



Functional Independence

ELA, Mathematics, and Science

Michigan's Alternate Assessment Program

Michigan Department of Education
Bureau of Assessment and Accountability

Technical Report
2010-2011 Addendum

August 2011

TABLE OF CONTENTS

LIST OF TABLES.....	iii
INTRODUCTION.....	1
SECTION 1: FORM DESIGN.....	3
SECTION 2: PARTICIPATION.....	5
SECTION 3: ITEM ANALYSIS TO FACILITATE EQUATING.....	19
SECTION 4: SCORE RELIABILITY & SUMMARY STATISTICS.....	21
SECTION 5: RATER CONSISTENCY OF ELA EXPRESSING IDEAS PROMPT SCORES.....	22
SECTION 6: CONDITIONAL STANDARD ERROR OF MEASUREMENT AT CUT-POINTS.....	23
SECTION 7: CLASSIFICATION ACCURACY AND CONSISTENCY.....	24
SECTION 8: DIFFERENTIAL ITEM FUNCTIONING.....	28
SECTION 9: INTERRELATIONS AMONG STRANDS WITHIN MEASURES.....	34
SECTION 10: SUMMARY OF ITEMS FLAGGED FOR DIFFICULTY, DISCRIMINATION, AND MODEL FIT.....	42
SECTION 11: VERIFICATION OF PSYCHOMETRIC PROCEDURES.....	44
REFERENCES.....	51
APPENDIX A: ANCHOR ITEM PLOTS.....	53
APPENDIX B: TEST CHARACTERISTIC CURVES AND STANDARD ERROR CURVES.....	63

LIST OF TABLES

TABLE 1.1: OPERATIONAL MATHEMATICS TEST BLUEPRINT GRADES 3 TO 8	3
TABLE 1.2: OPERATIONAL MATHEMATICS TEST BLUEPRINT GRADE 11	3
TABLE 1.3: OPERATIONAL ENGLISH LANGUAGE ARTS TEST BLUEPRINT GRADES 3 TO 11	4
TABLE 1.4: OPERATIONAL SCIENCE TEST BLUEPRINT GRADES 5, 8, AND 11	4
TABLE 2.1: 2010–2011 N-COUNTS AND PERCENTS BY GENDER AND GRADE FOR ELA	5
TABLE 2.2: 2010–2011 N-COUNTS AND PERCENTS BY GENDER AND GRADE FOR MATHEMATICS	6
TABLE 2.3: 2010–2011 N-COUNTS AND PERCENTS BY GENDER AND GRADE FOR SCIENCE	6
TABLE 2.4: 2010–2011 N-COUNTS AND PERCENTS BY ETHNICITY AND GRADE FOR ELA	7
TABLE 2.5: 2010–2011 N-COUNTS AND PERCENTS BY ETHNICITY AND GRADE FOR MATHEMATICS	7
TABLE 2.6: 2010–2011 N-COUNTS AND PERCENTS BY ETHNICITY AND GRADE FOR SCIENCE	8
TABLE 2.7: 2010–2011 N-COUNTS AND PERCENTS BY ECONOMICALLY DISADVANTAGED AND GRADE FOR ELA	8
TABLE 2.8: 2010–2011 N-COUNTS AND PERCENTS BY ECONOMICALLY DISADVANTAGED AND GRADE FOR MATHEMATICS	9
TABLE 2.9: 2010–2011 N-COUNTS AND PERCENTS BY ECONOMICALLY DISADVANTAGED AND GRADE FOR SCIENCE	9
TABLE 2.10: 2010–2011 N-COUNTS AND PERCENTS BY ACCOMMODATION AND GRADE FOR ELA	10
TABLE 2.11: 2010–2011 N-COUNTS AND PERCENTS BY ACCOMMODATION AND GRADE FOR MATHEMATICS	10
TABLE 2.12: 2010–2011 N-COUNTS AND PERCENTS BY ACCOMMODATION AND GRADE FOR SCIENCE	10
TABLE 2.13: 2010–2011 MI-ACCESS FUNCTIONAL INDEPENDENCE—ELA, PERCENT OF STUDENTS BY SUBGROUP AND FORM	11
TABLE 2.14: 2010–2011 MI-ACCESS FUNCTIONAL INDEPENDENCE—MATHEMATICS, PERCENT OF STUDENTS BY SUBGROUP AND FORM	13
TABLE 2.15: 2010–2011 MI-ACCESS FUNCTIONAL INDEPENDENCE—SCIENCE, PERCENT OF STUDENTS BY SUBGROUP AND FORM	14
TABLE 2.16: RAW SCORE SUMMARY STATISTICS BY GENDER*	15
TABLE 2.17: RAW SCORE SUMMARY STATISTICS BY WHITE AND BLACK STUDENTS*	16
TABLE 2.18: RAW SCORE SUMMARY STATISTICS BY ECONOMICALLY DISADVANTAGED*	17
TABLE 2.19: RAW SCORE SUMMARY STATISTICS BY MAJOR ACCOMMODATION GROUP*	18
TABLE 4.1: SCORE RELIABILITY AND SUMMARY STATISTICS BY GRADE	21
TABLE 5.1: INTERRATER AGREEMENT RATES FOR OPERATIONAL EXPRESSING IDEAS PROMPT SCORES	22
TABLE 5.2: INTERRATER AGREEMENT RATES FOR FIELD TEST EXPRESSING IDEAS PROMPT SCORES	22

TABLE 6.1: CONDITIONAL STANDARD ERROR OF MEASUREMENT OF CUT-POINTS BY SUBJECT AND GRADE	23
TABLE 7.1: PERCENT OF STUDENTS BY PROFICIENCY LEVEL.....	24
TABLE 7.2: ESTIMATED CLASSIFICATION ACCURACY AND CONSISTENCY BY SUBJECT AND GRADE	25
TABLE 7.3 NUMBER AND PERCENT OF STUDENTS BY PERFORMANCE LEVEL CHANGE BETWEEN 2009 AND 2010— ACCESSING PRINT	26
TABLE 7.4 NUMBER AND PERCENT OF STUDENTS BY PERFORMANCE LEVEL CHANGE BETWEEN 2009 AND 2010— MATHEMATICS	27
TABLE 8.1: ELA DIF SUMMARY	29
TABLE 8.2: MATHEMATICS DIF SUMMARY	30
TABLE 8.3: SCIENCE DIF SUMMARY.....	31
TABLE 8.4: DIF SUMMARY ACROSS GRADES	32
TABLE 9.1: CORRELATIONS BETWEEN MULTIPLE CHOICE (ACCESSING PRINT) AND CONSTRUCTED RESPONSE (EXPRESSING IDEAS) SCORES.....	35
TABLE 9.2: MEAN ACCESSING PRINT SCORE BY EXPRESSING IDEAS PROMPT SCORE	35
TABLE 9.3: ELA STRAND PEARSON PRODUCT-MOMENT INTERCORRELATIONS BY GRADE	36
TABLE 9.4: MATHEMATICS STRAND PEARSON PRODUCT-MOMENT INTERCORRELATIONS FOR GRADES 3–8	37
TABLE 9.5: MATHEMATICS STRAND PEARSON PRODUCT-MOMENT INTERCORRELATIONS FOR GRADE 11.....	38
TABLE 9.6: SCIENCE STRAND PEARSON PRODUCT-MOMENT INTERCORRELATIONS	38
TABLE 9.7: ELA STRAND SUMMARY STATISTICS.....	39
TABLE 9.8: MATHEMATICS STRAND SUMMARY STATISTICS.....	40
TABLE 9.9: SCIENCE STRAND SUMMARY STATISTICS.....	41
TABLE 10.1: NUMBER OF FLAGGED ITEMS.....	43
TABLE 11.1: EXAMPLE OF WINSTEPS VERIFICATION RECORD PROVIDED TO MDE	45
TABLE 11.2: LINKING CONSTANT (LC) COMPARISON.....	46
TABLE 11.3: HUMRRO'S VERIFICATION TABLE FOR CLASSICAL STATISTICS ON SCIENCE FIELD TEST ITEMS, BY SUBGROUP	47
TABLE 11.4: VERIFICATION OF MATCHES OF WINSTEPS OUTPUT BETWEEN HUMRRO AND QUESTAR FOR MATHEMATICS, ELA, AND SCIENCE	48
TABLE 11.5: VERIFICATION OF THE DIF ANALYSES OUTPUT BETWEEN HUMRRO AND QUESTAR FOR MATHEMATICS, ELA, AND SCIENCE	49

INTRODUCTION

The MI–Access Technical Reports provide information about (a) the nature of the tests; (b) their intended uses; (c) the processes involved in their development; (d) technical information related to scoring, interpretation, and evidence of reliability and validity; (e) scaling and equating; and (f) guidelines for test administration and interpretation, as recommended by the Standards for Educational and Psychological Testing (1999, p. 67). Technical Reports have been developed for the Functional Independence assessments and the Participation/Supported Independence level of assessments.

The following Technical Reports have been developed:

Functional Independence ELA/Mathematics, March 2007

Participation and Supported Independence ELA/Mathematics, June 2007

Participations/Supported Independence/Functional Independence Science, August 2008

Each year, an addendum will be produced to provide the technical quality evidence for the most recent operational administrations of the tests. This is the fifth annual addendum and includes the Functional Independence ELA, Mathematics, and Science tests administered in the 2010 – 2011 school year.

As indicated in the full technical reports for MI–Access, the reports are designed to communicate with multiple users, including state policy makers and their staffs, school and district administrators, teachers, and parents and other advocates interested in such documentation. However, the addendums are designed to provide annual technical quality updates for a much smaller audience. The addendums will focus on reliability and validity evidence gathered at the time of test administration, scoring and equating, and reporting.

1. Form Design

The form design of the 2010 – 2011 operational tests was unchanged from last year’s design. Tables 1.1 to 1.4 contain the test blueprints. In ELA, 6 forms were developed for grades 3 – 8 and four forms for grade 11. In Mathematics, 3 forms were developed for grades 3 – 8 and two forms for grade 11. In Science, two forms were developed at each of the three grades.

