-12	9.7
IIIB. Other Functions	2-4	4-8	6.2

* More than one standard was identified for some items.

In general, Table 2 indicates that the percentages of items found by the alignment committee to address the standards are within the ranges specified in the test design listed in the test Blueprint. This finding provides supporting evidence that the field test is closely aligned with the state Algebra Standards.

It is important to note that some components of the Algebra Standards are not directly assessed in the end of course assessment. These components are not assessed with the end of course test because they are very difficult to assess with a paper and pencil, multiple choice test administered in a single day. It is more appropriate to assess these standards as part of daily classroom instruction and assessment. An example of a standard component not assessed in the end of course test is the statement, "gather and record data" in Standard I.A.2, "Gather and record data or use data sets to determine functional (systematic) relationships between quantities." Given the multiple choice format for the end of course test, it would not be practical to expect students to gather data for a problem, but it is practical to assess the students' ability to use data sets to determine relationships between quantities.

Alignment committee members also judged each item for the level of thinking or cognitive demand it would pose to students. The committee members individually assigned a number ranging from one to six that corresponds to the hierarchy of cognitive levels described by Bloom (Appendix E). The committee members' ratings are compiled in Table 3.

Table 3
Cognitive Levels of Items By Standard
Cognitive Levels Based on Taxonomy of Educational Objectives
EOC Assessment and Standards Alignment Committee
Algebra I/Math for Technology II End of Course Field Test, July 30, 2002

Standard		Median Cognitive Level*	Mean Cognitive Level	Range of Cognitive Levels (1-6)
I. Understanding Functions		3	2.83	1-5
	IA. Relationships	2	2.53	1-4
	IB. Linear & Quadratic Functions & Data Representation	3	3.15	2-5
	IC. Generalizations, Algebraic Symbols, & Matrices	3	3.02	2-5
	ID. Algebraic Expressions in Problem Solving Situations	3	2.68	1-4
II. Linear Functions		3	2.81	1-6
	IIA. Representations	3	2.73	1-5
	IIB. Interpretations	3	2.87	1-5
	IIC. Equations & Inequalities	3	2.61	1-4
	IID. Systems of Linear Equations	3	3.13	1-6
III. Quadratic and Other Functions		2	2.37	1-4
	IIIA. Quadratic Functions	2	2.39	1-4
	IIIB. Other Functions	2.5	2.35	1-4

* Bloom Benjamin S. and David R. Krathwohl. Taxonomy of Educational Objectives: The Classification of Educational Goals, by a committee of college and university examiners. [Handbook I: Cognitive Domain](#). New York: Longmans, Green, 1956.

The committee members judged that the midpoint (median) in cognitive levels for the items addressing Understanding Functions and Linear Functions was at level three, Application. The midpoint for Quadratic and Other Functions was lower at level two, Comprehension. However, in most cases the cognitive demands of the items addressing each standard covered a broad range. The items addressing Linear Functions had the broadest range of cognitive demands, and Quadratic and Other Functions had the narrowest. For most standards listed in Table 3 the mean rating was lower than the median rating, suggesting that there were relatively more items at lower cognitive levels for these standards than at higher levels. This pattern was reversed for I.B, Linear and Quadratic Functions and Data Representation and II.D, Systems of Linear Equations.

The committee judgments indicated that approximately 37% of the field test items were targeted at the two lowest levels of the hierarchy (Knowledge and Comprehension); 61% were at the middle two levels (Application and Analysis); and approximately 2% targeted the two highest cognitive levels (Synthesis and Evaluation).

The taxonomy of process levels employed by the item developers for the State Department of Education provides additional insight into the levels of cognitive demands made by the Algebra I/Math for the Technologies II field test items. This taxonomy is simplified to three levels, and is based on the National Assessment of Educational Progress framework and the National Council of Teachers of Mathematics publication, *Principles and Standards for School Mathematics*. The lowest level in this system is Conceptual and Procedural Understanding; the middle level is termed Application; and the highest level is Problem Solving and Reasoning. Based on data provided by the State Department of Education, approximately 51% of the field test items are at the Conceptual and Procedural Understanding level, 42% are at the Application level, and 7% are at the Problem Solving and Reasoning level. These proportions are similar to those based on committee members' judgments using Bloom's taxonomy, and reflect the difficulty encountered when writing multiple choice items to address the highest levels of cognitive processing.

Technical Study

The technical analysis of the Algebra I/Math for the Technologies II field test data conducted by Dr. Edward Wolfe and colleagues at Michigan State University (Appendix D) indicates that the field test is of good overall technical quality. The reliability of the items is adequate, the field test appears to be unidimensional in its assessment of algebra proficiency, and there is no evidence for differential item functioning among students having similar levels of overall algebra proficiency but who belong to different demographic groups. The authors of the study did identify 12 of the 285 field test items that should be reviewed for potential technical problems.

The technical review indicates that the field test is very difficult for the students in the sample who took it in spring 2002. In addition, students reporting they were enrolled in Math for the Technologies II performed at a much lower level on the test than students enrolled in Algebra I.

Discussion

Based on the findings of the alignment and technical studies, the Algebra I/Math for the Technologies II field test is aligned with the academic standards and is of good technical quality. The field test results indicate that the test is rigorous and sets a high standard for performance.

The overall difficulty of the field test, especially for students enrolled in Math for the Technologies II, is an area for possible future concern, especially for curriculum development and instructional practice. The report on setting standards for performance on the Algebra I/Math for the

Technologies II test presented to the State Board of Education on August 13, 2002 (Appendix F) suggests that the mathematics teachers on the standard-setting committee supported the high standards and level of rigor represented by the test. It is reported that, by and large, the teachers believed that students should be able to demonstrate their proficiency on a certain portion of the algebra standards to pass the test. Based on the field test data, approximately 51% of the students in the sample would have met the standard set for passing the Algebra I/Math for the Technologies II end of course test. Information was not reported on the relative predicted pass rates for students belonging to different demographic groups or enrolled in Algebra I or Math for the Technologies II.

It is reasonable to expect that student performance on the field test, which is not reported to the students and does not have an impact on student grades, is lower than performance on future administrations of the test that will affect student grades. It is also likely that teachers will become more aware of ways to teach the algebra standards more effectively and students will become more motivated to learn when the test begins to "count" in the 2003-2004 school year.

Students' responses to the questionnaire provide some insight into these issues, especially when responses from students enrolled in Algebra I are compared to those enrolled in Math for the Technologies II. When asked how important the field test was to them, 18% of the Algebra I students reported that it was not important, while 24.5% of the Math for the Technologies II students indicated it was not important. When asked the extent to which the content on the field test was covered in class, 63% of the Algebra I students reported that most or all of the test content had been covered in class, compared to only 45% of the Math for the Technologies II students. The students participating in the field test represent a sample of students enrolled in Algebra I or Math for the Technologies II statewide so these survey results are preliminary, but they do suggest some areas to address.

Finally, the use (or lack of use) of calculators by students taking the Algebra I/Math for the Technologies II end of course test needs to be examined. Students were allowed to use calculators for the field test (see the Calculator Use Policy from the State Department of Education, Appendix H), but not all students did. Approximately 30% of students enrolled in Algebra I reported they did not use a calculator on the test, and 40% of Math for the Technologies II students did not. Differences among classes in the use of calculators appear to reflect a lack of resources in some cases, and philosophical objections to the use of calculators in others. It is not clear what effect, if any, the use or non-use of calculators will have on student scores, but the policy regarding calculator use for the end of course test may need to be reviewed and clarified.

Recommendations

The Algebra I/Math for the Technologies II end of course field test appears to be well aligned with the academic standards, to be of high quality, and to reflect high standards for performance. Based on these findings, it is recommended that the test be approved with the following recommendations for future improvement:

1. Review the items identified as having potential technical problems and revise as needed.
2. As proposed in the report on standard-setting made to the State Board of Education (Appendix G), monitor the performance of students on the Algebra I/Math for the Technologies II tests administered in the 2002-2003 school year for possible review and revision of the standards set on the field test data.
3. Examine the extent and effects of the use of calculators on the end of course tests administered in 2002-2003 and make needed revisions to the policy on the use of calculators for the test.

Appendix A

South Carolina Mathematics Grades 9-12 Course Achievement Standards Algebra I

In Algebra 1, students build upon the mathematical understandings that are addressed in prekindergarten through the eighth grade. Students will

- use symbolic reasoning to represent mathematical situations, to express generalizations, and to study relationships among quantities;
- use functions to represent and model problem situations as well as to analyze and interpret relationships;
- set up equations in a wide range of situations and use a variety of methods to solve them; and
- use problem solving, representation, reasoning and proof, language and communication, and connections both within and outside mathematics.

In Algebra 1, hand-held graphing calculators are required as part of instruction and assessment. Students should use a variety of representations (concrete, numerical, algorithmic, graphical), tools (matrices, data), and technologies to model mathematical situations to solve meaningful problems. The technologies include, but are not limited to, powerful and accessible hand-held calculators as well as computers with graphing capabilities.

I. Understanding Functions

A. Relationships

1. Describe independent and dependent quantities in functional relationships.
2. Gather and record data or use data sets to determine functional (systematic) relationships between quantities.
3. Describe functional relationships for given problem situations and write equations, inequalities, and recursive relations to answer questions arising from the situations.
4. Represent relationships among quantities using concrete models, tables, graphs, diagrams, verbal descriptions, equations, and inequalities including representations involving computer algebra systems, spreadsheets, and graphing calculators.
5. Make judgments about units of measure and scales within a system and between systems.
6. Interpret and make inferences from explicit and recursive functional relationships.

B. Linear and Quadratic Functions and Data Representations

1. Identify and sketch the general forms of linear ($y = x$) and quadratic ($y = x^2$) parent functions.
2. For a variety of situations, identify and determine reasonable domain and range values for given situations.
3. Interpret situations in terms of given graphs or create situations that fit given graphs.
4. Represent, display, and interpret data using scatterplots, bar graphs, stem-and-leaf plots, and box-and-whiskers diagrams, including representations on graphing calculators and computers.
5. Write a linear equation that fits a data set, check the model for "goodness of fit," and make predictions using the model.

C. Generalizations, Algebraic Symbols, and Matrices

1. Read, write, and represent very large and very small numbers in a variety of forms including exponential.
2. Use unit analysis to check measurement computations.
3. Given situations, determine patterns and represent generalizations algebraically.
4. Use symbolic representation, reasoning, and proof to verify statements about numbers.

5. Recognize and justify the relationship between the magnitude of a number and the application of specific operations.
6. Identify and use properties related to operations with matrices (addition, subtraction, and scalar multiplication) to solve applied problems.

D. Algebraic Expressions in Problem Solving Situations

1. Find specific function values and evaluate expressions.
2. Simplify polynomial expressions and perform polynomial arithmetic.
3. Transform and solve equations and inequalities, factoring as necessary in problem situations.
4. Given a problem situation, determine whether to use a rough estimate, an approximation, or an exact answer. Select a suitable method of computing from techniques such as the use of mental mathematics, paper-and-pencil combinations, calculators, and computers.
5. Use supporting data to explain why a solution is mathematically reasonable.
6. Use the commutative, associative, and distributive properties to simplify algebraic expressions.

II. Linear Functions

A. Representations

1. Determine whether or not given situations can be represented by linear functions.
2. Based on the constraints of the problem, determine the domain and range values for linear functions.
3. Translate among and use algebraic, tabular, graphical, or verbal descriptions of linear functions using computer algebra systems, spreadsheets, and graphing calculators.

B. Interpretations

1. Develop the concept of slope as rate of change and determine slope from graphs, tables, and algebraic representations.
2. Interpret the meaning of slope and intercepts in situations using data, symbolic representations, or graphs.
3. With and without using a graphing calculator, investigate, describe, and predict the effects of changes in m and b on the graph of $y = mx + b$.
4. Graph and write equations of lines given characteristics such as two points, a point and a slope, or a slope and y -intercept.
5. Determine the intercepts of linear functions from graphs, tables, and algebraic representations.
6. With and without using a graphing calculator, interpret and predict the effects of changing slope and y -intercept in applied situations.
7. Relate direct variation to linear functions and solve problems involving proportional change.

C. Equations and Inequalities

1. Analyze situations involving linear functions and formulate linear equations or inequalities to solve problems.
2. Investigate methods for solving linear equations and inequalities using concrete models, graphs, and the properties of equality; select a method and solve the equations and inequalities.
3. Use the commutative, associative, distributive, equality, and identity properties to justify the steps in solving equations and inequalities.
4. Using concrete models for given contexts, interpret and determine the reasonableness of solutions to linear equations and inequalities.

D. Systems of Linear Equations

1. Analyze situations and formulate systems of linear equations to solve problems.

2. Solve systems of linear equations using concrete models, graphs, tables, and algebraic methods including computer algebra systems, spreadsheets, and graphing calculators.
3. For given contexts, interpret and determine the reasonableness of solutions to systems of linear equations.

III. Quadratic and Other Functions

A. Quadratic Functions

1. Given the constraints of the problem, determine the domain and range values for quadratic functions.
2. With and without using a graphing calculator, investigate, describe, and predict the effects of changes in the coefficient a on the graph of $y = ax^2$.
3. With and without using a graphing calculator, investigate, describe, and predict the effects of changes in the constant c on the graph of $y = x^2 + c$.
4. For problem situations, analyze graphs of quadratic functions and draw conclusions.
5. Solve quadratic equations using concrete models, tables, graphs, and algebraic methods that include factoring and using the quadratic formula as well as computer algebra systems, spreadsheets, and graphing calculators.
6. Relate the solutions of quadratic equations to the roots of their functions.

B. Other Functions

1. Use patterns to generate the laws of exponents and apply the laws of exponents in problem-solving situations.
2. Analyze data and represent situations involving inverse variation using concrete models, tables, graphs, or algebraic methods as well as computer algebra systems, spreadsheets, and graphing calculators.
3. Analyze data and represent situations involving exponential growth and decay using concrete models, tables, graphs, or algebraic methods as well as computer algebra systems, spreadsheets, and graphing calculators.

Appendix B

**Members of Algebra I Assessment Alignment
Review Committee
July 30, 2002**

Algebra I Alignment Committee Members

Mary Davis
Georgetown School District
Georgetown, SC

Angie McCune
Horry County Schools
Myrtle Beach, SC

Kathy Ewing
Richland District Two
Columbia, SC

Ann Shreve
Sumter School District 17
Sumter, SC

Donna Foster
Piedmont TEC
Greenwood, SC

Dianne Steelman
Lexington School District One
Lexington, SC

Bill Gillam
Richland District One
Columbia, SC

Ann Waddell
Spartanburg School District 5
Duncan, SC

Greg Hall
Beaufort Middle School
Beaufort, SC

Bobbie Kelley
Darlington County School District
Quimby, SC

Jan Massey
North Central Math/Science HUB
Rock Hill, SC

Terrie Mayfield
Greenville Math/Science HUB
Greenville, SC

Appendix C

**Materials for Assessment Alignment Review Meeting
July 30, 2002**



SOUTH CAROLINA ALGEBRA I/MATH FOR TECHNOLOGIES II ASSESSMENT REVIEW MEETING

July 30, 2002
9:00 a.m. - 4:30 p.m.
Rutledge Building
Columbia, SC

A G E N D A

- I. Introductions and Overview - David Potter
- II. Development of the Algebra I/Math for Technologies II Field Test - Dr. Lynne Mappus, INSITE
- III. Standards Alignment
 - A. Overview of task
 - B. Individual/small group work
- IV. Break for lunch
- V. Continue work on standards alignment
- VI. Closing discussion and adjournment

**Review of South Carolina Algebra I/Math for Technologies II Assessment Field Test
July 30, 2002**

Purpose: To fulfill the requirements of the Education Accountability Act of 1998.

Section 59-18-310 (A) Notwithstanding any other provision of law, the State Board of Education, through the Department of Education, is required to develop or adopt a statewide assessment program to measure student performance on state standards and:

- (1) identify areas in which students need additional support;
- (2) indicate the academic achievement for schools, districts, and the State; and
- (3) satisfy federal reporting requirements.

All assessments required to be developed or adopted under the provisions of this section or chapter must be objective and reliable.

Section 59-18-320. (A) After the first statewide field test of the assessment program in each of the four academic areas, and after the field tests of the end of course assessments of benchmark courses, the Education Oversight Committee, established in Section 59-6-10, will review the state assessment program and the course assessments for alignment with the state standards, level of difficulty and validity, and for the ability to differentiate levels of achievement, and will make recommendations for needed changes, if any. The review will be provided to the State Board of Education, the State Department of Education, the Governor, the Senate Education Committee, and the House Education and Public Works Committee as soon as feasible after the field tests. The Department of Education will then report to the Education Oversight Committee no later than one month after receiving the reports on the changes made to the assessments to comply with the recommendations.

Tasks to be accomplished:

1. To determine the degree of alignment between the Algebra I/Math for Technologies II field test items and the South Carolina High School Algebra I Standards;
2. To classify the cognitive demands of the Algebra I/Math for Technologies II field test items.

Organization of teams:

Form	Group Members
1	Kathy Ewing, Greg Hall
2	Mary Davis, Terri Mayfield
3	Donna Foster, Ann Waddell
4	Bobbie Kelley, Angie McCune
5	Ann Shreve, Diane Steelman
6	Bill Gillam, Jan Massey

DIRECTIONS

Review of South Carolina Algebra I/Math for Technologies II Assessment Field Test July 30, 2002

TASK 1: To determine the degree of alignment between the Algebra I/Math for Technologies II field test items and the South Carolina High School Algebra I Standards

The purpose of this activity is to determine the degree to which Algebra I assessment items match the Algebra I curriculum standards. Each analyst should work on this task individually. The task will be accomplished in several steps:

Step 1 - Read and answer the test question. As you are doing so, reflect on the kinds of algebra knowledge and skills needed to correctly answer the question and on the level of cognitive challenge the question presents to students.

Step 2 - Review the standards document to identify the standard(s) you believe the item best addresses. The standard(s) you identify may or may not match those previously identified.

Step 3 - Record the standard(s) you believe the item is addressing in the space provided. Use the numbering system in the standards document (e. g., IA6, etc.) to identify the standard(s). If you identify more than one standard, CIRCLE the standard you believe is the primary one addressed.

TASK 2: To identify the level of cognitive demands made by the item which must be met to correctly answer it.

The purpose of this task is to make a judgment regarding the cognitive difficulty of each test question. Refer to the document, "Major Categories in the Taxonomy of Educational Objectives (Bloom, 1956)."

- Level One (1) - Knowledge
- Level Two (2) - Comprehension
- Level Three (3) - Application
- Level Four (4) - Analysis
- Level Five (5) - Synthesis
- Level Six (6) - Evaluation.

Based on your reading of the question and on the cognitive challenges you believe students will face to correctly answer it, identify the level of cognitive difficulty you believe best describes the item and record it in the space provided.

Appendix D

Technical Review of Algebra I/Math for the Technologies II Field Test

A Review of the South Carolina End of Course Examination in Algebra I/Math for the
Technologies II:

A Report to the Education Oversight Committee

Edward W. Wolfe

Benita Barnes

Linda Chard

Michigan State University

Table of Contents

Procedures.....	4
Results.....	6
Dimensionality.....	7
Reliability & Precision.....	7
Fit & Biserial Correlations.....	8
Logit Interpretations.....	9
Differential Item Functioning.....	11
Output.....	11

List of Tables and Figures

Figure 1: Scree Plot for First Six Dimensions	7
Table 1: Precision of Person Measures	8
Table 2: Summary of Item Fit and Biserial Statistics	8
Table 3: Summary of Person Fit Statistics.....	8
Figure 2: Examinee Ability & Item Difficulty Map	10
Table 4: Item-Level Statistics	11
Table 5: Summary of Item DIF Statistics	18

A Review of the South Carolina End of Course Examination in Algebra I/Math for the Technologies II

A Report to the Education Oversight Committee

We have completed our review of the six forms of the Algebra I/Math for the Technologies II tests that were field tested in the spring of 2002. Our report contains three sections: (a) a description of the procedures and indices we used when analyzing the data, (b) a summary of the results of our review, and (c) output that we produced for the review.

Procedures

We received a data set and record layout from David Potter of the Education Oversight Committee. There were six forms, each containing 45 unique dichotomous items and 15 linking items that were common to all forms—a total of 285 items. We scaled the data using a dichotomous Rasch model using *Winsteps* (Linacre & Wright, 2001). This model depicts the log odds (*logit*) of an examinee answering a particular item correctly (versus incorrectly) as a function of two parameters: θ , the examinees ability, and δ , the item's difficulty,

$$\ln\left(\frac{\pi_{x=1}}{\pi_{x=0}}\right) = \theta_n - \delta_i.$$

Each form was scaled several times—once as a pool with all forms and all linking items simultaneously (*joint analyses*), once as a separate form deleting responses to all other forms (*separate analyses*), and twice for each of a series of differential item functioning (DIF) studies so that a separate scaling was produced for the reference group and the focus group for each DIF study (*DIF analyses*).

We focused on the following indices in the output of the joint and separate analyses.

First, we performed a principal component analysis of the standardized residuals of the observed scores from the expectations of the dichotomous Rasch model. We then considered the magnitude of the *eigenvalue* of each extracted factor and the magnitude of the *factor loading* for each item on the instrument. If the eigenvalue for a particular factor is very small, relative to the total variance in the data (i.e., the proportion of variance accounted for by a particular factor), then little evidence exists for multidimensionality in the data. Similarly, if factor loadings for each item are small (a value of .40), then individual items exhibit little evidence of multidimensionality (Stevens, 1996).

Second, we examined the *reliability of separation*, which is an index analogous to coefficient α and indicates the degree to which the instrument is internally consistent.

Third, we examined the values of the *standard errors* of the item calibrations and person measures to determine the level of precision with which item calibration and person measures have been estimated. Because the standard errors on the logit metric may be difficult to understand, we transformed the standard errors to the probability metric by first creating a 95% confidence band around logits at various locations on the person and item distribution and then transforming the upper and lower limits of those confidence bands to the probability scale via the following equation,

$$p = \frac{\exp(\gamma_n - \bar{\lambda})}{1 + \exp(\gamma_n - \bar{\lambda})},$$

where γ_n is the parameter estimate for person or item n and $\bar{\lambda}$ is the average parameter estimate for the other measurement facet (i.e., $\bar{\lambda}$ is the average item difficulty when γ_n is a person measure and $\bar{\lambda}$ is the average person ability when γ_n is an item calibration).

Fourth, we examined the *standardized mean-square outfit statistics*,

$$z_{outfit} = \frac{\sum_{i=1}^I z_{ni}^2}{I},$$

where z_{ni}^2 is the square of the standardized residual for the response of person n to item i . This index can be generated for each person and each item in the data, and it indicates the degree to which the observed scores are consistent with the expectations generated by the Rasch model. If there is consistency, then it is assumed that the observations behave in the manner that the Rasch model depicts. If consistency is lacking, then the individual item or person is behaving in a way that is inconsistent with the Rasch model. These inconsistencies may arise due to a variety of reasons (e.g., multidimensionality in the instrument, aberrant behavior on the part of the examinee, incorrectly specified latent trait model). In this study, items were flagged as being potentially problematic if the standardized mean-square outfit index exceeded 4.00, and persons were flagged as being potentially problematic if the standardized mean-square outfit index exceeded 2.00 (Richard Smith, personal communication).

Fifth, we examined the *raw score biserial correlation* for each item between the item scores and total scores across examinees. This index depicts the degree to which individual items rank order examinees in a manner that is consistent with all other items on the instrument. The index assumes that the underlying distribution from which the dichotomous correct/incorrect distinction is made is normal in shape, and in doing so, reduces the influence of item difficulty on the magnitude of the correlation. We considered items to be potentially problematic if the biserial correlation was less than 0.15.

Sixth, we examined the values and distributions of the *logits* for each item and each person to determine whether there are items that are too easy or too difficult to contribute to information about examinees.

Seventh, we examined Raju's *Signed Area Index (SAI)*, which was generated for each item for the purpose of evaluating threats due to differential item functioning (DIF). We evaluated items for DIF based on several groupings. Specifically, we considered gender (male versus female), race (white versus all others), parent education level (college versus no college), and mathematics course (Algebra I versus Math for the Technologies II). For each of these pairings, our DIF assessment procedure was conducted by first scaling the data to the dichotomous Rasch model for each of the groupings separately. The estimated item difficulties were then scaled to have the same mean and standard deviation using the mean and sigma method (Hambleton & Swaminathan, 1985). Next, Raju's Signed Area Index (Raju, 1988, 1990) was computed for the two calibrations for each item,

$$SAI_i = \delta_{reference} - \delta_{focus},$$

where δ_g is the item calibration obtained from the data for group g . The SAI can be tested for statistical significance via the following formula,

$$z_{SAI} = \frac{SAI}{\sqrt{V_{\delta_{reference}} + V_{\delta_{focus}}}},$$

where V_{δ_g} is the variance of the δ_g parameter estimate.

With the large sample sizes we encountered, it is likely that some items would be flagged for DIF using this standardized difference simply because of the sensitivity of the z_{SAI} to large sample sizes. Therefore, we chose to interpret the SAI directly as an effect size index using the 0.50 criterion value suggested by Draba (Draba, 1977).

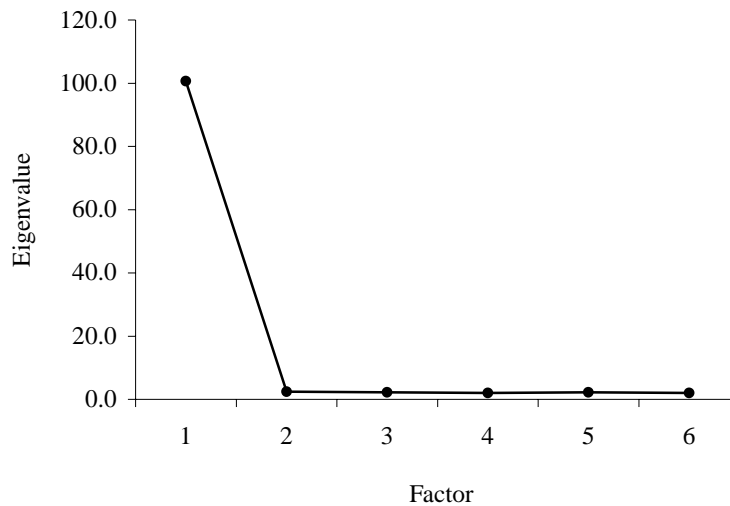
Results

The results of our analyses, described in the following sections, indicate that the Algebra I/Math for the Technologies II instrument produces unidimensional measures. Specifically, the factor accounted for by the Rasch model accounted for 35% of the total variance, with the remaining factors accounting for less than 1% of the variance each. The reliability of separation was reasonably high, about .81. The precision of the 60-item forms was also reasonable, with a 95% confidence band around the probabilities that average-ability examinees would answer an average difficulty item ranging from 0.37 to 0.63. Item fit and biserial correlations resulted in flags for 14 (5%) and 45 (16%) of the items, respectively. Of these, we recommend careful review of the content of 12 items that were flagged for both high standardized mean-square outfit values and low biserial correlations. In addition, only 3% of the examinees were flagged for exhibiting misfit. The match between person abilities and item difficulties is fairly good. However, the item difficulties tend to be somewhat higher than necessary for this population, and the item pool is somewhat thin at the highest levels of difficulty. In addition, the test developers should demonstrate that the empirical item difficulties are consistent with the difficulty or complexity of the underlying categories for which the items were written. Finally, there is minimal evidence of DIF relating to ethnicity, gender, or parent education.