Each form also contained a set of anchor items that were used to facilitate equating to the score scale originally developed in 2005 – 2006 for ELA and Mathematics, and in 2007 – 2008 for Science. Anchor items were included among the core items as they counted toward the total score. See Section 3 for the number of anchor items by test and grade.

Table 1.1
Operational Mathematics Test Blueprint Grades 3 to 8

Strand	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Numbers and Operations	10	16	16	18	17	17
Algebra						2
Measurement	8	8	10	12	12	10
Geometry	9	4	2	2	3	3
Data and Probability	3	2	2	3	3	3
Total Core Items	30	30	30	35	35	35
Embedded Field-test Items	8	8	8	10	10	10
Total Test Items	38	38	38	45	45	45

Table 1.2
Operational Mathematics Test Blueprint Grade 11

Strand	Grade 11
Patterns and Relationships	4
Geometry and Measurement	16
Data analysis and Statistics	2
Number Sense and Numeration	15
Numerical and Algebraic Operations	3
Total Core Items	40
Embedded Field-test Items	10
Total Test Items	50

Table 1.3
Operational English Language Arts Test Blueprint Grades 3 to 11

Strand	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
Word Recognition	20	20	20	20	20	20	20
Text Comprehension							
Narrative Text	7	7	7	7	7	7	7
Informational Text	7	7	7	7	7	7	7
Functional Text	7	7	7	7	7	7	7
Accessing Print Total	41	41	41	41	41	41	41
Expressing Ideas Prompt	0	1	0	0	1	0	1
Total Core Items	41	42	41	41	42	41	42
Embedded Field-test Items	11	12	11	11	12	11	12
Total Test Items	52	54	52	52	54	52	54

Table 1.4
Operational Science Test Blueprint Grades 5, 8, and 11

Strand	Grade 5	Grade 8	Grade 11
Constructing & Reflecting	4	4	4
Life Science	13	14	14
Physical Science	12	14	15
Earth Science	6	8	12
Total Core Items	35	40	45
Embedded Field-test Items	8	10	10
Total Test Items	43	50	55

2. Participation

Participation in the assessments is monitored by gender, by racial/ethnic group, by economically disadvantaged, and by accommodation. These student-level characteristics are also used to evaluate differential item functioning (DIF) when the groups are large enough to support the analysis. These results are reported in Section 8.

Participation counts and percentages by gender and grade are given in Tables 2.1 – 2.3 for ELA, Mathematics, and Science, respectively, participation counts and percentages by race/ethnicity and grade are given in Tables 2.4 – 2.6, participation counts and percentages by economically disadvantaged and grade are given in Tables 2.7 – 2.9, and participation counts and percentages by accommodation and grade are given in Tables 2.10 – 2.12. In general, there are roughly twice as many males as females. The largest racial/ethnic group is White students with 62% to 66% of the students, followed by Black students with 23% to 29% of the students, Hispanic students with 3% to 7% of the students, and Asian with about 1% of the students. In general, there are approximately two to three times as many economically disadvantaged students as non-economically disadvantaged students. The largest accommodation is for Reader (e.g., Read aloud accommodation) with 28% to 62% of the students. The highest percentage is at grade 3 and it becomes progressively smaller as grade level increases. The next largest percentage is for Audio CD with 8% to 15% of the students, followed by Other with 2% to 9% of the students. For ELA at grades 4, 7, and 11, where students are administered a prompt, 17%, 10%, and 6% of the students, respectively, have a scribe as an accommodation.

Table 2.1
2010–2011 N-Counts and Percents by Gender and Grade for ELA

Grade	Female		Male		Total
	N	%	N	%	N
3	591	32.5	1229	67.5	1820
4	707	32.2	1489	67.8	2196
5	728	33.4	1454	66.6	2182
6	761	35.0	1415	65.0	2176
7	782	35.8	1402	64.2	2184
8	752	34.8	1411	65.2	2163
11	676	37.8	1112	62.2	1788

Table 2.2
2010–2011 N-Counts and Percents by Gender and Grade for Mathematics

Grade	Female		Male		Total
	N	%	N	%	N
3	548	34.8	1025	65.2	1573
4	662	35.1	1225	64.9	1887
5	700	35.3	1284	64.7	1984
6	760	37.0	1296	63.0	2056
7	784	36.7	1355	63.3	2139
8	764	36.7	1318	63.3	2082
11	676	37.7	1118	62.3	1794

Table 2.3
2010–2011 N-Counts and Percents by Gender and Grade for Science

Grade	Female		Male		Total
	N	%	N	%	N
5	628	34.6	1188	65.4	1816
8	709	36.2	1247	63.8	1956
11	675	37.8	1110	62.2	1785

Table 2.4
2010–2011 N-Counts and Percents by Ethnicity and Grade for ELA

Grade	Native Hawaiian or Other Pacific Islander		American Indian or Alaskan Native		Black or African American		Hispanic of any race		White		Multi-racial		Asian		Unknown		Total
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
3	3	0.2	21	1.2	414	22.7	135	7.4	1194	65.6	41	2.3	12	0.7	0	0.0	1820
4	4	0.2	24	1.1	559	25.5	142	6.5	1400	63.8	43	2.0	24	1.1	0	0.0	2196
5	1	0.0	35	1.6	526	24.1	131	6.0	1406	64.4	52	2.4	31	1.4	0	0.0	2182
6	2	0.1	24	1.1	589	27.1	113	5.2	1396	64.2	34	1.6	18	0.8	0	0.0	2176
7	0	0.0	19	0.9	625	28.6	114	5.2	1359	62.2	50	2.3	17	0.8	0	0.0	2184
8	2	0.1	26	1.2	615	28.4	108	5.0	1367	63.2	26	1.2	19	0.9	0	0.0	2163
11	2	0.1	24	1.3	476	26.6	55	3.1	1186	66.3	19	1.1	26	1.5	0	0.0	1788

Table 2.5
2010–2011 N-Counts and Percents by Ethnicity and Grade for Mathematics

Grade	Native Hawaiian or Other Pacific Islander		American Indian or Alaskan Native		Black or African American		Hispanic of any race		White		Multi-racial		Asian		Unknown		Total
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
3	3	0.2	16	1.0	388	24.7	117	7.4	1001	63.6	37	2.4	11	0.7	0	0.0	1573
4	3	0.2	21	1.1	502	26.6	115	6.1	1187	62.9	37	2.0	22	1.2	0	0.0	1887
5	1	0.1	29	1.5	498	25.1	116	5.8	1264	63.7	51	2.6	25	1.3	0	0.0	1984
6	2	0.1	24	1.2	580	28.2	104	5.1	1296	63.0	35	1.7	15	0.7	0	0.0	2056
7	0	0.0	21	1.0	617	28.8	112	5.2	1322	61.8	52	2.4	15	0.7	0	0.0	2139
8	2	0.1	22	1.1	597	28.7	103	4.9	1313	63.1	25	1.2	20	1.0	0	0.0	2082
11	2	0.1	24	1.3	477	26.6	57	3.2	1189	66.3	19	1.1	26	1.4	0	0.0	1794

Table 2.6
2010–2011 N-Counts and Percents by Ethnicity and Grade for Science

Grade	Native Hawaiian or Other Pacific Islander		American Indian or Alaskan Native		Black or African American		Hispanic of any race		White		Multi-racial		Asian		Unknown		Total
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
5	1	0.1	30	1.7	478	26.3	113	6.2	1123	61.8	45	2.5	26	1.4	0	0.0	1816
8	2	0.1	21	1.1	569	29.1	97	5.0	1223	62.5	25	1.3	19	1.0	0	0.0	1956
11	2	0.1	24	1.3	472	26.4	57	3.2	1185	66.4	19	1.1	26	1.5	0	0.0	1785

Table 2.7
2010-2011 N-Counts and Percents by Economically Disadvantaged and Grade for ELA

Grade	Economically Disadvantaged		Non-Economically Disadvantaged		Total
	N	%	N	%	
3	1329	73.0	491	27.0	1820
4	1615	73.5	581	26.5	2196
5	1593	73.0	589	27.0	2182
6	1560	71.7	616	28.3	2176
7	1563	71.6	621	28.4	2184
8	1465	67.7	698	32.3	2163
11	1151	64.4	637	35.6	1788

Table 2.8

2010-2011 N-Counts and Percents by Economically Disadvantaged and Grade for Mathematics

Grade	Economically Disadvantaged		Non-Economically Disadvantaged		Total
	N	%	N	%	
3	1150	73.1	423	26.9	1573
4	1391	73.7	496	26.3	1887
5	1427	71.9	557	28.1	1984
6	1496	72.8	560	27.2	2056
7	1523	71.2	616	28.8	2139
8	1399	67.2	683	32.8	2082
11	1154	64.3	640	35.7	1794

Table 2.9

2010-2011 N-Counts and Percents by Economically Disadvantaged and Grade for Science

Grade	Economically Disadvantaged		Non-Economically Disadvantaged		Total
	N	%	N	%	
5	1324	72.9	492	27.1	1816
8	1315	67.2	641	32.8	1956
11	1146	64.2	639	35.8	1785

Table 2.10
2010–2011 N-Counts and Percents by Accommodation and Grade for ELA

Grade	Reader		Audio		Braille		Enlarged Print		Other		Scribe		Word Processed		All Students
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
3	1038	57.0	149	8.2	2	0.1	7	0.4	152	8.4	NA		0	0.0	1820
4	1194	54.4	205	9.3	3	0.1	6	0.3	198	9.0	361	7.1	7	0.3	2196
5	1110	50.9	252	11.5	3	0.1	10	0.5	160	7.3	NA		0	0.0	2182
6	964	44.3	267	12.3	0	0.0	1	0.0	131	6.0	NA		0	0.0	2176
7	948	43.4	277	12.7	6	0.3	4	0.2	139	6.4	207	9.8	34	1.6	2184
8	846	39.1	269	12.4	2	0.1	9	0.4	106	4.9	NA		0	0.0	2163
11	553	30.9	155	8.7	0	0.0	9	0.5	29	1.6	101	5.6	16	0.9	1788

Table 2.11
2010–2011 N-Counts and Percents by Accommodation and Grade for Mathematics

Grade	Reader		Audio		Braille		Enlarged Print		Other		Scribe		Word Processed		All Students
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
3	971	61.7	153	9.7	2	0.1	5	0.3	109	6.9	NA		0	0.0	1573
4	1111	58.9	216	11.4	3	0.2	6	0.3	121	6.4	NA		0	0.0	1887
5	1124	56.7	265	13.4	2	0.1	9	0.5	127	6.4	NA		0	0.0	1984
6	971	47.2	279	13.6	0	0.0	2	0.1	125	6.1	NA		0	0.0	2056
7	936	43.8	275	12.9	5	0.2	4	0.2	112	5.2	NA		0	0.0	2139
8	767	36.8	301	14.5	2	0.1	7	0.3	107	5.1	NA		0	0.0	2082
11	495	27.6	146	8.1	0	0.0	9	0.5	60	3.3	NA		0	0.0	1794

Table 2.12
2010–2011 N-Counts and Percents by Accommodation and Grade for Science

Grade	Reader		Audio		Braille		Enlarged Print		Other		Scribe		Word Processed		All Students
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
5	1091	60.1	230	12.7	1	0.1	8	0.4	105	5.8	NA		0	0.0	1816
8	812	41.5	283	14.5	2	0.1	6	0.3	85	4.3	NA		0	0.0	1956
11	532	29.8	147	8.2	0	0.0	10	0.6	29	1.6	NA		0	0.0	1785

Form Distribution

Recall from Section 1, six forms were developed for ELA at grades 3 – 8 and four at grade 11, three forms were developed for Mathematics at grades 3 – 8 and two at grade 11, and two forms were developed for Science at each of the three grades. These forms were distributed to districts and schools according to the guidelines from the Michigan Department of Education, Bureau of Assessment and Accountability. The sampling unit was the school. Forms were randomly assigned using stratified random sampling where stratification was based on the enrollment counts provided to Questar. Except for Detroit, each district received up to two forms at a grade. For Detroit, each school received the same form. One additional condition was imposed on the distribution of forms. Due to cost considerations, only Form 1 was developed for the audio, Braille, and enlarged print accommodations. Hence, all students with these accommodations were administered Form 1.

The percent of students by various subgroups and form for the 2010 – 2011 school year are given in the tables below. Each table contains the number of students tested by form at each grade, as well as the grade total. At each grade, the percent of students for the various subgroups is given by form as well as for the grade total. The percents for ELA, Mathematics, and Science are given in Tables 2.10, 2.11, and 2.12, respectively. The subgroups consist of gender, three racial/ethnic groups (Black, Hispanic, and White), and three other subgroups (Economically Disadvantaged, English Language Learners or ELL, and Formerly Limited English Proficient or FLEP). As seen from the tables for all three content areas, each form is well represented by the various subgroups. Moreover, for each form at a grade, the percent of students across the subgroups is generally consistent with the percents for the grade population.

Table 2.13

2010–2011 MI–Access Functional Independence—ELA Percent of Students by Subgroup and Form

	N	Female	Male	Black	Hispanic	White	Economic Disadv	ELL	FLEP
Grade 3									
All Forms	1820	32.5	67.5	22.7	7.4	65.6	73.0	5.0	0.6
Form 1*	540	29.8	70.2	23.5	7.2	64.1	74.3	3.9	1.1
Form 2	265	36.6	63.4	31.3	9.1	56.2	70.6	4.5	0.4
Form 3	239	35.1	64.9	20.1	8.8	66.9	75.7	4.2	0.0
Form 4	247	34.0	66.0	27.1	6.1	62.8	74.5	8.1	0.4
Form 5	229	28.4	71.6	23.6	4.8	68.1	71.6	7.0	0.9
Form 6	300	33.3	66.7	11.7	8.3	76.0	70.7	4.0	0.3
Grade 4									
All Forms	2196	32.2	67.8	25.5	6.5	63.8	73.5	4.1	0.4
Form 1*	664	32.5	67.5	26.5	7.4	62.3	74.7	6.0	0.6
Form 2	320	32.5	67.5	29.7	5.6	61.3	75.6	4.1	0.3
Form 3	269	32.3	67.7	31.2	5.6	58.7	71.7	3.3	0.0
Form 4	309	30.1	69.9	19.4	7.4	68.3	75.4	4.9	0.3
Form 5	299	30.1	69.9	23.4	6.0	65.9	70.2	2.3	0.3
Form 6	335	34.9	65.1	22.1	5.7	66.9	71.9	1.8	0.3
Grade 5									
All Forms	2182	33.4	66.6	24.1	6.0	64.4	73.0	3.3	0.7
Form 1*	644	32.0	68.0	20.0	4.3	69.9	73.1	2.6	0.6
Form 2	301	33.6	66.4	31.9	5.0	57.1	75.7	2.3	0.7

Table 2.13 (cont.)
2010–2011 MI-Access Functional Independence—ELA Percent of Students by Subgroup and Form

	N	Female	Male	Black	Hispanic	White	Economic Disadv	ELL	FLEP
Form 3	285	35.8	64.2	23.9	8.8	63.5	74.0	1.8	0.0
Form 4	280	34.6	65.4	22.9	7.9	65.0	67.9	5.0	0.4
Form 5	347	30.3	69.7	25.4	7.2	62.5	75.5	3.7	0.3
Form 6	325	36.0	64.0	24.9	4.9	62.8	71.1	4.9	2.2
Grade 6									
All Forms	2176	35.0	65.0	27.1	5.2	64.2	71.7	2.6	0.2
Form 1*	691	32.9	67.1	27.8	6.1	62.2	73.5	3.6	0.0
Form 2	320	33.8	66.3	37.5	5.3	53.4	78.4	0.9	0.0
Form 3	286	37.4	62.6	21.3	4.2	72.0	66.8	4.2	1.0
Form 4	313	34.8	65.2	24.6	4.5	66.5	68.4	1.9	0.0
Form 5	291	39.2	60.8	29.2	3.8	64.6	74.6	2.7	0.0
Form 6	275	34.9	65.1	19.6	6.2	70.2	65.1	1.1	0.4
Grade 7									
All Forms	2184	35.8	64.2	28.6	5.2	62.2	71.6	2.7	0.1
Form 1*	728	36.7	63.3	25.8	4.7	65.1	67.4	2.5	0.3
Form 2	313	35.8	64.2	29.1	8.9	59.1	71.6	5.1	0.0
Form 3	276	37.0	63.0	20.7	3.3	72.5	72.8	1.1	0.0
Form 4	268	34.7	65.3	31.3	4.9	60.4	79.1	3.0	0.0
Form 5	282	35.8	64.2	30.5	5.3	59.2	72.0	4.6	0.0
Form 6	317	33.8	66.2	37.5	4.7	53.9	73.2	0.6	0.0
Grade 8									
All Forms	2163	34.8	65.2	28.4	5.0	63.2	67.7	2.6	0.5
Form 1*	714	34.7	65.3	28.4	5.6	62.6	66.4	2.7	0.6
Form 2	304	33.6	66.4	31.6	5.9	58.9	69.7	4.6	0.3
Form 3	292	31.8	68.2	27.4	4.1	65.4	73.6	1.4	0.0
Form 4	297	35.0	65.0	28.6	4.4	62.6	67.0	2.4	0.3
Form 5	249	34.9	65.1	25.3	3.6	68.7	67.1	0.8	0.8
Form 6	307	38.4	61.6	28.7	5.2	62.9	64.5	3.6	0.7
Grade 11									
All Forms	1788	37.8	62.2	26.6	3.1	66.3	64.4	1.4	0.1
Form 1*	663	39.5	60.5	26.5	3.6	66.7	63.2	0.9	0.0
Form 2	407	36.4	63.6	27.5	2.5	65.4	64.4	1.7	0.0
Form 3	368	38.0	62.0	29.3	2.4	64.1	69.6	1.9	0.0
Form 4	350	36.0	64.0	22.9	3.4	69.1	61.1	1.4	0.3

*Form 1 is administered to all students who require the audio, Braille, or enlarged print accommodation. N-counts by accommodation are given in Table 2.7.

Table 2.14
2010–2011 MI–Access Functional Independence—Mathematics
Percent of Students by Subgroup and Form

	N	Female	Male	Black	Hispanic	White	Economic Disadv	ELL	FLEP
Grade 3									
All Forms	1573	34.8	65.2	24.7	7.4	63.6	73.1	5.0	0.7
Form 1*	668	35.5	64.5	26.0	6.3	63.5	73.7	3.3	0.3
Form 2	487	34.3	65.7	25.7	8.4	61.2	73.5	5.5	0.6
Form 3	418	34.4	65.6	21.3	8.1	66.7	71.8	7.2	1.4
Grade 4									
All Forms	1887	35.1	64.9	26.6	6.1	62.9	73.7	4.0	0.4
Form 1*	815	34.1	65.9	26.5	6.3	64.2	73.3	3.8	0.4
Form 2	561	34.8	65.2	26.6	5.9	62.7	71.8	3.7	0.5
Form 3	511	37.0	63.0	26.8	6.1	61.1	76.5	4.5	0.4
Grade 5									
All Forms	1984	35.3	64.7	25.1	5.8	63.7	71.9	3.1	0.7
Form 1*	943	35.5	64.5	24.5	6.4	64.2	73.0	2.8	0.4
Form 2	531	35.4	64.6	26.7	5.3	62.3	71.0	3.4	0.2
Form 3	510	34.7	65.3	24.5	5.5	64.3	71.0	3.5	1.8
Grade 6									
All Forms	2056	37.0	63.0	28.2	5.1	63.0	72.8	2.3	0.0
Form 1*	924	37.9	62.1	30.1	6.0	60.2	75.1	2.6	0.0
Form 2	553	35.8	64.2	26.4	5.1	64.9	70.5	2.9	0.2
Form 3	579	36.6	63.4	26.9	3.6	65.8	71.2	1.4	0.0
Grade 7									
All Form	2139	36.7	63.3	28.8	5.2	61.8	71.2	2.5	0.1
Form 1*	996	37.8	62.2	25.6	5.2	64.2	69.8	3.0	0.2
Form 2	533	35.3	64.7	31.7	5.3	59.1	71.3	3.0	0.2
Form 3	610	36.1	63.9	31.6	5.2	60.3	73.4	1.3	0.0
Grade 8									
All Forms	2082	36.7	63.3	28.7	4.9	63.1	67.2	2.6	0.4
Form 1*	1016	36.3	63.7	27.3	5.3	64.5	67.7	2.2	0.5
Form 2	535	36.1	63.9	32.5	5.0	58.7	66.7	3.6	0.2
Form 3	531	38.0	62.0	27.5	4.1	64.8	66.7	2.6	0.6
Grade 11									
All Forms	1794	37.7	62.3	26.6	3.2	66.3	64.3	1.4	0.1
Form 1*	1039	39.7	60.3	26.3	3.8	65.8	64.8	1.6	0.1
Form 2	755	35.0	65.0	27.0	2.4	66.9	63.7	1.1	0.0

*Form 1 is administered to all students who require the audio, Braille, or enlarged print accommodation. N-counts by accommodation are given in Table 2.8.