Dimensionality

Figure 1 depicts the scree plot of the first six factors generated by the joint analysis of the Algebra I/Math for the Technologies II data. Specifically, the dichotomous Rasch model was fit to the data, residuals were generated for the observed data from the Rasch expectations, and a principal component analysis was performed on the residuals. The eigenvalue of the Rasch and each subsequent dimension was then computed, and Figure 1 plots the factor number on the x-axis and the value of the eigenvalue on the y-axis. From this figure, it is clear that the Rasch factor (the first dimension of the Algebra algebra test) accounts for a large proportion of the total variance (35% to be precise). The remaining dimensions account for less than 1% of the total variance. Hence, the scree plot provides very strong evidence that the Algebra I/Math for the Technologies II instrument is unidimensional. In addition, examination of the factor loadings for the second dimension indicates that only one item from the pool of 285 items has a loading on that factor greater than 0.40. Hence, we conclude that the Algebra I/Math for the Technologies II instrument produces unidimensional measures.

Figure 1: Scree Plot for First Six Dimensions



Reliability & Precision

The reliability of separation for the joint analysis equals 0.81 when the index is based upon parametric standard errors. Because the data frequently exhibit some misfit to the Rasch model, the asymptotic standard errors tend to overestimate the precision of the measures, an alternative standard error, called the *real* standard error in *Winsteps*, can be computed by multiplying the standard errors by a proportionality constant, the model's deviation statistic divided by its degrees of freedom (Linacre & Wright, 2001). The reliability of the joint analysis based on the real standard errors equals 0.80.

Table 1 depicts the precision of person measures at several points on the underlying measurement scale for the joint analyses. That table contains the logit and associated probability values (i.e., the probability of the person answering an item of average difficulty correctly) and the upper and lower bounds of the 95% confidence interval on both the logit and probability scales. These figures reveal that the confidence bands on the probability scale range from about .16 probability points to about .26 probability points between the extreme and the middle ranges of the logit scale. These are not particularly large confidence bands, but if more precision is required, a larger number of items should be included on the operational test forms.

Table 1: Precision of Person Measures

Logit	SE	Logit_{LL}	Logit_{UL}	P_{LL}	P	P_{UL}
-2.01	0.39	-2.77	-1.25	0.06	0.12	0.22
-1.00	0.30	-1.59	-0.41	0.17	0.27	0.40
0.00	0.27	-0.53	0.53	0.37	0.50	0.63
1.00	0.31	0.39	1.61	0.60	0.73	0.83
2.00	0.38	1.26	2.74	0.78	0.88	0.94

Fit & Biserial Correlations

Table 2 displays the summary statistics for the standardized mean-square outfit statistics and biserial correlations for the 285 items in the joint analyses. From these figures, it is clear that only a small proportion of the items were flagged for exhibiting misfit according to the standardized mean-square outfit statistic. A considerably greater number of items were flagged for having a biserial correlation less than 0.15. It should also be noted that nearly every item that was flagged for having a large fit statistic was also flagged for having a small biserial correlation. The following items—items flagged for both high standardized mean-square outfit statistics and low biserial correlations—should be examined carefully to rule out problems caused by miskeying, misleading options, or unclear wording: Linking—#4, #13; Form 1—#42; Form 2—#1; Form 3—#1; Form 4—#2, #25, #42; Form 5—#12; Form 6—#18, #19, #41, #45.

Table 2: Summary of Item Fit and Biserial Statistics

Statistic	Mean	SD	Minimum	Maximum	P_{flagged}	N_{flagged}
Z_{outfit}	-0.09	2.85	-9.90	9.68	0.05	14
r_{biserial}	0.31	0.15	-0.13	0.63	0.16	45

Table 3 displays the summary statistics for the standardized mean-square outfit statistics and biserial correlations for the 4278 persons in the joint analysis. From these figures, it is clear that only a very small proportion of the persons were flagged for exhibiting misfit according to the standardized mean-square outfit statistic. We believe that this proportion is small enough to dismiss the notion that there are substantial differences between students with respect to the validity of the items as measures of their algebra knowledge.

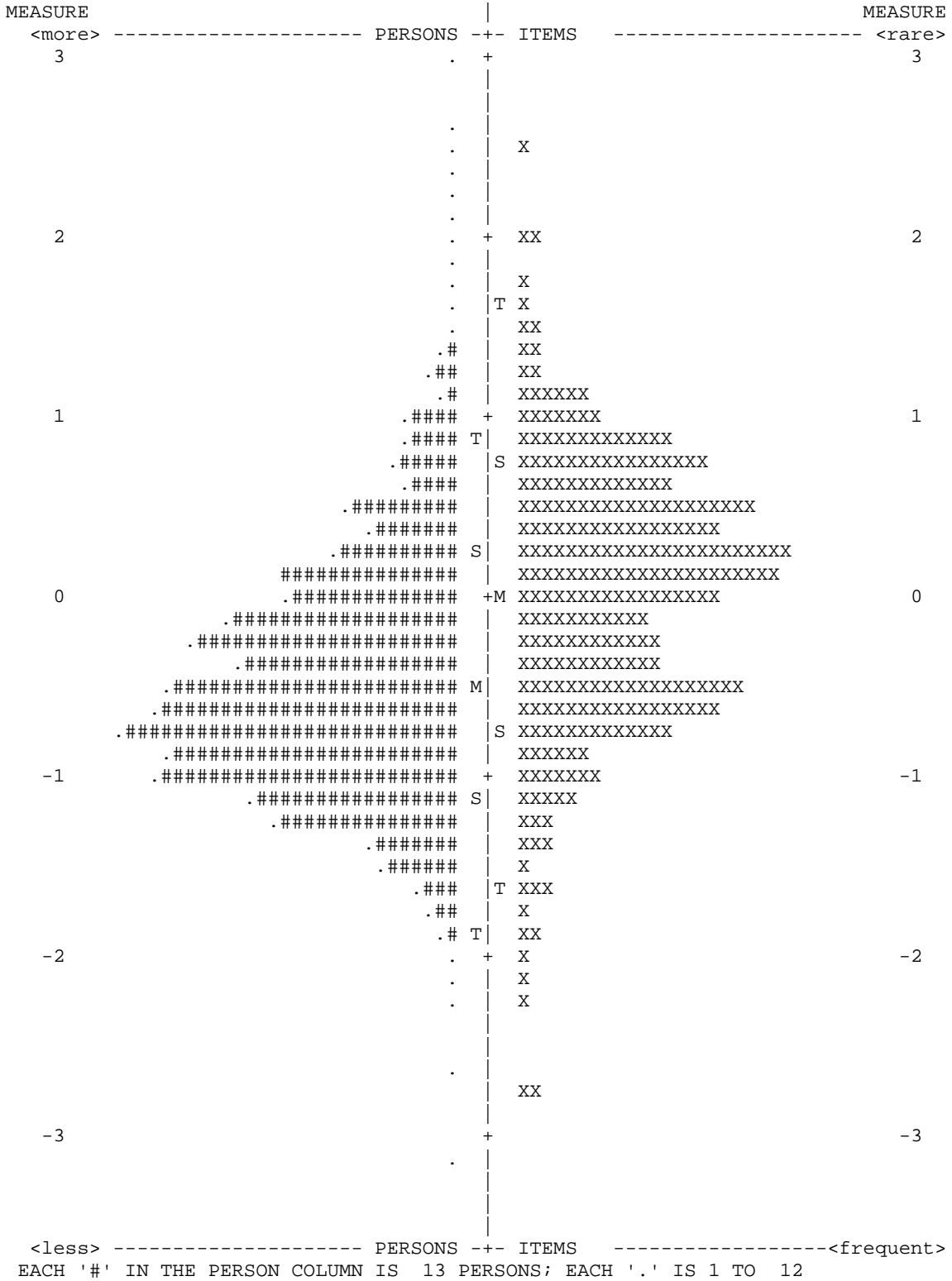
Table 3: Summary of Person Fit Statistics

Statistic	Mean	SD	Minimum	Maximum	P_{flagged}	N_{flagged}
Z_{outfit}	-0.09	1.01	-3.20	4.78	0.03	111

Logit Interpretations

Figure 1 displays the joint examinee measure and item calibration map. This figure shows the correspondence between the algebra abilities of the examinees and the difficulties of the 285 items. Persons are shown on the left side of the figure, and items are shown on the right side of the figure. In general, the distribution of items and examinees should be well-matched (i.e., there should be item coverage at each level for which there is an examinee). From this display, we see that the distribution of examinee abilities looks normal in shape. The distribution of item difficulties also seems fairly normal in shape, but it also seems to be slightly bimodal. This phenomenon could be caused by content specifications or because of domain characteristics that cause items to cluster around two points on the difficulty continuum with respect to difficulty. It should be noted that there is good coverage of the examinee ability range with respect to the range of the item difficulties. It should also be noted, however, that the item pool tends to be somewhat difficult, on average, relative to the examinee abilities. This is evidenced by the fact that the larger of the two item difficulty modes lies approximately one logit above the mode of the examinee ability distribution. Hence, it may be advisable to choose a larger proportion of items that lie below the mean difficulty of the pool of items for the operational form(s), assuming that the field-test sample is representative of the target population and that the operational form(s) will be interpreted in a norm-referenced manner. In addition, the difficult tail of the item distribution is somewhat sparse relative to the examinee ability distribution. Assuming that the field-test sample is a random sample from the population and assuming that differentiating among high-ability examinees is desirable in operational testing, it may be advisable to develop a few additional items that are highly difficult. Finally, additional analyses should be performed, comparing substantive classifications of items with the empirical difficulties. The test developers should, at the very least, provide evidence that the items rank order themselves in a manner that is consistent with the substantive theory upon which the items were developed. That is, items that were written to be difficult should be shown to exhibit higher levels of difficulty and items that were written to be easy should be shown to exhibit lower levels of difficulty.

Figure 2: Examinee Ability & Item Difficulty Map



Differential Item Functioning

Overall, it seems that the threat of differential item functioning is minimal. Race DIF analyses resulted in 7 of the 285 items (2%) being flagged. Gender DIF analyses resulted in 8 of the items (3%) being flagged. Parent education DIF analyses resulted in 9 of the items (3%) being flagged. Flagged items can be identified by finding the item numbers in Table X of this document. If the test developer has not already performed sensitivity reviews of the content of any flagged items following the field test, such reviews should be performed. It should also be noted that we performed a t-test to compare students who were enrolled in Algebra I/Math for the Technologies II to students who were enrolled in a general mathematics course. The difference in the ability estimates of these groups was statistically significant with a very large effect size, $t(3628) = 32.88$, $p > .0001$, $d = 1.00$.

Output

Table 4 displays the item statistics for the 285 items in the joint analysis.

Table 4: Item-Level Statistics

Form	Item	Logit	Zoutfit	Biserial	p-value	p high	p low	z high	z low	biserial low	z & bis
Link	1	-1.25	-0.59	0.30	0.67	0	0	0	0	0	0
Link	2	0.48	-4.68	0.48	0.30	0	0	0	1	0	0
Link	3	0.25	0.63	0.33	0.34	0	0	0	0	0	0
Link	4	0.75	6.80	0.12	0.25	0	0	1	0	1	1
Link	5	0.65	1.81	0.30	0.27	0	0	0	0	0	0
Link	6	-0.36	-9.42	0.48	0.48	0	0	0	1	0	0
Link	7	-0.55	-9.29	0.47	0.52	0	0	0	1	0	0
Link	8	0.69	1.05	0.31	0.26	0	0	0	0	0	0
Link	9	1.38	1.23	0.34	0.16	0	1	0	0	0	0
Link	10	0.78	5.05	0.18	0.25	0	0	1	0	0	0
Link	11	-1.13	-9.90	0.55	0.65	0	0	0	1	0	0
Link	12	-0.01	0.05	0.32	0.40	0	0	0	0	0	0
Link	13	0.41	9.68	0.05	0.31	0	0	1	0	1	1
Link	14	0.70	-2.29	0.41	0.26	0	0	0	0	0	0
Link	15	0.22	-3.35	0.41	0.35	0	0	0	0	0	0
1	1	-0.61	-2.87	0.42	0.52	0	0	0	0	0	0
1	2	0.97	2.32	0.12	0.20	0	0	0	0	1	0
1	3	-1.93	-0.71	0.27	0.79	0	0	0	0	0	0
1	4	1.04	0.47	0.30	0.19	0	1	0	0	0	0
1	5	1.79	2.72	-0.03	0.11	0	1	0	0	1	0
1	6	1.37	2.14	0.08	0.15	0	1	0	0	1	0
1	7	0.18	2.47	0.13	0.34	0	0	0	0	1	0
1	8	-0.52	-1.60	0.35	0.50	0	0	0	0	0	0
1	9	-0.52	-1.11	0.32	0.50	0	0	0	0	0	0
1	10	0.11	3.90	0.05	0.36	0	0	0	0	1	0
1	11	0.40	0.46	0.26	0.30	0	0	0	0	0	0
1	12	0.31	0.84	0.25	0.32	0	0	0	0	0	0
1	13	-0.44	-5.00	0.51	0.48	0	0	0	1	0	0
1	14	0.14	3.73	0.09	0.35	0	0	0	0	1	0
1	15	1.62	-1.93	0.55	0.12	0	1	0	0	0	0
1	16	0.27	-2.44	0.48	0.33	0	0	0	0	0	0
1	17	0.88	-0.23	0.39	0.22	0	0	0	0	0	0

1	18	-0.23	-2.51	0.42	0.43	0	0	0	0	0	0
1	19	0.47	1.41	0.20	0.29	0	0	0	0	0	0
1	20	1.25	1.77	0.11	0.16	0	1	0	0	1	0
1	21	0.38	3.01	0.09	0.30	0	0	0	0	1	0
1	22	-0.59	-0.61	0.30	0.51	0	0	0	0	0	0
1	23	-0.46	-1.10	0.33	0.48	0	0	0	0	0	0
1	24	0.13	1.12	0.22	0.35	0	0	0	0	0	0
1	25	-1.13	-2.02	0.36	0.63	0	0	0	0	0	0
1	26	0.47	-1.01	0.41	0.29	0	0	0	0	0	0
1	27	0.54	-1.30	0.46	0.27	0	0	0	0	0	0
1	28	0.07	0.88	0.28	0.37	0	0	0	0	0	0
1	29	-0.48	-3.28	0.43	0.49	0	0	0	0	0	0
1	30	0.02	2.56	0.15	0.38	0	0	0	0	0	0
1	31	-0.06	0.02	0.29	0.39	0	0	0	0	0	0
1	32	-0.65	-6.34	0.56	0.53	0	0	0	1	0	0
1	33	-0.73	-4.13	0.49	0.55	0	0	0	1	0	0
1	34	-0.25	1.19	0.24	0.44	0	0	0	0	0	0
1	35	0.13	2.01	0.19	0.35	0	0	0	0	0	0
1	36	0.18	-2.73	0.49	0.34	0	0	0	0	0	0
1	37	-1.18	-3.00	0.42	0.65	0	0	0	0	0	0
1	38	0.01	-0.59	0.32	0.38	0	0	0	0	0	0
1	39	0.65	2.70	0.08	0.25	0	0	0	0	1	0
1	40	-0.45	-6.52	0.58	0.48	0	0	0	1	0	0
1	41	0.48	0.70	0.27	0.28	0	0	0	0	0	0
1	42	0.18	4.89	-0.02	0.34	0	0	1	0	1	1
1	43	-0.57	2.16	0.19	0.51	0	0	0	0	0	0
1	44	-0.95	-4.30	0.48	0.59	0	0	0	1	0	0
1	45	0.28	3.33	0.07	0.32	0	0	0	0	1	0
2	1	0.04	4.01	0.09	0.38	0	0	1	0	1	1
2	2	0.62	0.85	0.28	0.26	0	0	0	0	0	0
2	3	0.49	3.39	0.08	0.29	0	0	0	0	1	0
2	4	-0.73	-1.83	0.39	0.55	0	0	0	0	0	0
2	5	-0.97	-3.89	0.49	0.60	0	0	0	0	0	0
2	6	1.98	1.07	0.19	0.09	0	1	0	0	0	0
2	7	-0.39	-3.15	0.45	0.47	0	0	0	0	0	0
2	8	-2.13	-1.15	0.33	0.82	0	0	0	0	0	0
2	9	0.31	1.96	0.21	0.32	0	0	0	0	0	0
2	10	0.20	-0.98	0.40	0.34	0	0	0	0	0	0
2	11	-1.10	0.49	0.25	0.63	0	0	0	0	0	0
2	12	-0.49	-2.69	0.43	0.49	0	0	0	0	0	0
2	13	-0.03	-1.27	0.40	0.39	0	0	0	0	0	0
2	14	0.34	-2.82	0.52	0.32	0	0	0	0	0	0
2	15	-0.30	-3.12	0.46	0.45	0	0	0	0	0	0
2	16	-0.02	1.85	0.24	0.39	0	0	0	0	0	0
2	17	-0.37	1.49	0.23	0.47	0	0	0	0	0	0
2	18	-0.11	-0.04	0.33	0.41	0	0	0	0	0	0
2	19	1.29	3.27	-0.07	0.16	0	1	0	0	1	0
2	20	0.44	2.67	0.16	0.30	0	0	0	0	0	0
2	21	-0.60	-0.68	0.40	0.52	0	0	0	0	0	0
2	22	0.25	-2.85	0.55	0.33	0	0	0	0	0	0

2	23	0.10	-1.34	0.41	0.36	0	0	0	0	0	0
2	24	-0.48	-2.13	0.40	0.49	0	0	0	0	0	0
2	25	-0.65	-1.26	0.38	0.53	0	0	0	0	0	0
2	26	0.54	1.86	0.24	0.28	0	0	0	0	0	0
2	27	0.78	-0.45	0.41	0.24	0	0	0	0	0	0
2	28	0.61	-0.40	0.42	0.27	0	0	0	0	0	0
2	29	0.11	-1.28	0.40	0.36	0	0	0	0	0	0
2	30	0.51	0.96	0.30	0.28	0	0	0	0	0	0
2	31	0.93	1.67	0.22	0.21	0	0	0	0	0	0
2	32	-0.76	-3.81	0.49	0.55	0	0	0	0	0	0
2	33	0.95	2.37	0.14	0.21	0	0	0	0	1	0
2	34	0.22	1.67	0.23	0.34	0	0	0	0	0	0
2	35	-1.66	-2.79	0.44	0.74	0	0	0	0	0	0
2	36	0.49	0.90	0.28	0.29	0	0	0	0	0	0
2	37	-0.23	1.25	0.25	0.43	0	0	0	0	0	0
2	38	1.94	1.48	0.12	0.10	0	1	0	0	1	0
2	39	-0.01	1.69	0.25	0.39	0	0	0	0	0	0
2	40	0.10	-1.28	0.42	0.36	0	0	0	0	0	0
2	41	0.37	3.08	0.13	0.31	0	0	0	0	1	0
2	42	-0.85	3.09	0.16	0.58	0	0	0	0	0	0
2	43	-0.68	-3.14	0.44	0.54	0	0	0	0	0	0
2	44	-0.40	-1.20	0.38	0.48	0	0	0	0	0	0
2	45	0.36	2.79	0.16	0.31	0	0	0	0	0	0
3	1	-0.79	8.94	-0.10	0.57	0	0	1	0	1	1
3	2	0.71	1.50	0.16	0.25	0	0	0	0	0	0
3	3	-0.76	-2.43	0.39	0.56	0	0	0	0	0	0
3	4	0.16	1.93	0.17	0.35	0	0	0	0	0	0
3	5	0.82	-0.65	0.40	0.23	0	0	0	0	0	0
3	6	1.49	1.93	0.21	0.14	0	1	0	0	0	0
3	7	-1.54	-1.42	0.38	0.72	0	0	0	0	0	0
3	8	-1.62	-1.24	0.34	0.74	0	0	0	0	0	0
3	9	0.88	1.68	0.15	0.22	0	0	0	0	1	0
3	10	-2.80	-1.39	0.34	0.90	0	0	0	0	0	0
3	11	0.54	1.57	0.17	0.28	0	0	0	0	0	0
3	12	0.84	-0.35	0.37	0.23	0	0	0	0	0	0
3	13	-0.95	-1.95	0.37	0.60	0	0	0	0	0	0
3	14	0.22	1.09	0.27	0.34	0	0	0	0	0	0
3	15	-1.19	-3.81	0.49	0.65	0	0	0	0	0	0
3	16	-0.59	-3.67	0.46	0.52	0	0	0	0	0	0
3	17	-0.97	-4.30	0.50	0.60	0	0	0	1	0	0
3	18	-0.20	-1.20	0.38	0.43	0	0	0	0	0	0
3	19	0.70	2.81	0.05	0.25	0	0	0	0	1	0
3	20	0.67	1.25	0.20	0.26	0	0	0	0	0	0
3	21	-0.59	-3.56	0.45	0.52	0	0	0	0	0	0
3	22	0.59	2.45	0.13	0.27	0	0	0	0	1	0
3	23	0.32	-2.87	0.52	0.32	0	0	0	0	0	0
3	24	0.40	1.22	0.22	0.31	0	0	0	0	0	0
3	25	1.06	3.47	0.05	0.19	0	1	0	0	1	0
3	26	0.09	2.12	0.19	0.37	0	0	0	0	0	0
3	27	0.70	0.91	0.20	0.25	0	0	0	0	0	0

3	28	-0.44	-3.63	0.47	0.49	0	0	0	0	0	0
3	29	-1.00	0.15	0.30	0.61	0	0	0	0	0	0
3	30	-0.18	-3.96	0.50	0.43	0	0	0	0	0	0
3	31	0.73	0.34	0.30	0.25	0	0	0	0	0	0
3	32	-0.15	-4.03	0.50	0.42	0	0	0	1	0	0
3	33	0.06	1.80	0.19	0.38	0	0	0	0	0	0
3	34	0.99	3.06	0.07	0.20	0	0	0	0	1	0
3	35	-0.45	-4.30	0.51	0.49	0	0	0	1	0	0
3	36	0.70	1.63	0.25	0.25	0	0	0	0	0	0
3	37	-0.62	-1.78	0.36	0.53	0	0	0	0	0	0
3	38	0.69	3.38	0.05	0.25	0	0	0	0	1	0
3	39	1.15	0.99	0.20	0.18	0	1	0	0	0	0
3	40	0.42	1.19	0.22	0.30	0	0	0	0	0	0
3	41	0.02	-2.25	0.42	0.38	0	0	0	0	0	0
3	42	-0.14	-0.55	0.32	0.42	0	0	0	0	0	0
3	43	-0.30	-0.86	0.33	0.45	0	0	0	0	0	0
3	44	0.12	0.73	0.28	0.36	0	0	0	0	0	0
3	45	0.05	-3.88	0.54	0.38	0	0	0	0	0	0
4	1	0.82	1.70	0.18	0.25	0	0	0	0	0	0
4	2	-0.61	6.42	0.04	0.55	0	0	1	0	1	1
4	3	-1.40	-1.76	0.37	0.71	0	0	0	0	0	0
4	4	-0.74	-0.87	0.33	0.58	0	0	0	0	0	0
4	5	1.16	-0.79	0.46	0.19	0	1	0	0	0	0
4	6	0.12	0.63	0.26	0.38	0	0	0	0	0	0
4	7	-0.36	0.00	0.32	0.49	0	0	0	0	0	0
4	8	-0.14	0.59	0.29	0.44	0	0	0	0	0	0
4	9	0.10	-0.10	0.32	0.39	0	0	0	0	0	0
4	10	0.93	2.24	0.18	0.23	0	0	0	0	0	0
4	11	0.24	-2.65	0.48	0.36	0	0	0	0	0	0
4	12	-0.83	-2.57	0.41	0.60	0	0	0	0	0	0
4	13	-0.36	-0.63	0.33	0.49	0	0	0	0	0	0
4	14	0.29	-2.13	0.46	0.35	0	0	0	0	0	0
4	15	0.34	1.43	0.23	0.34	0	0	0	0	0	0
4	16	-0.24	-0.60	0.33	0.46	0	0	0	0	0	0
4	17	-0.40	-0.94	0.33	0.50	0	0	0	0	0	0
4	18	-0.55	0.01	0.29	0.53	0	0	0	0	0	0
4	19	0.96	3.21	0.02	0.22	0	0	0	0	1	0
4	20	0.71	2.48	0.12	0.27	0	0	0	0	1	0
4	21	0.82	1.90	0.19	0.25	0	0	0	0	0	0
4	22	0.74	-0.74	0.44	0.26	0	0	0	0	0	0
4	23	0.20	2.58	0.17	0.37	0	0	0	0	0	0
4	24	0.39	3.08	0.11	0.33	0	0	0	0	1	0
4	25	-0.36	4.35	0.08	0.49	0	0	1	0	1	1
4	26	-1.12	-2.20	0.41	0.66	0	0	0	0	0	0
4	27	0.06	-1.48	0.41	0.40	0	0	0	0	0	0
4	28	0.25	2.61	0.15	0.36	0	0	0	0	0	0
4	29	-0.87	-3.14	0.45	0.60	0	0	0	0	0	0
4	30	-0.83	-0.64	0.33	0.60	0	0	0	0	0	0
4	31	-1.39	-3.91	0.53	0.71	0	0	0	0	0	0
4	32	-1.56	-3.69	0.53	0.74	0	0	0	0	0	0