Table 2.15

2010–2011 MI–Access Functional Independence—Science Percent of Students by Subgroup and Form

	N	Female	Male	Black	Hispanic	White	Economic Disadv	ELL	FLEP
Grade 5									
All Forms	1816	34.6	65.4	26.3	6.2	61.8	72.9	3.5	0.8
Form 1*	1089	34.8	65.2	24.5	6.5	63.5	72.8	3.5	0.8
Form 2	727	34.3	65.7	29.0	5.8	59.3	73.0	3.4	0.7
Grade 8									
All Forms	1956	36.2	63.8	29.1	5.0	62.5	67.2	2.5	0.3
Form 1*	1157	36.6	63.4	26.6	5.5	64.2	65.9	2.9	0.4
Form 2	799	35.7	64.3	32.7	4.1	60.1	69.2	1.8	0.1
Grade 11									
All Forms	1785	37.8	62.2	26.4	3.2	66.4	64.2	1.4	0.1
Form 1*	1032	38.0	62.0	27.8	3.9	64.1	65.5	1.6	0.1
Form 2	753	37.6	62.4	24.6	2.3	69.5	62.4	1.2	0.0

*Form 1 is administered to all students who require the audio, Braille, or enlarged print accommodation. N-counts by accommodation are given in Table 2.9.

Subgroup Analysis

The raw score mean and standard deviation by content area and grade for the major subgroups are given in the following tables; Table 2.16 for males and females, Table 2.17 for White and Black students, Table 2.18 for economically disadvantaged and non-economically disadvantaged students, and Table 2.19 for accommodated groups. The males scored higher than the females at all grades of Mathematics and Science and at grades 4 and 5 of ELA. The White students scored higher than the Black students at all grades and content areas. The economically disadvantaged students scored higher than the non-economically disadvantaged students at grades 4 – 7 Mathematics, grade 5 Science, and grade 6 ELA, and the two groups scored about the same at grade 7 ELA and grade 8 Mathematics.

Table 2.19 contains the raw score mean and standard deviation for the largest accommodated groups and for those students who had no accommodation. At all three grades where the ELA prompt is given, those students receiving the scribe accommodation scored lower than the non-accommodated group. Also for ELA, those students receiving the audio accommodation scored higher than the non-accommodated group except at grade 11, those students receiving the Reader accommodation scored higher at grades 3 – 6, and those students receiving the Other accommodation scored higher at grades 3 – 5. For Mathematics the results were completely mixed between the accommodated and non-accommodated groups; at some grades the accommodated group scored higher, at other grades the non-accommodated group scored higher, and at some grades the groups scored about the same. For Science, those students receiving the Reader or Audio accommodation scored higher than the non-accommodated group. Those students receiving the Other accommodation scored higher at grade 5, the same at grade 8, and lower at grade 11.

Table 2.16
Raw Score Summary Statistics by Gender*

	Male		Female	
Grade	Mean	SD	Mean	SD
ELA				
3	29.59	6.88	29.58	6.57
4	26.79	7.87	26.58	7.50
5	29.45	7.73	29.17	7.50
6	28.93	7.98	29.07	7.63
7	30.40	7.67	31.42	6.97
8	31.75	7.63	32.47	6.93
11	32.72	7.67	33.84	6.93
Mathematics				
3	21.91	5.13	21.32	4.81
4	22.82	5.55	21.89	5.64
5	21.10	5.52	19.92	5.65
6	25.40	5.92	23.83	5.92
7	22.49	6.08	21.89	5.80
8	22.21	5.99	20.89	5.60
11	25.97	7.51	24.57	7.49
Science				
5	21.51	5.31	20.70	5.03
8	23.84	6.45	22.96	5.62
11	23.90	7.02	23.33	6.27

*N-counts are given in Tables 2.1 – 2.3.

Table 2.17
Raw Score Summary Statistics by White and Black Students*

	White		Black	
Grade	Mean	SD	Mean	SD
ELA				
3	30.10	6.64	28.04	7.07
4	27.29	7.67	25.35	7.91
5	29.84	7.63	28.23	7.75
6	29.41	7.90	27.78	7.82
7	31.39	7.23	29.32	7.88
8	32.64	7.20	30.72	7.68
11	34.26	6.89	30.62	8.01
Mathematics				
3	22.11	4.91	20.45	5.31
4	22.84	5.50	21.64	5.82
5	21.03	5.58	19.61	5.72
6	25.38	5.87	23.54	6.05
7	23.03	5.77	20.67	6.11
8	22.27	5.78	20.55	6.04
11	26.89	7.30	21.89	6.94
Science				
5	21.72	5.19	20.09	5.23
8	24.51	6.17	21.59	5.86
11	25.35	6.53	19.79	5.60

*N-counts are given in Tables 2.4 – 2.6.

Table 2.18
Raw Score Summary Statistics by Economically Disadvantaged*

Grade	Economically Disadvantaged		Non-Economically Disadvantaged	
	Mean	SD	Mean	SD
ELA				
3	29.37	6.78	30.17	6.74
4	26.45	7.85	27.47	7.43
5	29.41	7.56	29.22	7.91
6	29.21	7.73	28.39	8.15
7	30.75	7.43	30.80	7.48
8	31.91	7.46	32.19	7.28
11	32.64	7.66	34.04	6.87
Mathematics				
3	21.63	5.02	21.89	5.06
4	22.53	5.56	22.39	5.69
5	20.96	5.50	19.96	5.76
6	25.06	5.86	24.16	6.18
7	22.35	5.92	22.07	6.15
8	21.73	5.79	21.72	6.07
11	25.07	7.45	26.11	7.64
Science				
5	21.43	5.11	20.70	5.49
8	23.46	6.08	23.63	6.37
11	23.13	6.69	24.68	6.74

*N-counts are given in Tables 2.7 – 2.9.

Table 2.19
Raw Score Summary Statistics by Major Accommodation Group*

Grade	Reader		Audio		Other		Scribe		Non-Accommodated	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
ELA										
3	30.24	6.47	29.91	6.40	29.56	7.39	NA		28.69	7.13
4	27.14	7.37	28.44	6.94	26.12	8.05	25.29	7.67	25.72	8.29
5	29.77	7.20	31.17	6.67	29.35	8.30	NA		28.20	8.29
6	29.32	7.53	29.95	7.45	28.46	7.87	NA		28.44	8.21
7	30.54	7.11	31.90	7.01	30.45	7.35	29.42	7.60	30.68	7.87
8	31.76	7.13	32.74	7.19	31.90	8.41	NA		31.98	7.65
11	33.08	6.92	32.92	7.10	30.48	9.45	28.99	7.71	33.40	7.60
Mathematics										
3	21.73	4.93	21.39	5.18	22.46	5.45	NA		21.73	5.18
4	22.48	5.46	22.76	5.64	23.67	4.99	NA		22.29	5.92
5	20.77	5.42	20.51	5.65	21.03	5.11	NA		20.49	5.86
6	24.46	6.10	24.71	6.01	25.44	5.97	NA		25.27	5.72
7	22.05	5.73	22.64	5.96	21.46	5.89	NA		22.39	6.25
8	21.39	5.64	22.02	5.88	22.09	6.11	NA		21.83	6.03
11	25.26	7.46	25.95	6.75	22.72	7.60	NA		25.61	7.61
Science										
5	21.46	5.12	21.75	4.80	20.89	5.42	NA		20.47	5.54
8	23.67	6.00	24.07	5.80	23.31	6.67	NA		23.25	6.43
11	24.23	6.41	24.57	6.16	20.41	4.84	NA		23.38	6.97

*N-counts are given in Tables 2.10 – 2.12.

3. Item Analysis to Facilitate Equating

New secure forms must continually be constructed for future test administrations. The test forms are equated so as to convert the raw scores obtained from two forms of the test so that the scale scores derived from the two forms after conversion will be directly equivalent. Different forms of the test are designed to have comparable item content and similar distributions of item statistics based on field testing. The equating adjusts for unintended differences in difficulty of the forms. The equating also adjusts raw test scores from different forms to a common scale so that identical scale scores earned this year and last year reflect the same level of student achievement, even though the corresponding raw scores may differ.

Equating of the MI-Access Functional Independence ELA, Mathematics, and Science assessments was done using a common item or anchor test design. The description of equating is based on the Fall 2009 and Fall 2010 forms for grades 3 – 8 and the Spring 2010 and Spring 2011 forms for grade 11, but applies to all future forms. Anchor items are the same, identical items that appeared in both the 2009 school year form and in the 2010 school year form. For each assessment at each grade, at least 20% of the items were in common between the two forms. The anchor items were used to develop a linking constant that places the Rasch item difficulties from the 2010 school year form on the same logit scale as the 2009 school year form. The linking constant is computed as the difference between the average Rasch difficulty for the anchor items from the 2009 school year form's Winsteps analysis, minus the average Rasch difficulty from the 2010 school year form's Winsteps analysis. For all three content areas, linking constants are computed at each grade the assessment is administered (grades 3 – 8 and 11 for ELA and Mathematics and grades 5, 8, and 11 for Science).

Adding this linking constant to the Rasch difficulties for each of the items in the 2010 school year form places all of the 2010 school year form's Rasch difficulties (and log ability estimates) on the same Rasch logit scale as the 2009 school year form. Then previous years' linking constants are added to the current year's linking constant to place the 2010 school year form's Rasch log ability scale on the original 2005 scale. Recall that scale scores were developed for each assessment at each grade in the first year by setting the attained cut score to a pre-specified value and the standard deviation to 25. The same linear transformation that was developed in the first year for each assessment at each grade was then applied to the equated Rasch log ability scale for the 2010 school year form to yield equated scale scores.