4	33	1.04	2.69	0.07	0.21	0	0	0	0	1	0
4	34	-0.27	-2.58	0.41	0.47	0	0	0	0	0	0
4	35	-0.80	-4.90	0.53	0.59	0	0	0	1	0	0
4	36	0.26	2.05	0.19	0.35	0	0	0	0	0	0
4	37	-0.47	-3.03	0.43	0.52	0	0	0	0	0	0
4	38	-0.04	1.47	0.23	0.42	0	0	0	0	0	0
4	39	0.71	-0.30	0.30	0.27	0	0	0	0	0	0
4	40	0.21	0.01	0.33	0.37	0	0	0	0	0	0
4	41	0.89	1.22	0.22	0.24	0	0	0	0	0	0
4	42	-0.16	4.18	0.13	0.45	0	0	1	0	1	1
4	43	-0.76	-1.68	0.37	0.58	0	0	0	0	0	0
4	44	-0.10	-0.38	0.32	0.43	0	0	0	0	0	0
4	45	0.59	3.67	0.03	0.29	0	0	0	0	1	0
5	1	0.39	1.77	0.23	0.33	0	0	0	0	0	0
5	2	0.67	1.99	0.21	0.28	0	0	0	0	0	0
5	3	-0.54	-1.85	0.40	0.53	0	0	0	0	0	0
5	4	-0.58	-2.40	0.42	0.54	0	0	0	0	0	0
5	5	-0.62	-0.37	0.33	0.55	0	0	0	0	0	0
5	6	1.07	2.10	0.19	0.21	0	0	0	0	0	0
5	7	-2.28	-0.79	0.41	0.85	0	0	0	0	0	0
5	8	-2.79	-2.36	0.49	0.90	1	0	0	0	0	0
5	9	-0.35	0.63	0.32	0.49	0	0	0	0	0	0
5	10	-1.77	-2.12	0.43	0.78	0	0	0	0	0	0
5	11	0.81	-0.99	0.45	0.26	0	0	0	0	0	0
5	12	0.60	4.36	0.06	0.29	0	0	1	0	1	1
5	13	-0.14	-4.78	0.57	0.45	0	0	0	1	0	0
5	14	0.40	-1.18	0.45	0.33	0	0	0	0	0	0
5	15	-0.81	-2.96	0.45	0.60	0	0	0	0	0	0
5	16	0.20	-0.42	0.38	0.37	0	0	0	0	0	0
5	17	0.48	2.62	0.15	0.32	0	0	0	0	0	0
5	18	0.01	-1.87	0.43	0.41	0	0	0	0	0	0
5	19	0.47	2.47	0.17	0.32	0	0	0	0	0	0
5	20	-0.47	1.04	0.29	0.52	0	0	0	0	0	0
5	21	-0.44	-6.64	0.61	0.51	0	0	0	1	0	0
5	22	0.26	1.75	0.22	0.36	0	0	0	0	0	0
5	23	0.31	-1.53	0.44	0.35	0	0	0	0	0	0
5	24	-0.26	-2.59	0.44	0.47	0	0	0	0	0	0
5	25	-0.46	2.67	0.21	0.52	0	0	0	0	0	0
5	26	-0.07	0.31	0.32	0.43	0	0	0	0	0	0
5	27	0.09	-0.57	0.39	0.40	0	0	0	0	0	0
5	28	-0.21	-1.21	0.38	0.46	0	0	0	0	0	0
5	29	-0.19	-2.04	0.43	0.46	0	0	0	0	0	0
5	30	0.03	-2.05	0.45	0.41	0	0	0	0	0	0
5	31	0.51	1.59	0.26	0.31	0	0	0	0	0	0
5	32	-1.38	-2.06	0.44	0.71	0	0	0	0	0	0
5	33	0.51	-0.23	0.41	0.31	0	0	0	0	0	0
5	34	-0.93	-5.36	0.59	0.62	0	0	0	1	0	0
5	35	-0.02	-0.55	0.39	0.42	0	0	0	0	0	0
5	36	1.10	2.27	0.15	0.21	0	0	0	0	0	0
5	37	0.08	1.29	0.29	0.40	0	0	0	0	0	0

5	38	-0.52	2.51	0.23	0.53	0	0	0	0	0	0
5	39	0.31	1.92	0.24	0.35	0	0	0	0	0	0
5	40	0.51	2.56	0.18	0.31	0	0	0	0	0	0
5	41	0.77	2.29	0.18	0.26	0	0	0	0	0	0
5	42	0.89	0.40	0.28	0.24	0	0	0	0	0	0
5	43	-0.83	-0.96	0.36	0.60	0	0	0	0	0	0
5	44	0.44	1.43	0.28	0.32	0	0	0	0	0	0
5	45	0.53	2.70	0.14	0.31	0	0	0	0	1	0
6	1	0.67	3.63	0.06	0.27	0	0	0	0	1	0
6	2	0.34	0.57	0.29	0.33	0	0	0	0	0	0
6	3	-0.78	-2.32	0.41	0.58	0	0	0	0	0	0
6	4	-0.62	-7.17	0.62	0.55	0	0	0	1	0	0
6	5	2.49	1.74	0.20	0.06	0	1	0	0	0	0
6	6	0.65	3.13	0.09	0.28	0	0	0	0	1	0
6	7	-0.79	2.26	0.24	0.58	0	0	0	0	0	0
6	8	-0.95	1.29	0.27	0.62	0	0	0	0	0	0
6	9	-0.10	2.31	0.22	0.43	0	0	0	0	0	0
6	10	-1.01	-4.12	0.51	0.63	0	0	0	1	0	0
6	11	-2.06	-2.55	0.47	0.82	0	0	0	0	0	0
6	12	0.32	-0.87	0.40	0.34	0	0	0	0	0	0
6	13	-1.31	-2.44	0.43	0.69	0	0	0	0	0	0
6	14	0.82	-1.91	0.51	0.25	0	0	0	0	0	0
6	15	1.55	1.31	0.24	0.14	0	1	0	0	0	0
6	16	0.97	-0.63	0.42	0.22	0	0	0	0	0	0
6	17	0.15	2.19	0.21	0.38	0	0	0	0	0	0
6	18	-0.53	5.68	0.06	0.53	0	0	1	0	1	1
6	19	0.82	5.22	-0.13	0.25	0	0	1	0	1	1
6	20	1.17	2.78	0.10	0.19	0	1	0	0	1	0
6	21	-0.71	-4.79	0.53	0.57	0	0	0	1	0	0
6	22	0.18	-3.88	0.56	0.37	0	0	0	0	0	0
6	23	-0.30	-1.94	0.41	0.47	0	0	0	0	0	0
6	24	-0.11	-1.70	0.40	0.43	0	0	0	0	0	0
6	25	-1.82	-1.64	0.40	0.78	0	0	0	0	0	0
6	26	0.70	-0.20	0.41	0.27	0	0	0	0	0	0
6	27	0.28	-1.03	0.41	0.35	0	0	0	0	0	0
6	28	-0.38	0.58	0.29	0.49	0	0	0	0	0	0
6	29	-0.36	1.03	0.27	0.49	0	0	0	0	0	0
6	30	0.49	0.22	0.33	0.31	0	0	0	0	0	0
6	31	0.31	0.14	0.34	0.34	0	0	0	0	0	0
6	32	-0.61	-3.81	0.48	0.54	0	0	0	0	0	0
6	33	-0.72	-7.05	0.63	0.57	0	0	0	1	0	0
6	34	0.15	-3.12	0.52	0.37	0	0	0	0	0	0
6	35	-0.40	0.21	0.33	0.50	0	0	0	0	0	0
6	36	0.48	3.78	0.06	0.31	0	0	0	0	1	0
6	37	0.44	2.24	0.21	0.32	0	0	0	0	0	0
6	38	0.89	2.29	0.13	0.24	0	0	0	0	1	0
6	39	-0.66	0.53	0.31	0.55	0	0	0	0	0	0
6	40	0.74	0.00	0.32	0.26	0	0	0	0	0	0
6	41	0.16	2.78	0.16	0.37	0	0	0	0	0	0
6	42	-0.03	5.14	0.08	0.41	0	0	1	0	1	1

6	43	-0.64	-1.41	0.37	0.55	0	0	0	0	0	0
6	44	0.22	-0.99	0.39	0.36	0	0	0	0	0	0
6	45	0.26	5.58	-0.03	0.35	0	0	1	0	1	1

Table 5 displays the results of the differential item functioning analyses for

Table 5: Summary of Item DIF Statistics

Form	Item	White	Other	Race DIF	Male	Female	Gender DIF	College	No College	Parent DIF
Link	1	-1.34	-1.22	0	-1.34	-1.04	0	-1.34	-1.37	0
Link	2	0.12	1.08	1	0.12	0.76	1	0.27	0.71	0
Link	3	0.16	0.31	0	0.16	0.32	0	0.20	0.33	0
Link	4	0.66	0.59	0	0.66	0.71	0	0.85	0.77	0
Link	5	0.65	0.54	0	0.65	0.59	0	0.67	0.75	0
Link	6	-0.35	-0.26	0	-0.35	-0.31	0	-0.45	-0.35	0
Link	7	-0.42	-0.44	0	-0.42	-0.58	0	-0.70	-0.52	0
Link	8	0.66	0.65	0	0.66	0.61	0	0.65	0.86	0
Link	9	1.28	1.36	0	1.28	1.29	0	1.44	1.52	0
Link	10	0.72	0.68	0	0.72	0.74	0	0.82	0.87	0
Link	11	-0.94	-1.04	0	-0.94	-1.16	0	-1.26	-1.25	0
Link	12	-0.02	-0.03	0	-0.02	0.01	0	-0.03	0.00	0
Link	13	0.33	0.19	0	0.33	0.40	0	0.55	0.40	0
Link	14	0.84	0.72	0	0.84	0.52	0	0.71	0.77	0
Link	15	0.27	0.26	0	0.27	0.16	0	0.16	0.29	0
1	1	-0.64	-0.43	0	-0.64	-0.53	0	-0.60	-0.74	0
1	2	1.09	0.72	0	1.09	0.77	0	1.08	1.07	0
1	3	-2.00	-1.87	0	-2.00	-1.68	0	-1.97	-2.30	0
1	4	0.84	1.29	0	0.84	1.09	0	0.82	1.38	1
1	5	1.68	1.59	0	1.68	1.64	0	2.00	1.91	0
1	6	1.54	1.32	0	1.54	1.16	0	1.43	1.55	0
1	7	0.22	-0.03	0	0.22	0.11	0	0.39	0.09	0
1	8	-0.80	-0.23	1	-0.80	-0.24	1	-0.48	-0.64	0
1	9	-0.67	-0.50	0	-0.67	-0.33	0	-0.56	-0.52	0
1	10	0.00	0.11	0	0.00	0.20	0	0.01	0.22	0
1	11	0.37	0.26	0	0.37	0.35	0	0.46	0.53	0
1	12	0.22	0.51	0	0.22	0.31	0	0.37	0.29	0
1	13	-0.41	-0.27	0	-0.41	-0.38	0	-0.56	-0.51	0
1	14	0.26	0.07	0	0.26	0.03	0	0.13	0.12	0
1	15	1.88	1.64	0	1.88	1.27	1	1.50	1.94	0
1	16	0.49	0.38	0	0.49	0.13	0	0.22	0.35	0
1	17	0.88	0.99	0	0.88	0.76	0	1.00	0.90	0
1	18	-0.35	-0.39	0	-0.35	-0.14	0	-0.34	-0.22	0
1	19	0.47	0.28	0	0.47	0.39	0	0.31	0.65	0
1	20	1.06	1.13	0	1.06	1.19	0	1.37	1.30	0
1	21	0.28	0.21	0	0.28	0.40	0	0.51	0.31	0
1	22	-0.42	-0.65	0	-0.42	-0.62	0	-0.49	-0.71	0
1	23	-0.56	-0.52	0	-0.56	-0.35	0	-0.66	-0.40	0
1	24	0.22	-0.03	0	0.22	0.06	0	0.12	0.14	0
1	25	-0.94	-1.13	0	-0.94	-1.14	0	-1.20	-1.26	0
1	26	0.44	0.39	0	0.44	0.41	0	0.32	0.56	0
1	27	0.40	0.65	0	0.40	0.57	0	0.24	0.78	1
1	28	-0.22	0.26	0	-0.22	0.26	0	-0.21	0.28	0
1	29	-0.51	-0.62	0	-0.51	-0.43	0	-0.65	-0.44	0

1	30	0.02	-0.08	0	0.02	0.02	0	0.07	-0.07	0
1	31	0.11	-0.28	0	0.11	-0.19	0	-0.10	-0.06	0
1	32	-0.60	-0.34	0	-0.60	-0.61	0	-0.71	-0.76	0
1	33	-0.59	-0.63	0	-0.59	-0.74	0	-0.75	-0.87	0
1	34	-0.15	-0.24	0	-0.15	-0.28	0	-0.07	-0.44	0
1	35	0.33	-0.05	0	0.33	-0.02	0	0.24	0.15	0
1	36	0.22	0.25	0	0.22	0.15	0	0.06	0.24	0
1	37	-1.17	-1.10	0	-1.17	-1.06	0	-1.39	-1.27	0
1	38	-0.27	0.07	0	-0.27	0.18	0	-0.07	0.01	0
1	39	0.53	0.38	0	0.53	0.66	0	0.88	0.65	0
1	40	-0.45	-0.28	0	-0.45	-0.38	0	-0.71	-0.41	0
1	41	0.50	0.39	0	0.50	0.38	0	0.49	0.57	0
1	42	0.01	-0.07	0	0.01	0.29	0	0.33	0.07	0
1	43	-0.48	-0.62	0	-0.48	-0.57	0	-0.45	-0.69	0
1	44	-0.79	-0.78	0	-0.79	-0.93	0	-1.03	-1.04	0
1	45	0.12	0.21	0	0.12	0.35	0	0.15	0.40	0
2	1	-0.09	-0.02	0	-0.09	0.12	0	0.01	0.10	0
2	2	0.49	0.48	0	0.49	0.61	0	0.45	0.80	0
2	3	0.61	0.27	0	0.61	0.31	0	0.79	0.34	0
2	4	-0.98	-0.68	0	-0.98	-0.44	1	-0.70	-0.79	0
2	5	-1.08	-0.71	0	-1.08	-0.76	0	-1.08	-1.06	0
2	6	2.00	1.70	0	2.00	1.72	0	2.11	2.10	0
2	7	-0.59	-0.15	0	-0.59	-0.17	0	-0.49	-0.33	0
2	8	-2.19	-1.83	0	-2.19	-1.83	0	-2.20	-2.37	0
2	9	0.47	0.25	0	0.47	0.14	0	0.52	0.23	0
2	10	0.21	0.42	0	0.21	0.19	0	0.11	0.25	0
2	11	-0.93	-1.09	0	-0.93	-1.10	0	-1.15	-1.28	0
2	12	-0.58	-0.47	0	-0.58	-0.35	0	-0.62	-0.50	0
2	13	-0.13	0.11	0	-0.13	0.02	0	-0.02	-0.08	0
2	14	0.32	0.56	0	0.32	0.35	0	0.31	0.38	0
2	15	-0.06	-0.23	0	-0.06	-0.42	0	-0.32	-0.34	0
2	16	-0.09	-0.21	0	-0.09	0.04	0	0.05	-0.14	0
2	17	-0.28	-0.54	0	-0.28	-0.37	0	-0.25	-0.48	0
2	18	-0.06	-0.13	0	-0.06	-0.13	0	-0.03	-0.21	0
2	19	1.28	1.25	0	1.28	1.09	0	1.61	1.32	0
2	20	0.48	0.26	0	0.48	0.32	0	0.71	0.37	0
2	21	-0.31	-0.55	0	-0.31	-0.72	0	-0.70	-0.57	0
2	22	0.29	0.27	0	0.29	0.21	0	0.02	0.42	0
2	23	0.03	0.21	0	0.03	0.14	0	-0.07	0.22	0
2	24	-0.45	-0.35	0	-0.45	-0.44	0	-0.55	-0.57	0
2	25	-0.68	-0.32	0	-0.68	-0.55	0	-0.91	-0.62	0
2	26	0.45	0.33	0	0.45	0.58	0	0.65	0.52	0
2	27	0.81	0.78	0	0.81	0.67	0	0.76	0.89	0
2	28	0.40	0.70	0	0.40	0.71	0	0.56	0.72	0
2	29	-0.14	0.30	0	-0.14	0.35	0	-0.19	0.41	1
2	30	0.51	0.51	0	0.51	0.44	0	0.67	0.47	0
2	31	0.88	0.87	0	0.88	0.86	0	1.26	0.86	0
2	32	-0.53	-0.68	0	-0.53	-0.82	0	-0.70	-0.93	0
2	33	0.81	0.83	0	0.81	0.93	0	0.99	1.08	0
2	34	0.32	0.16	0	0.32	0.12	0	0.10	0.35	0

2	35	-1.39	-1.63	0	-1.39	-1.70	0	-1.90	-1.76	0
2	36	0.41	0.39	0	0.41	0.49	0	0.50	0.52	0
2	37	-0.14	-0.45	0	-0.14	-0.27	0	0.09	-0.49	1
2	38	1.71	1.85	0	1.71	1.86	0	2.03	2.30	0
2	39	-0.16	-0.09	0	-0.16	0.10	0	-0.06	0.02	0
2	40	0.00	0.28	0	0.00	0.18	0	0.08	0.15	0
2	41	0.45	0.28	0	0.45	0.23	0	0.27	0.48	0
2	42	-0.64	-0.89	0	-0.64	-0.92	0	-0.87	-1.08	0
2	43	-0.60	-0.57	0	-0.60	-0.67	0	-0.78	-0.72	0
2	44	-0.36	-0.58	0	-0.36	-0.37	0	-0.32	-0.49	0
2	45	0.34	0.28	0	0.34	0.31	0	0.54	0.28	0
3	1	-0.73	-1.13	0	-0.73	-0.78	0	-0.68	-0.89	0
3	2	0.55	0.59	0	0.55	0.76	0	0.86	0.67	0
3	3	-0.84	-0.76	0	-0.84	-0.62	0	-0.93	-0.76	0
3	4	0.10	0.10	0	0.10	0.19	0	0.15	0.22	0
3	5	0.67	0.92	0	0.67	0.85	0	0.74	0.93	0
3	6	1.25	1.55	0	1.25	1.52	0	1.57	1.63	0
3	7	-1.57	-1.23	0	-1.57	-1.33	0	-1.43	-1.77	0
3	8	-1.62	-1.46	0	-1.62	-1.43	0	-1.72	-1.86	0
3	9	0.65	0.82	0	0.65	1.00	0	0.72	1.21	0
3	10	-2.51	-2.86	0	-2.51	-2.80	0	-2.55	-3.39	1
3	11	0.24	0.42	0	0.24	0.76	1	0.51	0.68	0
3	12	1.00	0.94	0	1.00	0.63	0	0.89	0.98	0
3	13	-0.98	-0.88	0	-0.98	-0.80	0	-1.09	-0.99	0
3	14	0.04	0.23	0	0.04	0.31	0	0.12	0.33	0
3	15	-1.19	-1.13	0	-1.19	-1.06	0	-1.22	-1.33	0
3	16	-0.54	-0.61	0	-0.54	-0.57	0	-0.68	-0.65	0
3	17	-0.90	-0.75	0	-0.90	-0.90	0	-0.94	-1.20	0
3	18	-0.09	-0.06	0	-0.09	-0.27	0	-0.27	-0.16	0
3	19	0.59	0.53	0	0.59	0.70	0	1.08	0.67	0
3	20	0.57	0.51	0	0.57	0.65	0	0.83	0.66	0
3	21	-0.74	-0.58	0	-0.74	-0.39	0	-0.65	-0.62	0
3	22	0.46	0.40	0	0.46	0.63	0	0.79	0.63	0
3	23	0.17	0.55	0	0.17	0.41	0	0.06	0.57	1
3	24	0.09	0.52	0	0.09	0.61	1	0.35	0.45	0
3	25	0.98	0.83	0	0.98	1.00	0	1.41	1.08	0
3	26	0.22	-0.08	0	0.22	-0.03	0	0.25	0.04	0
3	27	0.73	0.55	0	0.73	0.59	0	0.83	0.68	0
3	28	-0.20	-0.36	0	-0.20	-0.56	0	-0.48	-0.39	0
3	29	-0.83	-0.92	0	-0.83	-1.02	0	-0.97	-1.16	0
3	30	-0.11	0.03	0	-0.11	-0.20	0	-0.29	-0.15	0
3	31	0.52	0.64	0	0.52	0.84	0	0.89	0.74	0
3	32	-0.05	0.00	0	-0.05	-0.20	0	-0.32	-0.18	0
3	33	-0.13	-0.18	0	-0.13	0.21	0	0.02	0.08	0
3	34	0.73	0.77	0	0.73	1.10	0	0.94	1.17	0
3	35	-0.33	-0.19	0	-0.33	-0.51	0	-0.51	-0.52	0
3	36	0.60	0.54	0	0.60	0.71	0	0.71	0.85	0
3	37	-0.71	-0.58	0	-0.71	-0.48	0	-0.68	-0.73	0
3	38	0.48	0.63	0	0.48	0.80	0	0.65	0.91	0
3	39	1.24	1.00	0	1.24	0.99	0	1.31	1.16	0

3	40	0.23	0.43	0	0.23	0.52	0	0.45	0.44	0
3	41	0.02	0.16	0	0.02	0.00	0	0.13	-0.07	0
3	42	-0.01	-0.25	0	-0.01	-0.24	0	-0.04	-0.20	0
3	43	-0.37	-0.45	0	-0.37	-0.22	0	-0.33	-0.28	0
3	44	0.17	0.09	0	0.17	0.06	0	0.33	0.08	0
3	45	0.03	0.18	0	0.03	0.06	0	0.06	-0.01	0
4	1	0.72	0.64	0	0.72	0.74	0	0.96	0.78	0
4	2	-0.50	-0.82	0	-0.50	-0.64	0	-0.53	-0.77	0
4	3	-1.50	-1.25	0	-1.50	-1.15	0	-1.48	-1.54	0
4	4	-0.79	-0.73	0	-0.79	-0.62	0	-0.74	-0.87	0
4	5	1.10	1.21	0	1.10	1.06	0	1.04	1.47	0
4	6	0.09	0.00	0	0.09	0.12	0	0.08	0.20	0
4	7	-0.44	-0.02	0	-0.44	-0.25	0	-0.34	-0.43	0
4	8	-0.18	0.05	0	-0.18	-0.12	0	-0.30	-0.01	0
4	9	-0.01	0.44	0	-0.01	0.18	0	0.01	0.18	0
4	10	0.80	0.80	0	0.80	0.91	0	0.90	1.08	0
4	11	0.15	0.31	0	0.15	0.27	0	-0.02	0.49	1
4	12	-0.92	-0.72	0	-0.92	-0.66	0	-0.92	-0.87	0
4	13	-0.23	-0.55	0	-0.23	-0.42	0	-0.39	-0.36	0
4	14	0.39	0.23	0	0.39	0.19	0	0.30	0.31	0
4	15	0.34	-0.05	0	0.34	0.30	0	0.67	0.11	1
4	16	-0.23	-0.18	0	-0.23	-0.22	0	-0.20	-0.29	0
4	17	-0.31	-0.41	0	-0.31	-0.42	0	-0.42	-0.45	0
4	18	-0.41	-0.64	0	-0.41	-0.58	0	-0.58	-0.60	0
4	19	1.16	0.70	0	1.16	0.67	0	1.23	0.86	0
4	20	0.74	0.54	0	0.74	0.59	0	0.71	0.81	0
4	21	0.66	0.77	0	0.66	0.85	0	0.71	1.06	0
4	22	0.76	0.75	0	0.76	0.65	0	0.74	0.83	0
4	23	0.25	-0.08	0	0.25	0.15	0	0.44	0.03	0
4	24	0.41	0.27	0	0.41	0.35	0	0.48	0.38	0
4	25	-0.17	-0.67	0	-0.17	-0.49	0	-0.19	-0.53	0
4	26	-0.88	-1.17	0	-0.88	-1.20	0	-1.26	-1.14	0
4	27	0.10	-0.04	0	0.10	0.01	0	0.13	0.02	0
4	28	0.26	0.16	0	0.26	0.19	0	0.21	0.33	0
4	29	-0.97	-0.56	0	-0.97	-0.66	0	-0.99	-0.88	0
4	30	-0.67	-0.67	0	-0.67	-0.86	0	-0.94	-0.84	0
4	31	-1.19	-1.20	0	-1.19	-1.38	0	-1.56	-1.45	0
4	32	-1.40	-1.42	0	-1.40	-1.52	0	-1.72	-1.67	0
4	33	1.14	0.89	0	1.14	0.83	0	1.28	0.98	0
4	34	-0.29	0.12	0	-0.29	-0.20	0	-0.49	-0.14	0
4	35	-0.48	-0.81	0	-0.48	-0.95	0	-0.88	-0.85	0
4	36	0.25	0.26	0	0.25	0.21	0	0.46	0.14	0
4	37	-0.32	-0.36	0	-0.32	-0.52	0	-0.54	-0.49	0
4	38	-0.02	-0.17	0	-0.02	-0.07	0	0.04	-0.13	0
4	39	0.53	0.66	0	0.53	0.77	0	0.77	0.74	0
4	40	0.15	0.11	0	0.15	0.27	0	0.13	0.31	0
4	41	0.65	0.72	0	0.65	0.96	0	0.95	0.95	0
4	42	-0.11	-0.31	0	-0.11	-0.19	0	-0.07	-0.29	0
4	43	-0.59	-0.85	0	-0.59	-0.84	0	-1.01	-0.67	0
4	44	-0.10	-0.20	0	-0.10	-0.09	0	-0.07	-0.15	0

4	45	0.71	0.29	0	0.71	0.45	0	0.69	0.56	0
5	1	0.38	0.23	0	0.38	0.34	0	0.53	0.35	0
5	2	0.59	0.69	0	0.59	0.62	0	0.77	0.73	0
5	3	-0.80	-0.42	0	-0.80	-0.23	1	-0.72	-0.52	0
5	4	-0.65	-0.50	0	-0.65	-0.41	0	-0.58	-0.67	0
5	5	-0.62	-0.46	0	-0.62	-0.53	0	-0.60	-0.76	0
5	6	1.20	1.11	0	1.20	0.83	0	1.29	1.05	0
5	7	-2.02	-2.33	0	-2.02	-2.31	0	-2.17	-2.69	1
5	8	-2.57	-2.53	0	-2.57	-2.69	0	-2.79	-3.27	0
5	9	-0.21	-0.21	0	-0.21	-0.43	0	-0.51	-0.28	0
5	10	-1.57	-1.61	0	-1.57	-1.75	0	-1.65	-2.09	0
5	11	0.91	0.92	0	0.91	0.65	0	0.64	1.01	0
5	12	0.55	0.63	0	0.55	0.55	0	0.75	0.64	0
5	13	-0.02	0.10	0	-0.02	-0.20	0	-0.22	-0.12	0
5	14	0.56	0.67	0	0.56	0.25	0	0.33	0.48	0
5	15	-0.81	-0.63	0	-0.81	-0.70	0	-0.82	-0.89	0
5	16	0.36	0.18	0	0.36	0.09	0	0.26	0.19	0
5	17	0.52	0.33	0	0.52	0.41	0	0.62	0.46	0
5	18	-0.10	-0.14	0	-0.10	0.13	0	-0.15	0.14	0
5	19	0.31	0.42	0	0.31	0.49	0	0.60	0.48	0
5	20	-0.48	-0.48	0	-0.48	-0.44	0	-0.48	-0.57	0
5	21	-0.25	-0.31	0	-0.25	-0.54	0	-0.56	-0.48	0
5	22	0.19	0.08	0	0.19	0.27	0	0.23	0.32	0
5	23	0.40	0.20	0	0.40	0.24	0	0.29	0.39	0
5	24	-0.12	-0.18	0	-0.12	-0.35	0	-0.31	-0.24	0
5	25	-0.44	-0.70	0	-0.44	-0.45	0	-0.38	-0.60	0
5	26	-0.04	0.08	0	-0.04	-0.08	0	-0.12	-0.05	0
5	27	0.41	0.35	0	0.41	-0.16	1	0.19	0.09	0
5	28	-0.17	-0.37	0	-0.17	-0.19	0	-0.17	-0.29	0
5	29	-0.18	-0.01	0	-0.18	-0.16	0	-0.08	-0.26	0
5	30	0.16	0.27	0	0.16	-0.04	0	-0.06	0.14	0
5	31	0.43	0.52	0	0.43	0.52	0	0.57	0.52	0
5	32	-1.09	-1.34	0	-1.09	-1.46	0	-1.44	-1.49	0
5	33	0.61	0.33	0	0.61	0.42	0	0.65	0.52	0
5	34	-0.67	-0.82	0	-0.67	-1.06	0	-0.96	-1.05	0
5	35	-0.12	0.34	0	-0.12	0.09	0	0.02	-0.01	0
5	36	1.00	0.90	0	1.00	1.04	0	1.19	1.15	0
5	37	0.00	0.07	0	0.00	0.16	0	0.16	0.04	0
5	38	-0.37	-0.61	0	-0.37	-0.60	0	-0.54	-0.56	0
5	39	0.29	0.21	0	0.29	0.32	0	0.24	0.37	0
5	40	0.46	0.41	0	0.46	0.49	0	0.69	0.48	0
5	41	0.83	0.68	0	0.83	0.70	0	0.98	0.73	0
5	42	1.06	0.71	0	1.06	0.70	0	0.88	1.06	0
5	43	-0.76	-0.98	0	-0.76	-0.78	0	-0.84	-0.94	0
5	44	0.34	0.30	0	0.34	0.45	0	0.34	0.59	0
5	45	0.50	0.34	0	0.50	0.46	0	0.67	0.52	0
6	1	0.82	0.37	0	0.82	0.49	0	0.90	0.53	0
6	2	0.44	0.35	0	0.44	0.20	0	0.43	0.30	0
6	3	-0.74	-0.81	0	-0.74	-0.74	0	-0.89	-0.82	0
6	4	-0.53	-0.28	0	-0.53	-0.62	0	-0.74	-0.65	0