Since equating involves comparing the Rasch difficulties for the anchor items from the 2010 school year form with those from the 2009 school year form, a plot of those difficulties provides information about the quality of the equating. The plot of the 2010 school year Rasch difficulties versus the 2009 school year Rasch difficulties for the anchor items for each assessment at each grade is given in Appendix A. For Accessing Print, there are 11 anchor or equating items at each grade. For Mathematics, there are 8 equating items at each of grades 3 to 8, and 10 equating items at grade 11. For Science, there are 8 equating items at grade 5, and 10 equating items at grades 8 and 11. Each plot also contains the 45-degree straight line that passes through the mean of the 2010 school year Rasch difficulties and the mean of the 2009 school year Rasch difficulties. The plots show that the Rasch difficulties fall along this 45-degree line as the model requires. Of course, not all points are on or right next to the line due to the inherent error that is in all measurement. Across the 17 assessments, all of the plotted points were close to the identity line. In addition, none of the items had a displacement value greater than the criterion of .5 logits given in the Winsteps manual for potentially deleting an item as an anchor item (Linacre, 2006).

Another way to evaluate the plots is to compute the correlation coefficient between the 2010 Rasch difficulties and the 2009 Rasch difficulties. The correlation coefficient (r) is given in the upper right-hand corner of each plot. Across all seventeen 2010 assessments, the correlations ranged from .907 to .997 with a median correlation of .978. These correlations are as close to 1 as can practically be expected.

Equating involved only the core, operational items on each content area and grade level test. Following the equating, the field-test items for each test were calibrated using a concurrent anchor test design. For each test, the core items plus the field test items across all forms were calibrated together in a single Winsteps run by fixing or anchoring the core items to the Rasch values obtained during equating. This single run placed all field test items on the same scale as the core operational items.

The Test Characteristic Curve (TCC) and Standard Error Curve (SEC) for each assessment at each grade are given in Appendix B. The raw score cuts are denoted in each plot of the TCC and the scale scores associated with the raw score cuts are denoted in each plot of the SEC with red square symbols.

4. Score Reliability & Summary Statistics

Score reliability is estimated by Cronbach's Coefficient Alpha using item raw score data in SPSS and by the model reliability estimated by the Rasch modeling in Winsteps version 3.67.0 (Linacre, 2006). Raw score and scale score summary statistics are also presented in Table 4.1 for all assessments. Across the grades, the traditional Cronbach's Coefficient Alpha reliability estimates ranged from .86 – .91 for ELA, .82 – .87 for mathematics, and .75 – .78 for science. While the science reliability estimates are lower than those for ELA and mathematics, they still indicate a reasonably high degree of internal consistency. The model based reliability estimates are somewhat similar to the Cronbach's Coefficient Alpha estimates. The biggest differences between the estimates are for the ELA tests in grades 5 – 8 and 11 and the mathematics tests in grades 3 – 5.

Table 4.1
Score Reliability and Summary Statistics by Grade

	Grade Level						
ELA	3	4	5	6	7	8	11
Model Reliability	0.82	0.85	0.84	0.84	0.81	0.79	0.77
Cronbach's Alpha	0.86	0.87	0.89	0.89	0.89	0.90	0.91
Raw Score Mean	29.6	26.7	29.4	29.0	30.8	32.0	33.1
Raw Score SD	6.8	7.8	7.7	7.9	7.4	7.4	7.4
Raw Score Max	41	45	41	41	45	41	45
Raw Score Min	0	0	0	0	0	0	0
Scale Score Mean	2317	2416	2521	2627	2730	2833	3134
Scale Score SD	21.6	22.1	24.6	24.8	23.7	24.9	28.2
	Grade Level						
Mathematics	3	4	5	6	7	8	11
Model Reliability	0.77	0.75	0.79	0.80	0.81	0.81	0.86
Cronbach's Alpha	0.82	0.86	0.84	0.84	0.82	0.81	0.87
Raw Score Mean	21.7	22.5	20.7	24.8	22.3	21.7	25.4
Raw Score SD	5.0	5.6	5.6	6.0	6.0	5.9	7.5
Raw Score Max	30	30	30	35	35	35	40
Raw Score Min	0	0	0	0	0	0	0
Scale Score Mean	2314	2423	2512	2617	2712	2816	3112
Scale Score SD	22.1	22.9	22.8	20.8	19.6	19.2	24.2
	Grade Level						
Science			5			8	11
Model Reliability			0.75			0.79	0.79
Cronbach's Alpha			0.75			0.78	0.78
Raw Score Mean			21.2			23.5	23.7
Raw Score SD			5.2			6.2	6.7
Raw Score Max			35			40	45
Raw Score Min			0			0	0
Scale Score Mean			2502			2799	3106
Scale Score SD			24.0			28.3	22.8

5. Rater Consistency of ELA Expressing Ideas Prompt Scores

The writing prompt responses are scored by human raters. Of the core EI prompts, 18%, 21%, and 20% were scored by two raters at grades 4, 7, and 11, respectively, with 92%, 92%, and 86% perfect agreement, respectively. Also, 17% to 20% of the field-test EI prompts were double-scored with 86% or 87% perfect agreement. As seen in Tables 5.1 and 5.2, differences of more than one point occur infrequently.

Table 5.1
Interrater Agreement Rates for Operational Expressing Ideas Prompt Scores

Grade	Perfect		1 Point Difference		More than 1 Point Difference	
	N	%	N	%	N	%
4	365	91.9	32	8.1	0	0.0
7	419	91.7	38	8.3	0	0.0
11	308	85.6	50	13.9	2	0.6

Table 5.2
Interrater Agreement Rates for Field Test Expressing Ideas Prompt Scores

Grade	Perfect		1 Point Difference		More than 1 Point Difference	
	N	%	N	%	N	%
4	375	87.0	51	11.8	5	1.2
7	325	85.5	52	13.7	3	0.8
11	311	87.1	45	12.6	1	0.3

6. Conditional Standard Error of Measurement at Cut-Points

The conditional standard error of measurement is estimated in the raw-score to scale-score conversion table after equating. These estimates are based on the ratio of raw-score and scale-score standard deviations to scale the conditional SEM associated with each theta as estimated by the Rasch model in Winsteps. See Appendix B for the plot of all conditional standard errors for each assessment, the standard error curve. The scale score cuts denoted with a red square in each assessment's standard error curve are at the first raw score with a scale score equal to or greater than the scale score cuts given in Table 6.1.

Table 6.1
Conditional Standard Error of Measurement of Cut-Points by Subject and Grade

Grade	Attained		Surpassed	
	Scale Score	Conditional SEM	Scale Score	Conditional SEM
ELA				
3	2300	7	2315	7
4	2400	6	2415	7
5	2500	6	2511	7
6	2600	6	2614	6
7	2700	6	2713	6
8	2800	6	2820	6
11	3100	6	3129	8
Mathematics				
3	2300	8	2314	9
4	2400	7	2417	8
5	2500	8	2515	9
6	2600	7	2617	8
7	2700	7	2714	8
8	2800	7	2817	7
11	3100	7	3135	10
Science				
5	2500	11	2517	12
8	2800	12	2816	13
11	3100	10	3122	10

7. Classification Accuracy and Consistency

The percent of students at each of the three performance levels by content area and grade are given in Table 7.1. For ELA, the largest percent is for Surpassed with 50% to 77%. For Mathematics, excluding grade 11, the largest group is Surpassed with 43% to 57%; there are only 16% Surpassed at grade 11. For Science, the largest group is Emerging with 43%, 55%, and 41% at grades 5, 8, and 11, respectively. The sum of the attained and surpassed percentages in Table 7.1 represents the percentages of students that would be considered proficient for Adequate Year Progress calculations under NCLB.

Table 7.1
Percent of Students by Proficiency Level

Grade	Emerging		Attained		Surpassed	
	N	%	N	%	N	%
ELA						
3	432	23.7	457	25.1	931	51.2
4	532	24.2	575	26.2	1089	49.6
5	448	20.5	337	15.4	1397	64.0
6	283	13.0	411	18.9	1482	68.1
7	178	8.2	330	15.1	1676	76.7
8	181	8.4	487	22.5	1495	69.1
11	216	12.1	562	31.4	1010	56.5
Math						
3	398	25.3	395	25.1	780	49.6
4	322	17.1	483	25.6	1082	57.3
5	564	28.4	578	29.1	842	42.4
6	404	19.6	744	36.2	908	44.2
7	621	29.0	596	27.9	921	43.1
8	438	21.0	660	31.7	984	47.3
11	562	31.3	938	52.3	294	16.4
Science						
5	784	43.2	510	28.1	522	28.7
8	1077	55.1	316	16.2	563	28.8
11	729	40.8	612	34.3	444	24.9

Classification accuracy and consistency are indices of agreement for performance-level classification as a score. Classification accuracy is a way to estimate the difference between true classification and observed classification due to measurement error. Classification consistency is a way to estimate the difference between the observed classification and the classification on a parallel form. The MI-Access Functional Independence classification accuracy and consistency indices were calculated by applying procedures given in Livingston and Lewis (1995) via the BB-CLASS computer program (Brennan, 2004). These indices are presented in the following table, Table 7.2. The accuracy indices can be interpreted as the proportion of examinees that would be classified accurately into the performance-level score categories given infinite replications of identical conditions. The consistency indices can be interpreted as the proportion of examinees that would be classified into the same performance-level score categories on the assessment and a parallel form of the assessment.