6	5	2.18	2.67	0	2.18	2.48	0	2.81	2.60	0
6	6	0.66	0.46	0	0.66	0.54	0	0.76	0.70	0
6	7	-0.84	-0.56	0	-0.84	-0.69	0	-0.89	-0.81	0
6	8	-1.03	-0.73	0	-1.03	-0.80	0	-0.99	-1.09	0
6	9	-0.13	-0.19	0	-0.13	-0.04	0	-0.10	-0.10	0
6	10	-1.08	-0.62	0	-1.08	-0.84	0	-1.19	-1.05	0
6	11	-1.82	-1.80	0	-1.82	-2.04	0	-2.35	-2.16	0
6	12	0.27	0.35	0	0.27	0.34	0	0.31	0.38	0
6	13	-1.10	-1.36	0	-1.10	-1.33	0	-1.33	-1.51	0
6	14	0.86	1.09	0	0.86	0.69	0	0.74	1.02	0
6	15	1.38	2.16	1	1.38	1.53	0	1.75	1.62	0
6	16	0.96	0.95	0	0.96	0.86	0	0.94	1.15	0
6	17	0.35	-0.19	1	0.35	-0.06	0	0.27	0.08	0
6	18	-0.53	-1.01	0	-0.53	-0.48	0	-0.46	-0.65	0
6	19	0.75	0.51	0	0.75	0.78	0	1.02	0.80	0
6	20	1.15	0.96	0	1.15	1.01	0	1.21	1.32	0
6	21	-0.81	-0.28	1	-0.81	-0.53	0	-1.03	-0.54	0
6	22	0.20	0.35	0	0.20	0.14	0	0.16	0.21	0
6	23	-0.31	-0.44	0	-0.31	-0.23	0	-0.46	-0.21	0
6	24	-0.28	0.10	0	-0.28	0.06	0	-0.25	0.02	0
6	25	-1.78	-1.26	1	-1.78	-1.67	0	-2.21	-1.83	0
6	26	0.64	0.60	0	0.64	0.66	0	0.69	0.79	0
6	27	0.24	0.28	0	0.24	0.29	0	0.33	0.32	0
6	28	-0.41	-0.31	0	-0.41	-0.30	0	-0.25	-0.56	0
6	29	-0.50	-0.15	0	-0.50	-0.22	0	-0.27	-0.47	0
6	30	0.37	0.33	0	0.37	0.54	0	0.46	0.58	0
6	31	0.32	0.11	0	0.32	0.25	0	0.27	0.41	0
6	32	-0.47	-0.62	0	-0.47	-0.66	0	-0.71	-0.60	0
6	33	-0.60	-0.65	0	-0.60	-0.72	0	-1.01	-0.62	0
6	34	-0.06	0.52	1	-0.06	0.35	0	0.04	0.29	0
6	35	-0.46	-0.42	0	-0.46	-0.32	0	-0.36	-0.48	0
6	36	0.50	0.31	0	0.50	0.40	0	0.49	0.56	0
6	37	0.33	0.40	0	0.33	0.48	0	0.52	0.43	0
6	38	0.87	0.75	0	0.87	0.77	0	0.85	1.07	0
6	39	-0.45	-0.77	0	-0.45	-0.76	0	-0.60	-0.78	0
6	40	0.89	0.62	0	0.89	0.55	0	0.76	0.82	0
6	41	0.08	0.16	0	0.08	0.21	0	0.31	0.08	0
6	42	-0.02	-0.17	0	-0.02	-0.04	0	0.06	-0.11	0
6	43	-0.61	-0.79	0	-0.61	-0.60	0	-0.58	-0.74	0
6	44	0.08	0.38	0	0.08	0.32	0	0.11	0.33	0
6	45	0.07	0.17	0	0.07	0.38	0	0.29	0.32	0

References

- Draba, R. E. (1977). *The identification and interpretation of item bias* (Research Memorandum No. 26). Chicago, IL: University of Chicago.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, MA: Kluwer-Nijhoff.
- Linacre, J. M., & Wright, B. D. (2001). *A User's Guide to WINSTEPS/MINISTEP Rasch-Model Computer Program (Version 3.31)*. Chicago, IL: MESA Press.
- Raju, N. S. (1988). The Area Between Two Item Characteristic Curves. *Psychometrika*, 53, 495-502.
- Raju, N. S. (1990). Determining the Significance of Estimated Signed and Unsigned Areas Between Two Item Response Functions. *Applied Psychological Measurement*, 14, 197-207.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Appendix E

Bloom's Taxonomy of Educational Objectives

Major Categories in the Taxonomy of Educational Objectives (Bloom 1956)

(<http://faculty.washington.edu/krumme/guides/bloom.html>)

Categories in the Cognitive Domain: (with Outcome-Illustrating Verbs)

1. **Knowledge** of terminology; specific facts; ways and means of dealing with specifics (conventions, trends and sequences, classifications and categories, criteria, methodology); universals and abstractions in a field (principles and generalizations, theories and structures): Knowledge is (here) defined as the remembering (recalling) of appropriate, previously learned information.
 - defines; describes; enumerates; identifies; labels; lists; matches; names; reads; records; reproduces; selects; states; views.
2. **Comprehension**: Grasping (understanding) the meaning of informational materials.
 - classifies; cites; converts; describes; discusses; estimates; explains; generalizes; gives examples; makes sense out of; paraphrases; restates (in own words); summarizes; traces; understands.
3. **Application**: The use of previously learned information in new and concrete situations to solve problems that have single or best answers.
 - acts; administers; articulates; assesses; charts; collects; computes; constructs; contributes; controls; determines; develops; discovers; establishes; extends; implements; includes; informs; instructs; operationalizes; participates; predicts; prepares; preserves; produces; projects; provides; relates; reports; shows; solves; teaches; transfers; uses; utilizes.
4. **Analysis**: The breaking down of informational materials into their component parts, examining (and trying to understand the organizational structure of) such information to develop divergent conclusions by identifying motives or causes, making inferences, and/or finding evidence to support generalizations.
 - breaks down; correlates; diagrams; differentiates; discriminates; distinguishes; focuses; illustrates; infers; limits; outlines; points out; prioritizes; recognizes; separates; subdivides.
5. **Synthesis**: Creatively or divergently applying prior knowledge and skills to produce a new or original whole.
 - adapts; anticipates; categorizes; collaborates; combines; communicates; compares; compiles; composes; contrasts; creates; designs; devises; expresses; facilitates; formulates; generates; incorporates; individualizes; initiates; integrates; intervenes; models; modifies; negotiates; plans; progresses; rearranges; reconstructs; reinforces; reorganizes; revises; structures; substitutes; validates.
6. **Evaluation**: Judging the value of material based on personal values/opinions, resulting in an end product, with a given purpose, without real right or wrong answers.
 - appraises; compares & contrasts; concludes; criticizes; critiques; decides; defends; interprets; judges; justifies; reframes; supports.

Appendix F

Test Blueprint for Algebra I and Mathematics for the Technologies II

South Carolina End-of-Course Examination

Test Blueprint for Algebra I and Mathematics for the Technologies II

OPERATIONAL FORMS

Item Allocation Rules for Test Form Construction	Number of Items
I. Understanding Functions	19-21
A. Relationships	4-6
B. Linear and Quadratic Functions and Data Representations	4-6
C. Generalizations, Algebraic Symbols, and Matrices	3-5
D. Algebraic Expressions in Problem Solving Situations	5-7
II. Linear Functions	21-23
A. Representations	4-6
B. Interpretations	7-9
C. Equations and Inequalities	5-7
D. Systems of Linear Equations	2-4
III. Quadratic and Other Functions	7-9
A. Quadratic Functions	4-6
B. Other Functions	2-4

For the December 2002 and May 2003 administrations, the test will contain 50 operational items, distributed as represented in the table above, and an additional 10 embedded field-test items.

Appendix G

**Report on Standard Setting
Algebra I/Math for the Technologies II
Reported to the State Board of Education
August 13, 2002**

Setting Achievement Levels (Performance Standards) for the Algebra I/ Mathematics for the Technologies II End-of-Course Examination

The purpose of the South Carolina End-of-Course Examination Program (EOCEP) is defined by State Board of Education regulation:

- (1) The tests shall promote instruction in the specific academic standards for the courses, encourage student achievement, and document the level of student mastery of the curriculum standards.
- (2) The tests shall serve as indicators of the program, school, and school district effectiveness in the manner prescribed by the Education Oversight Committee in accordance with the provisions of the Education Accountability Act of 1998 (EAA).
- (3) The tests shall be weighted 20 percent in the determination of the students' final grades in the gateway courses.

The first subject area addressed by this new program is Algebra I, the content of which is covered by both Algebra I and Mathematics for the Technologies II courses. The results of the examination will be used, beginning in the 2003-2004 school year, as 20% of a student's final grade in the course.

A vital part of the EOCEP is to make the results of the examinations translatable to teachers for use in the final grading of students. As a first step in making the results understandable and useful to teachers, staff from the American Institutes for Research (AIR) conducted a standard setting session on July 15 and 16, 2002 in Columbia, SC. The Bookmark Method, which was required by the contract for EOCEP and which was recommended and used in the PACT program, was the standard setting method used. The purpose of this standard setting was for the content experts to recommend a cut point at the place that would determine the amount of knowledge a minimally passing student (one at the D/F cut point) would need to know.

South Carolina Department of Education (SDE) staff recruited 30 mathematics content experts to set standards for the Algebra I/Mathematics for the Technologies II End-of-Course Examination Program. These experts were divided into 5 groups of 6 panelists each for the purpose of facilitating small group discussions. The Department assigned one member of each table to be the table leader. AIR staff trained the table leaders during the lunch break on the first day, before any of their table leader responsibilities began. AIR began the standard setting workshop by first presenting the panelists with an overview, training them on both the Bookmark Method and on the South Carolina standards and materials, followed by three rounds of bookmark placement, and provision of summary results and feedback.

Standard Setting Workshop Overview

The training was organized into two parts: a general overview of standard setting in general and the Bookmark Method in particular, that included time for practice, and a specific orientation to South Carolina course standards, test items, and the Performance Level Descriptor (PLD). The session began with a review of the purpose and agenda. The purpose, as stated for the panelists, was:

To recommend to the South Carolina Department of Education the level of achievement, relative to the Algebra I Course Standards as measured by the EOCEP Algebra I/Mathematics for the Technologies II examination, that a student should achieve to pass minimally (at the D/F cut point).

Panelists were told that their task was to determine how much of the South Carolina Algebra I course standards, as measured by the EOCEP examination, a student must know and be able to do in order to pass minimally (i.e., just meet "D" level performance). AIR staff also discussed the issue of confidentiality and reminded panelists that the test specifications, items, and cut scores must remain confidential and secure. All panelists signed non-disclosure forms, agreeing not to discuss the content of anything they saw.

Training

AIR staff led the training of both the table leaders and the group as a whole. During lunch on Day 1, an AIR staff member trained the table leaders on their role as small group leaders and facilitators. Another

AIR staff member trained all the panelists on using the Bookmark Method, and a third member trained the panelists on the South Carolina Algebra standards and materials.

Training on the Bookmark Method

AIR staff presented an overview of standard setting, including a discussion on scaling, and then focused specifically on the bookmark task. The panelists learned that the Bookmark Method is a procedure for setting performance standards that has been used in more than 20 states and districts and that has withstood legal challenges. The panelists received the following instructions for placing a bookmark.

1. Read each item in the Ordered Item Booklet and identify the knowledge and skills required to respond successfully to the item.
2. Review the definition for passing minimally at the D/F level and compare it to the course standards.
3. Find the location in the item set that separates groups of examinees into categories and then place a bookmark at that location in the Ordered Item Booklet.

The panelists practiced using nine items from the NAEP 4th grade mathematics assessment.

Training on SDE Course Standards, Test Design, and the Performance Level Descriptor

AIR's lead mathematics developer conducted the training on the South Carolina course standards and test specifications. She gave each panelist a copy of the South Carolina Algebra I Course Standards and explained how the pool of items was developed using the standards. These specifications included a general description of the test, a content outline indicating the percentage of items assessing each strand, the goals for the different content strands, and the alignment between the item content and the course standards. Panelists were instructed to use this document to familiarize themselves with the content standards, with how the test was designed, and with what students were specifically expected to know.

As a next step, panelists were asked to answer all items in their ordered item booklet as if they were taking a test. It took approximately 40 minutes for the panelists to answer all 80 items. Answer keys were then distributed and panelists were given the opportunity to ask any questions.

Next, the South Carolina Performance Level Descriptor, defining minimally passing performance at the D/F level for Algebra I/Mathematics for the Technologies II, was introduced:

A minimally passing student (a student at the D/F cut point) operates at a predominately concrete rather than an abstract level. Whereas minimally passing students may demonstrate conceptual understanding of simple algebraic topics in familiar contexts, they exhibit difficulty transferring knowledge to new situations. Minimally passing students have a limited understanding of procedural knowledge, and have minimal facility with algebraic problem solving.