Table 7.2
Estimated Classification Accuracy and Consistency by Subject and Grade

	2 Categories Emerging vs Attained plus Surpassed		3 Categories Emerging vs Attained vs Surpassed	
Grade	Accuracy	Consistency	Accuracy	Consistency
ELA				
3	0.92	0.88	0.80	0.73
4	0.91	0.88	0.81	0.74
5	0.93	0.91	0.85	0.80
6	0.94	0.92	0.86	0.81
7	0.96	0.94	0.89	0.85
8	0.96	0.95	0.89	0.84
11	0.96	0.95	0.87	0.82
Mathematics				
3	0.90	0.85	0.77	0.69
4	0.93	0.90	0.83	0.76
5	0.90	0.85	0.78	0.70
6	0.92	0.88	0.79	0.71
7	0.88	0.83	0.76	0.68
8	0.90	0.85	0.77	0.69
11	0.90	0.86	0.83	0.76
Science				
5	0.84	0.78	0.71	0.63
8	0.85	0.79	0.75	0.69
11	0.85	0.79	0.75	0.66

The classification accuracy when categorizing students into the NCLB categories of proficient (attained + surpassed) and not proficient (emerging), is at least 88% for ELA and Mathematics, and the classification consistency is at least 83%. For Science, the accuracy and consistency indices are somewhat smaller. Across all grades and the three content areas, the classification accuracy when categorizing students into three categories (emerging, attained, and surpassed) is 71% or higher and the classification consistency is 63% or higher. The 71% and 63% are for the tests with the lowest reliability where a three category classification would have the greatest effect on the agreement indices. The accuracy indices will be higher than the consistency indices because the former estimates accuracy between observed scores containing measurement error and true scores with no error, whereas the latter estimates consistency between observed scores on parallel forms of the assessment where both scores contain measurement error.

These estimates represent strong proportions of students classified accurately for an assessment of the length appropriate for students with disabilities such as those that take the MI-Access Functional Independence assessments.

The State Summary Reports for Accessing Print and Mathematics contain longitudinal performance level change between last year and this year for grades 4 – 8. Those results are summarized in Table 7.3 for Accessing Print and in Table 7.4 for Mathematics. The first column contains the total number of students assessed in 2010 and matched to 2009 and the second column is that number expressed as a percent. The five subsequent columns show the number and percent of students by performance level change category; significant decline, decline, maintaining, improvement, and significant improvement. Except for grade 4 Accessing Print, a much higher percent of students improved (improvement plus significant improvement) than declined (significant decline plus decline). The results were quite mixed for Mathematics. A much higher percent of students improved at grades 4 and 8, but a much higher percent declined at grades 5 and 7. Grade 6 had about the same percent.

Table 7.3
Number and Percent of Students by Performance Level Change Between 2009 and 2010
Accessing Print

Grade	N	Percent Matched	Significant Decline	Decline	Maintaining	Improvement	Significant Improvement
4	1391	63.3	143	418	379	361	90
			10.3	30.1	27.2	26.0	6.5
5	1543	70.7	87	277	407	550	222
			5.6	18.0	26.4	35.6	14.4
6	1522	69.9	120	324	462	467	149
			7.9	21.3	30.4	30.7	9.8
7	1553	71.1	57	300	560	504	132
			3.7	19.3	36.1	32.5	8.5
8	1559	72.0	66	381	598	430	84
			4.2	24.4	38.4	27.6	5.4

Table 7.4
 Number and Percent of Students by Performance Level Change Between 2009 and 2010
 Mathematics

Grade	N	Percent Matched	Significant Decline	Decline	Maintaining	Improvement	Significant Improvement
4	1189	63.0	61	242	303	471	112
			5.1	20.4	25.5	39.6	9.4
5	1382	69.7	181	526	348	273	54
			13.1	38.1	25.2	19.8	3.9
6	1408	68.5	117	412	374	405	100
			8.3	29.3	26.6	28.8	7.1
7	1489	69.6	149	507	374	390	69
			10.0	34.0	25.1	26.2	4.6
8	1492	71.7	48	355	438	517	134
			3.2	23.8	29.4	34.7	9.0

8. Differential Item Functioning

Differential Item Functioning (DIF) is assessed via the DIFAS computer software (Penfield, 2007). Several DIF statistics are computed with the program, but the ETS category (Zieky, 1993) is the one reported to the Content Advisory Committee (CAC) and the Sensitivity Review Committee (SRC). The ETS scheme categorizes items as follows: A represents negligible DIF, B represents moderate DIF, and C represents large DIF. DIF analyses were run on the following four pairs of comparison groups; gender, white students versus black students, non-economically disadvantaged students versus economically disadvantaged students, and non-accommodated students versus accommodated students. DIF analyses were run on all core and field-test items, but only results for the field test items went to the CAC and the SRC.

The number of items by ETS category and grade for each of the four comparison groups is reported in Table 8.1 for ELA, in Table 8.2 for Mathematics, and in Table 8.3 for Science. Also reported in each table is the number of items favoring each group within each of the four comparison groups. Table 8.4 shows the percent of items by ETS category across all grades for each content area. Across the grades for all three content areas for the gender, race, and economic DIF comparisons, at least 94% of the items were classified with negligible DIF. ELA and Mathematics gender and race had one item classified with large DIF, and ELA economically disadvantaged had two items classified with large DIF. Science had no items classified with large DIF. Across all grades for the accommodation comparison, 92% of the Mathematics items and 99% of the Science items were classified with negligible DIF, and only 1% of the Mathematics items and none of the Science items were classified with large DIF. The results for ELA for the accommodation comparison, however, were noticeably different from all the other results. Here only 79% of the items were classified with negligible DIF and 3.5% were classified with large DIF. It would appear having an accommodation had a much greater effect on the ELA assessments than on the mathematics and science assessments. Within ELA, the grades 3 and 5 results were noticeably different from the results for the other grades where only 60 and 66 percent of the items, respectively, were classified with negligible DIF whereas the other grades had at least 85% classified with negligible DIF. For items classified with moderate or large DIF, the number and percent of such items favoring the reference group (male, white, non-economically disadvantaged, and non-accommodated) and the focal group was determined. Across all grades, the favored group results were mixed by content area and comparison group. In some cases, about the same percentage of items favored each group and in the other cases sometimes a greater percentage of items favored the reference group and sometimes it favored the focal group. All field-test items classified with moderate or large DIF were given special attention during the review process by the SRC.

Table 8.1
ELA DIF Summary

	Male vs. Female	White vs. Black	Non-Econ Dis vs. Econ Dis	Non-Accommodated vs. Accommodated
Grade 3				
A: Negligible DIF	95	95	97	60
B: Moderate DIF	5	5	2	32
C: Large DIF	0	0	1	8
Favoring Reference ¹	0	3	1	23
Favoring Focal ¹	5	2	2	17
Grade 4				
A or AA	102	103	101	90
B or BB	4	3	5	14
C or CC	0	0	0	2
Favoring Reference ¹	2	2	2	7
Favoring Focal ¹	2	1	3	9
Grade 5				
A: Negligible DIF	102	95	96	67
B: Moderate DIF	0	6	6	28
C: Large DIF	0	1	0	7
Favoring Reference ¹	0	3	3	20
Favoring Focal ¹	0	4	3	15
Grade 6				
A: Negligible DIF	95	96	97	85
B: Moderate DIF	7	6	5	14
C: Large DIF	0	0	0	3
Favoring Reference ¹	3	2	3	7
Favoring Focal ¹	4	4	2	10
Grade 7				
A or AA	99	102	100	86
B or BB	7	4	5	17
C or CC	0	0	1	3
Favoring Reference ¹	6	2	3	6
Favoring Focal ¹	1	2	3	14
Grade 8				
A: Negligible DIF	99	97	97	86
B: Moderate DIF	2	5	5	15
C: Large DIF	1	0	0	1
Favoring Reference ¹	0	3	4	9
Favoring Focal ¹	3	2	1	7
Grade 11				
A or AA	63	62	64	65
B or BB	2	3	1	0
C or CC	0	0	0	0
Favoring Reference ¹	1	3	0	0
Favoring Focal ¹	1	0	1	0

¹Number of Category B or C items favoring the Reference or first subgroup given in the column head, e.g. Male, and the number favoring the focal or second subgroup given in the column head, e.g., Female.

Table 8.2
Mathematics DIF Summary

	Male vs. Female	White vs. Black	Non-Econ Dis vs. Econ Dis	Non-Accommodated vs. Accommodated
Grade 3				
A: Negligible DIF	53	52	49	50
B: Moderate DIF	1	2	5	4
C: Large DIF	0	0	0	0
Favoring Reference ¹	1	0	3	1
Favoring Focal ¹	0	2	2	3
Grade 4				
A: Negligible DIF	53	49	52	48
B: Moderate DIF	1	5	2	5
C: Large DIF	0	0	0	1
Favoring Reference ¹	1	3	0	3
Favoring Focal ¹	0	2	2	3
Grade 5				
A: Negligible DIF	50	49	52	45
B: Moderate DIF	4	4	2	8
C: Large DIF	0	1	0	1
Favoring Reference ¹	4	1	0	3
Favoring Focal ¹	0	4	2	6
Grade 6				
A: Negligible DIF	60	63	63	61
B: Moderate DIF	5	2	2	4
C: Large DIF	0	0	0	0
Favoring Reference ¹	1	1	0	2
Favoring Focal ¹	4	1	2	2
Grade 7				
A: Negligible DIF	63	65	64	64
B: Moderate DIF	2	0	1	1
C: Large DIF	0	0	0	0
Favoring Reference ¹	1	0	0	1
Favoring Focal ¹	1	0	1	0
Grade 8				
A: Negligible DIF	61	60	64	64
B: Moderate DIF	4	5	1	1
C: Large DIF	0	0	0	0
Favoring Reference ¹	3	4	0	0
Favoring Focal ¹	1	1	1	1
Grade 11				
A: Negligible DIF	57	55	58	51
B: Moderate DIF	2	5	2	7
C: Large DIF	1	0	0	2
Favoring Reference ¹	2	3	2	8
Favoring Focal ¹	1	2	0	1

¹Number of Category B or C items favoring the Reference or first subgroup given in the column head, e.g. Male, and the number favoring the focal or second subgroup given in the column head, e.g., Female.

Table 8.3
Science DIF Summary

	Male vs. Female	White vs. Black	Non-Econ Dis vs. Econ Dis	Non-Accommodated vs. Accommodated
Grade 5				
A: Negligible DIF	50	51	50	49
B: Moderate DIF	1	0	1	2
C: Large DIF	0	0	0	0
Favoring Reference ¹	1	0	0	0
Favoring Focal ¹	0	0	1	2
Grade 8				
A: Negligible DIF	59	56	58	60
B: Moderate DIF	1	4	2	0
C: Large DIF	0	0	0	0
Favoring Reference ¹	1	1	1	0
Favoring Focal ¹	0	3	1	0
Grade 11				
A: Negligible DIF	63	62	64	65
B: Moderate DIF	2	3	1	0
C: Large DIF	0	0	0	0
Favoring Reference ¹	1	3	0	0
Favoring Focal ¹	1	0	1	0

¹Number of Category B or C items favoring the Reference or first subgroup given in the column head, e.g. Male, and the number favoring the focal or second subgroup given in the column head, e.g., Female.

Table 8.4
DIF Summary Across Grades

	Male vs. Female	White vs. Black	Non-Econ Dis vs. Econ Dis	Non-Accommodated vs. Accommodated
ELA				
Total Across All Grades				
A: Negligible DIF	655	650	652	539
B: Moderate DIF	27	32	29	120
C: Large DIF	1	1	2	24
Favoring Reference ¹	12	18	16	72
Favoring Focal ¹	16	15	15	72
Total	683	683	683	683
Percents of Total Across All Grades				
A: Negligible DIF	95.9%	95.2%	95.5%	78.9%
B: Moderate DIF	4.0%	4.7%	4.2%	17.6%
C: Large DIF	0.1%	0.1%	0.3%	3.5%
Favoring Reference ¹	1.8%	2.6%	2.3%	10.5%
Favoring Focal ¹	2.3%	2.2%	2.2%	10.5%
Mathematics				
Total Across All Grades				
A: Negligible DIF	397	393	402	383
B: Moderate DIF	19	23	15	30
C: Large DIF	1	1	0	4
Favoring Reference ¹	13	12	5	18
Favoring Focal ¹	7	12	10	16
Total	417	417	417	417
Percents of Total Across All Grades				
A: Negligible DIF	95.2%	94.2%	96.4%	91.8%
B: Moderate DIF	4.6%	5.5%	3.6%	7.2%
C: Large DIF	0.2%	0.2%	0.0%	1.0%
Favoring Reference ¹	3.1%	2.9%	1.2%	4.3%
Favoring Focal ¹	1.7%	2.9%	2.4%	3.8%

Table 8.4 (cont.)
DIF Summary Across Grades

	Male vs. Female	White vs. Black	Non-Econ Dis vs. Econ Dis	Non-Accommodated vs. Accommodated
Science				
Total Across All Grades				
A: Negligible DIF	172	169	172	174
B: Moderate DIF	4	7	4	2
C: Large DIF	0	0	0	0
Favoring Reference ¹	3	4	1	0
Favoring Focal ¹	1	3	3	2
Total	176	176	176	176
Percents of Total Across All Grades				
A: Negligible DIF	97.7%	96.0%	97.7%	98.9%
B: Moderate DIF	2.3%	4.0%	2.3%	1.1%
C: Large DIF	0.0%	0.0%	0.0%	0.0%
Favoring Reference ¹	1.7%	2.3%	0.6%	0.0%
Favoring Focal ¹	0.6%	1.7%	1.7%	1.1%

¹Number of Category B or C items favoring the Reference or first subgroup given in the column head, e.g. Male, and the number favoring the focal or second subgroup given in the column head, e.g., Female.

9. Interrelations Among Strands within Measures

One important source of validity evidence is the consistency of the relations of test subcomponents – interrelations among strands within the test. The correlations were computed based on subscore raw scores and estimated as Pearson product-moment correlations in SPSS.

The correlation between multiple choice (Accessing Print) and constructed response (Expressing Ideas) scores by grade is given in Table 9.1. The correlations are .30, .37, and .45 at grades 4, 7, and 11, respectively. Table 9.2 contains mean Accessing Print scores for each possible Expressing Ideas score, where Accessing Print scores increase consistently as Expressing Ideas scores increase from 1 to 4. Expressing Ideas scores of 0 are difficult to interpret in a consistent way since this score results from a number of alternative non–scorable responses or condition codes. Table 9.3 contains the ELA strand intercorrelations by grade. Across the grades, the three types of passages are typically correlated among each other in the .50s. They are each correlated very highly with text comprehension, but this is not surprising since text comprehension consists of the three passages. At grades 4, 7, and 11, Expressing Ideas is correlated from the high teens to the high .30s with each of the passage types and with Word Recognition. These are moderately high correlations given the maximum score is only four for Expressing Ideas. The correlation between Word Recognition and Text Comprehension is from .29 at grade 4 to a high of .58 at grade 8.

Table 9.4 contains the Mathematics strand intercorrelations for grades 3 to 8, and Table 9.5 contains the intercorrelations for grade 11. Across the grades, the intercorrelations among Mathematics strands range from the high .20s to the mid .60s. Lower correlations are associated with strands with only two or three items and the items are very difficult.

Table 9.6 contains the Science strand intercorrelations. These intercorrelations are in the mid .20s to the mid .50s where the lower correlations are for the three- or four-item Constructing and Reflecting strand with the other strands.

The N, mean, standard deviation, and Cronbach's Coefficient Alpha along with the minimum and maximum score of the strand scores are also provided. These summary statistics are given in Table 9.7 for ELA, in Table 9.8 for Mathematics, and in Table 9.9 for Science.

Table 9.1
Correlations between Multiple Choice (Accessing Print) and Constructed Response
(Expressing Ideas) Scores

Grade	N	Correlation
4	2,196	0.30
7	2,184	0.37
11	1,788	0.45

Table 9.2
Mean Accessing Print Score by Expressing Ideas Prompt Score

Grade	Expressing Ideas Score	N	Accessing Print	
			Mean	Standard Deviation
4	0	138	22.0	8.2
	1	501	23.5	7.5
	2	1098	27.5	7.3
	3	352	29.3	7.3
	4	107	30.9	6.8
7	0	153	26.5	8.4
	1	375	26.2	7.8
	2	1225	31.3	6.8
	3	357	34.5	5.5
	4	74	36.6	4.8
11	0	191	28.3	8.8
	1	151	25.2	7.9
	2	680	32.5	7.2
	3	510	35.9	4.9
	4	256	37.6	3.7

Table 9.3
 ELA Strand Pearson Product-Moment Intercorrelations by Grade

	Informational Passage	Narrative Passage	Functional Passage	Word Recognition	Text Comprehension
Grade 3					
Narrative Passage	0.53				
Functional Passage	0.53	0.55			
Word Recognition	0.32	0.24	0.29		
Text Comprehension	0.82	0.85	0.84	0.34	
Grade 4					
Narrative Passage	0.51				
Functional Passage	0.45	0.48			
Word Recognition	0.38	0.39	0.47		
Text Comprehension	0.81	0.82	0.80	0.51	
Expressing Ideas	0.17	0.17	0.23	0.29	0.23
Grade 5					
Narrative Passage	0.53				
Functional Passage	0.50	0.54			
Word Recognition	0.36	0.40	0.46		
Text Comprehension	0.82	0.83	0.82	0.49	
Grade 6					
Narrative Passage	0.52				
Functional Passage	0.59	0.57			
Word Recognition	0.44	0.50	0.49		
Text Comprehension	0.84	0.82	0.87	0.57	
Grade 7					
Narrative Passage	0.51				
Functional Passage	0.55	0.56			
Word Recognition	0.43	0.50	0.47		
Text Comprehension	0.83	0.82	0.86	0.56	
Expressing Ideas	0.25	0.30	0.29	0.32	0.34
Grade 8					
Narrative Passage	0.50				
Functional Passage	0.61	0.56			
Word Recognition	0.45	0.52	0.50		
Text Comprehension	0.85	0.80	0.87	0.58	
Grade 11					
Narrative Passage	0.61				
Functional Passage	0.50	0.56			
Word Recognition	0.50	0.57	0.55		
Text Comprehension	0.85	0.86	0.80	0.64	
Expressing Ideas	0.35	0.36	0.34	0.39	0.42

Table 9.4
Mathematics Strand Pearson Product-Moment Intercorrelations for Grades 3 – 8

	Numbers & Operations	Measurement	Geometry	Data & Probability
Grade 3				
Measurement	0.57			
Geometry	0.51	0.52		
Data & Probability	0.41	0.40	0.39	
Grade 4				
Measurement	0.61			
Geometry	0.50	0.39		
Data & Probability	0.47	0.38	0.34	
Grade 5				
Measurement	0.64			
Geometry	0.37	0.35		
Data & Probability	0.45	0.43	0.27	
Grade 6				
Measurement	0.65			
Geometry	0.22	0.20		
Data & Probability	0.42	0.43	0.13	
Grade 7				
Measurement	0.61			
Geometry	0.32	0.32		
Data & Probability	0.52	0.49	0.27	
Grade 8				
Measurement	0.61			
Geometry	0.41	0.43		
Data & Probability	0.43	0.36	0.27	
Algebra	0.28	0.20	0.09	0.14

Table 9.5
Mathematics Strand Pearson Product-Moment Intercorrelations for Grade 11

	Patterns & Relationships	Geometry & Measurement	Data & Probability	Numbers & Operations
Grade 11				
Geometry & Measurement	0.58			
Data & Probability	0.38	0.46		
Numbers & Operations	0.49	0.66	0.40	
Algebra	0.35	0.46	0.27	0.52

Table 9.6
Science Strand Pearson Product-Moment Intercorrelations

	Constructing & Reflecting	Life Science	Physical Science
Grade 5			
Life Science	0.26		
Physical Science	0.32	0.46	
Earth Science	0.28	0.43	0.52
Grade 8			
Life Science	0.43		
Physical Science	0.45	0.48	
Earth Science	0.49	0.47	0.55
Grade 11			
Life Science	0.34		
Physical Science	0.35	0.49	
Earth Science	0.37	0.44	0.49