Panelists were informed that they would not be viewing an intact test form, rather that they would work with a representative subset of 80 items from the item pool. The field tests consisted of 60 items, while the operational tests will be composed of 50 scored items.

Reviewing the Algebra I/Mathematics for the Technologies II Ordered Item Booklet

After receiving training on the Algebra I/Mathematics for the Technologies II material, the panelists took about two hours to review the Ordered Item Booklets in their small groups. During this review, they were to ask themselves two questions:

1. What does this item measure?
2. What makes it more difficult than the previous item?

The table leaders facilitated this discussion. The purpose of this exercise was for the panelists to gain a common understanding about what knowledge and skills each item requires. This stage is considered key

to setting a reasonable standard based on the amount of knowledge and skills students should have to pass the Algebra I/Mathematics for the Technologies II test at the D/F line.

Round 1

In Round 1, panelists were instructed to place a bookmark after the hardest item that those students passing minimally, at the D/F cut point, would be able to answer correctly, and to record the page number of their bookmarks on their rating forms. Table leaders first ensured that each panelist understood the task and had no questions about any of the materials, they then instructed their table members to place their bookmark. After the table leaders checked the accuracy of the panelists' rating forms against their bookmarks in the Ordered Item Booklets, they gave the forms to an AIR staff member. This step ended Day 1 for the panelists who, after securing their materials, were then dismissed for the day.

The individual median was calculated.

Round 2

AIR staff presented the Round 1 results to the panelists on the morning of Day 2. The results were given first to the table leaders, who were informed of the overall median. The table leaders then conducted group discussions that emphasized a focus on the PLD, content standards, and the bookmark task.

Day 2 of the workshop began with a large group discussion led by AIR staff, who reiterated the task and focused on the definition of a minimally passing student. Then, the discussion was turned over to the table leaders who provided their table members with feedback for their table. The panelists then reviewed the items that fell between the high and low bookmarks at their table, and discussed the standards and the PLD. Panelists spent about an hour in these small group discussions.

After completing their discussions, panelists again placed their bookmarks. They submitted their rating forms to the table leaders who checked them for accuracy and then submitted all forms to an AIR staff member before taking a break.

Round 3

Round 3 began with a large group discussion in which the AIR facilitator presented the median for each table, the overall median, and the highest and lowest individual bookmarks.

Each table leader then presented a summary of their table's discussion. One particular debate occurred between two of the tables. One table was concerned about the students' opportunity to learn. Because the accountability was at the student level, panelists did not want to punish students with a failing grade if teachers had not taught the material. The other table wanted to use the performance standard to push teachers to teach the material and was willing to suffer through high student failure rates to achieve this.

Once each table had been given an opportunity to express their opinions and everyone had responded, the discussion was turned back over to the table leaders. The table leaders then presented their table members with their individual results and conducted a further discussion on the appropriateness of the table's highest and lowest cut scores.

Next, panelists were given an opportunity to place a final bookmark. Again, they submitted their forms to the table leaders who checked them for accuracy and then submitted all forms to an AIR staff member before taking a break.

Impact Data

At this point, AIR and SDE staff looked at the consequence, or impact, data from the field-test administration, which indicated the percentage of students that would pass Algebra I/Mathematics for the Technologies II with a cut score based on the committee recommendation. Overall, 22% of the students would pass. The Department wanted to share these results with the panelists and weigh their reactions. AIR then assembled the necessary forms for an extra (fourth) round and then presented these results to the panelists.

Extra Round

AIR staff explained to the panelists that one extra round of ratings would be conducted with an additional piece of information. Once again, panelists were told that their task was criterion based and dependent on

the standards, the PLD, and their knowledge of the items. However, they also were told that they would receive some normative data and that we wanted to learn whether that data would change their ratings. Panelists were advised that these data were based on a field test sample that potentially could have included non-motivated students and/or students who had not yet been taught all the standards. After this caveat, panelists were told that with the current cut score, only 22% of the students would pass.

The AIR facilitator asked the panelists to discuss how much, if at all, this information would change their placement of the bookmark. If they still felt strongly about their decision in Round 3, they were instructed to simply copy that line from the previous rating form onto their new rating form. If, however, they wanted to adjust their bookmark with the new data, they were instructed to go through the entire process of finding the point in their Ordered Item Booklet. Panelists then entered their new cut scores on the rating forms and submitted them to an AIR staff member.

Results

Overall, panelists did not change their ratings by very much. The overall median decreased slightly, increasing the pass rate from 22% to 25%. For the most part, the panelists felt that students needed to understand a certain portion of the material in order to pass the test. They were surprised by the high failure rate, and while some panelists moved their bookmarks in response, others were unwilling to modify their cut scores significantly, resulting in similar medians.

Modifying the Performance Level Descriptor

As a final step in the process, AIR staff described a method for modifying the PLD for a minimally passing student. Panelists were told that the purpose of this exercise was to expand the PLD with specific content assessed in the Algebra I/Mathematics for the Technologies II test that minimally passing students are expected to know and be able to do. AIR's lead developer provided a caveat that the new information should not be overly specific or address any single item. The new descriptor was to be generalizable to any Algebra I/Mathematics for the Technologies II course in South Carolina and was meant to clarify the reports of the test results to teachers, parents, and other stakeholders.

Each table produced one PLD to share with the large group. The AIR facilitator typed each new PLD into an MS PowerPoint slide and then facilitated the large group discussion.

Each table leader described to the large group how they had modified the new PLD and why they made the changes they did. Then, the entire group voted on their favorite PLD, which the group modified until everyone was satisfied with the final definition, which reads as follows:

A minimally passing student (a student at the D/F cut point)

- performs at a predominately concrete (numerical) rather than an abstract (symbolic) level
- may demonstrate conceptual understanding of simple algebraic topics in familiar contexts, such as numeric/tabular, graphical, verbal, or symbolic representations
- has limited procedural knowledge
- has minimal skills with algebraic thinking and problem solving
- has difficulty transferring knowledge to new situations.

Panelists agreed that it was important to define concrete and abstract thinking. They also wanted to add more to the description "simple algebraic topics in familiar contexts." There was one discussion about the difference between algebraic thinking and problem solving that was never resolved, but the consensus of the group was to use both terms. Finally, the panelists were a bit dissatisfied with the parallelism of the bullets and the punctuation, but agreed to leave those editorial modifications to the Department.

Overall Results

Throughout the process, the median cut score decreased and panelists drew closer to a convergence. The median cut score dropped dramatically between Rounds 1 and 2, and although it did not decrease between Rounds 2 and 3, the range decreased, indicating that more panelists were converging around the final cut score.¹

Evaluation Results

At the end of the standard setting workshop, panelists were asked to complete an evaluation form to rate the quality of the training, materials, process, and outcome of the workshop. Overall, the results were very positive. Out of 30 panelists,² 28 agreed or strongly agreed that

- purposes and goals of the workshop were clearly understood,
- training covered all the necessary information,
- the item map was easy to understand, and
- the ordered item booklet was easy to understand.

In addition, 27 panelists said they were satisfied with the final cut score, and 28 indicated that the final cut score is criterion based and that the group valued everyone's opinion. All 29 panelists who completed the evaluation indicated that they were satisfied with the assistance of facilitators, development and content leaders, and table leaders.

The following tables display some specific recommendations and comments.

¹ Note that we are not including Round 4 in the overall summary, because it was not part of the official, standard setting procedure.

² One panelist did not complete the evaluation.

Recommendations/Concerns

I suggest that the real implementation of this be done next year, 2003/2004 academic year, so that math teachers starting from 8th grade to 9th grade will be ready.
I'm concerned about validity of the Ordered Item Booklet given information from teachers concerning the students in the sample. This may have thrown off the entire process. Overall, this method of standard setting seems pretty fair, equitable, comprehensive, etc. I seriously question the degree of difficulty on this end of course test.
Regarding question #12: Confident except I question the original data set.
For the final activities, consider a "gallery wall" to see what others are doing. Then have each rewrite, then, whole group you did.
The PLD is not grammatically correct. Decision needs to be made on how state will test the students. For example, same day, time, etc. Will be necessary for security.

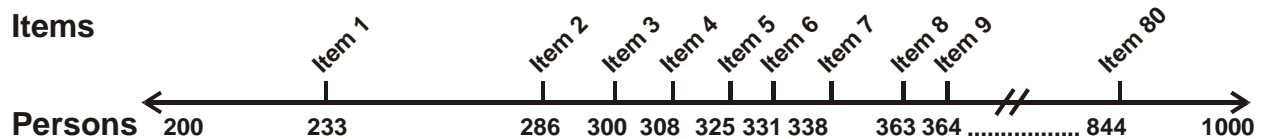
Accolades

Liked having the opportunity to pose questions, make suggestions and engage in open discussions.
Liked the excellent organization.
Great learning experience.
Excellent. I am impressed with the preparation that has already gone into this standard setting.
I am happy that the opinions of classroom teachers are being sought.
Small group discussions are excellent, I was happy to hear the opinions of my colleagues.
Thank you again for thinking about this standard setting. South Carolina really needs a standard for Algebra and education in general.
Workshop leaders were helpful to us in doing the difficult task of standard setting. A good workshop in profession
I liked having the opportunity of discussing in a small group before a large group discussion. Having the opportunity to give individual input as to what the cut-off should be.
The folks who were at my table were very responsive, willing to discuss and negotiate their thoughts and opinions, and caring about the needs of students. We had the best table!

The South Carolina Technical Advisory Committee (TAC) met on Friday, July 19, 2002, to review the results of the standard setting. After considering all of the factors, the committee recommended that the standard be set at the Round 3 bookmark minus two standard errors. This would all but eliminate the possibility of false negative classification (pass/fail) errors. Based on the distribution of student performance from the May field test, approximately 49% of the students would not achieve this level of performance.

Scaling

Test forms differ somewhat in difficulty due to differing item sets. The same percent correct score on different forms can mean different levels of student achievement due to the variation in the difficulty of the items on the forms.



Scales form the "common currency" on which to place and report scores. This process is called equating. When tests are equated, any particular scale score represents the same level of student achievement no matter which form is taken.

One step in the test development process for the Algebra I/ Math Tech. II End-of-Course tests consists of developing a scale for reporting test scores. A number of constraints influence the nature of that scale. Foremost among these is the requirement that students' scores account for twenty percent of their final course grade. Since the metric for final course grades is determined by the state's Uniform Grading Scale, the simplest scale for reporting EOCEP scores is one matching the uniform scale. This scale ranges from 0

to 100, with scores from 93 to 100 corresponding to the letter grade of “A”, 85-92 to “B”, 77-84 to “C”, 70-76 to “D”, and all grades below 70 to “F”.

Six field test forms were administered during the 2001-02 school year. Data from the field tests were used to calibrate the items using an IRT model (the one-parameter logistic or Rasch model). Once calibrated, each item has a unique item difficulty value used in subsequent procedures.

Another constraint on the scoring scale is that the scale be interval and linear with respect to achievement. It is characteristic of IRT scaling models that they produce scores that are on an interval scale. Thus the desired scoring scale for the EOCEP tests would correlate with, indeed be a linear transformation of, the Rasch scale. Two points are required to define such a linear transformation. One point (the Rasch score equivalent to 70) was defined by the standard-setting committee and the TAC. The other point can be selected logically. One obvious such point is zero. It seems quite reasonable to equate zero correct answers with zero points on the test scoring scale. The zero point is not available directly from the pool of item difficulties. A prototype test form must be developed in order for a Rasch ability score corresponding to zero correct answers to be determined. Then, the interval between the Rasch zero score and the Rasch score equivalent to 70 can be equated to the 70 point interval between 0 and 70, so that the Rasch scale increment for 1 point on the scoring scale can be computed. This increment could then be applied to determine the minimum Rasch score necessary to achieve each scale score point on the reporting scale.

A score of 0 will be assigned to students who answer no items correctly. All other scores will be derived through the linear transformation of the Rasch scale to the uniform grading scale (0-100) as described in the paragraph above. The highest score on the test will be 100.

Confirmatory Standard Setting

The standards that have been set will become operational for the baseline academic year, 2002-2003. Scores for students who take the Algebra I/Mathematics for the Technologies II tests at the end of the first semester and at the end of the school year will be reported to schools and districts in roster form for information. During the baseline year, the scores will not count as part of the students’ grades in the course nor will they count as part of the accountability system. Additional item development will continue during this baseline year as only half of the required number of test forms have been developed.

The initial achievement levels will be revisited after the spring test administration. Since the 2002-2003 administrations will be “operational” administrations (as opposed to field tests), it is likely that student motivation to perform well on the tests will be greater than it was for the field tests used in the initial standard setting. In addition, teachers are likely to be more familiar with the content standards for Algebra I and classroom instruction is likely to be more closely focused to those standards than during the previous school year. More teachers and students will be aware of the standards requirements and their basis for the examinations.

Appendix H

**Calculator Use Policy
Algebra I/Math for the Technologies II Field Test
Office of Assessment
SC Department of Education**

Calculator Use

The EOCEP test items were written so that a calculator is not required; however, calculator use is allowed.

To maintain test security, the test administrator **must** clear the memory of every test taker's calculator **before AND after** the testing period. This will prevent test takers from potentially carrying answers into the test as well as prevent them from taking copies of secure test items from the testing room. If a student refuse to have his or her calculator's memory cleared, that calculator **may not** be used during testing.

Test takers may use graphing and other calculators during the test. In the interest of test security, the following types of calculators **may not** be used: pocket organizers; Palm Pilots, Visor, or other Palm-based devices; handheld PCs that use a writing tablet or QWERTY (typewriter) keyboard; calculators that have symbolic manipulation capabilities, such as the Casio FX 2.0, Hewlett Packard hp40g or 49g, or Texas Instruments TI-89 or 92; or calculators that "talk" or otherwise make noise, such as the AudioCalc.

You may provide calculators, according to school policy, but they must meet the above criteria.