Table 9.7
ELA Strand Summary Statistics

	N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
Grade 3						
Informational Passage	1820	0	7	4.50	1.654	0.55
Narrative Passage	1820	0	7	4.43	1.849	0.63
Functional Passage	1820	0	7	4.32	1.758	0.57
Word Recognition	1820	0	20	16.34	3.906	0.86
Text Comprehension	1820	0	21	13.25	4.378	0.80
Grade 4						
Informational Passage	2196	0	7	3.74	1.754	0.53
Narrative Passage	2196	0	7	4.59	1.788	0.63
Functional Passage	2196	0	7	4.31	1.781	0.59
Word Recognition	2196	0	20	14.08	4.617	0.85
Text Comprehension	2196	0	21	12.64	4.309	0.78
Expressing Ideas	2196	0	4	1.90	0.909	NA
Grade 5						
Informational Passage	2182	0	7	4.07	1.773	0.56
Narrative Passage	2182	0	7	4.97	1.751	0.65
Functional Passage	2182	0	7	4.70	1.778	0.63
Word Recognition	2182	0	20	15.62	4.481	0.88
Text Comprehension	2182	0	21	13.74	4.376	0.81
Grade 6						
Informational Passage	2176	0	7	4.04	1.820	0.58
Narrative Passage	2176	0	7	5.09	1.724	0.65
Functional Passage	2176	0	7	4.66	1.913	0.68
Word Recognition	2176	0	20	15.19	4.292	0.86
Text Comprehension	2176	0	21	13.79	4.592	0.83
Grade 7						
Informational Passage	2184	0	7	4.28	1.795	0.58
Narrative Passage	2184	0	7	5.34	1.638	0.64
Functional Passage	2184	0	7	5.04	1.866	0.70
Word Recognition	2184	0	20	16.11	4.006	0.86
Text Comprehension	2184	0	21	14.65	4.416	0.82
Expressing Ideas	2184	0	4	1.92	0.863	NA
Grade 8						
Informational Passage	2163	0	7	4.55	1.815	0.61
Narrative Passage	2163	0	7	5.57	1.566	0.65
Functional Passage	2163	0	7	5.16	1.868	0.72
Word Recognition	2163	0	20	16.72	3.911	0.87
Text Comprehension	2163	0	21	15.28	4.417	0.84
Grade 11						
Informational Passage	1788	0	7	4.97	1.865	0.69
Narrative Passage	1788	0	7	5.20	1.711	0.67
Functional Passage	1788	0	7	5.91	1.531	0.71
Word Recognition	1788	0	20	17.06	3.899	0.89
Text Comprehension	1788	0	21	16.08	4.289	0.79
Expressing Ideas	1788	0	4	2.27	1.138	NA

Table 9.8
Mathematics Strand Summary Statistics

	N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
Grade 3						
Numbers & Operations	1573	0	10	6.57	2.105	0.62
Measurement	1573	0	8	5.39	1.859	0.58
Geometry	1573	0	9	7.15	1.623	0.55
Data & Probability	1573	0	3	2.60	0.716	0.48
Grade 4						
Numbers & Operations	1887	0	16	11.59	3.523	0.80
Measurement	1887	0	8	5.87	1.738	0.58
Geometry	1887	0	4	3.40	0.846	0.43
Data & Probability	1887	0	2	1.63	0.629	0.48
Grade 5						
Numbers & Operations	1984	0	16	10.79	3.393	0.75
Measurement	1984	0	10	6.63	2.108	0.60
Geometry	1984	0	2	1.58	0.624	0.30
Data & Probability	1984	0	2	1.68	0.574	0.39
Grade 6						
Numbers & Operations	2056	0	18	12.47	3.663	0.77
Measurement	2056	0	12	8.37	2.291	0.62
Geometry	2056	0	2	1.63	0.538	0.08
Data & Probability	2056	0	3	2.34	0.733	0.38
Grade 7						
Numbers & Operations	2138	0	17	9.99	3.306	0.69
Measurement	2138	0	12	8.10	2.333	0.60
Geometry	2138	0	3	2.18	0.773	0.21
Data & Probability	2138	0	3	2.00	0.937	0.42
Grade 8						
Numbers & Operations	2082	0	17	9.97	3.292	0.69
Measurement	2082	0	10	6.78	1.978	0.60
Geometry	2082	0	3	2.37	0.826	0.45
Data & Probability	2082	0	3	1.83	0.863	0.20
Algebra	2082	0	2	0.77	0.743	0.28
Grade 11						
Patterns & Relationships	1794	0	4	2.96	1.225	0.65
Geometry & Measurement	1794	0	16	10.92	3.196	0.74
Data Analysis & Statistics	1794	0	2	1.48	0.680	0.37
Numbers & Operations	1794	0	15	8.54	3.165	0.70
Algebra	1794	0	3	1.54	1.002	0.41

Table 9.9
Science Strand Summary Statistics

	N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
Grade 5						
Constructing and Reflecting	1816	0	3	2.19	0.776	0.21
Life Science	1816	0	13	7.64	2.110	0.43
Physical Science	1816	0	12	6.88	2.408	0.56
Earth Science	1816	0	7	4.51	1.544	0.46
Grade 8						
Constructing and Reflecting	1956	0	4	2.74	1.086	0.39
Life Science	1956	0	14	8.15	2.447	0.51
Physical Science	1956	0	14	7.46	2.495	0.50
Earth Science	1956	0	8	5.17	1.794	0.55
Grade 11						
Constructing and Reflecting	1785	0	4	2.38	1.024	0.23
Life Science	1785	0	14	7.03	2.795	0.61
Physical Science	1785	0	15	7.21	2.558	0.49
Earth Science	1785	0	12	7.07	2.378	0.55

10. Summary of Items Flagged for Difficulty, Discrimination, and Model Fit

The Michigan Department of Education, Bureau of Assessment and Accountability established the following criteria for flagging items based on their difficulty, discrimination and model fit:

PL flag if the adjusted p-value is less than 0.33

PH flag if the adjusted p-value is greater than 0.90

CL flag if the item-total correlation is less than 0.25

MH flag if infit mean-square or outfit mean-square is greater than 2.0

MM flag if infit mean-square is between 1.5 – 2.0 and outfit mean-square is less than 2.0 or outfit mean-square is between 1.5 – 2.0 and infit mean-square is less than 2.0

TP flag if infit mean-square is less than 0.5 and outfit mean-square is less than 1.5 or outfit mean-square is less than 0.5 and infit mean-square is less than 1.5.

The adjusted p-value is equal to the item mean divided by the maximum item score. The infit and outfit mean-squares are statistics that provide information on how well items are fitting the Rasch model. Both statistics have an expectation of 1. For infit, values substantially less than 1 indicate dependency in the data; values substantially above 1 indicate noise. For outfit, values substantially less than 1 indicate dependency in the data; values substantially greater than 1 indicate the presence of unexpected outliers. Infit is sensitive to unexpected behavior affecting responses to items near the person's measure level whereas outfit is more sensitive to unexpected behavior by persons on items far from the person's measure level (Linacre, 2006).

Table 10.1 summarizes the number of flagged items by level, content area, and grade. The table also shows the total number of items, core and field test, for each assessment. Across grades and content areas, zero to 5 items were flagged for the lower bound difficulty (PL) and zero to 7 items were flagged for the upper bound difficulty (PH) except at grade 11 ELA where 13 items were flagged. Across grades 3 – 8 and grade 11 for ELA and mathematics, 8 to 23 items were flagged for discrimination (CL) whereas 24 to 33 items were flagged for discrimination at the three grades of science. Across grades and content areas, only one item at grades 5 and 6 Mathematics was flagged for the most severe model fit flag (MH). For the MM model fit flag, one to 7 items were flagged across the grades of ELA, zero to 4 were flagged across the grades of mathematics, and no such items were flagged for science. For the TP model fit flag, no items were flagged for Mathematics and Science and grades 4 – 7 ELA. At grades 3 and 8 ELA, one item was flagged for TP and three items were flagged at grade 11.

Table 10.1
Number of Flagged Items

Grade	Total Number of Items	Difficulty Flag		Item-Total Correlation Flag ³	Mean-square Fit Flag		
		PL ¹	PH ²		MH ⁴	MM ⁵	TP ⁶
ELA							
3	100	0	4	20	0	2	1
4	106	2	0	23	0	1	0
5	102	2	1	17	0	2	0
6	102	0	0	13	0	2	0
7	106	1	3	11	0	5	0
8	102	2	7	8	0	6	1
11	84	0	13	10	0	7	3
Mathematics							
3	54	2	7	15	0	2	0
4	54	3	2	12	0	4	0
5	54	2	2	13	1	1	0
6	65	5	3	18	1	1	0
7	65	4	2	18	0	1	0
8	65	5	3	23	0	0	0
11	60	3	3	12	0	3	0
Science							
5	51	1	6	24	0	0	0
8	60	4	0	33	0	0	0
11	65	4	0	31	0	0	0

¹PL = p-value < 0.33

²PH = p-value > 0.90

³Item-total correlation < 0.25

⁴Infit Mean-square or Outfit Mean-square > 2.0

⁵Infit Mean-square between 1.5 – 2.0 and Outfit Mean-square < 2.0
or Outfit Mean-square between 1.5 – 2.0 and Infit Mean-square < 2.0

⁶Infit Mean-square < 0.5 and Outfit Mean-square < 1.5
or Outfit Mean-square < 0.5 and Infit Mean-square < 1.5

11. Verification of Psychometric Procedures

As the independent psychometric quality assurance provider for the MI-Access Functional Independence program, HumRRO was responsible for reviewing and assuring that all psychometric procedures were carried out accurately by Questar Assessment, Inc. at each step of the equating process for Accessing Print (grades 3 – 8 and 11), Mathematics (grades 3 – 8 and 11), and Science (grades 5, 8, and 11). Two phases of the process were checked: (a) equating with core items and (b) final item analyses with core and field-test items.

Equating with Core Items:

HumRRO checked and matched data from Questar for all assessments and grades at each of the following steps of the equating phase.

Classical Statistics:

The first step in the process was to check Questar's classical statistics results. HumRRO staff wrote custom SAS® programs to calculate a predetermined set of statistics variables. HumRRO compared their results to Questar's. The variables checked were:

Number of Students

P-value or item mean divided by maximum score

Item Standard Deviation

Corrected/Adjusted Item-Total Correlation—For core items the total score excludes the item. For field-test items, correlation is between the item and the total raw score for core items.

Number of students with multiple marks on MC items

Number of students with condition code A on prompt

Number of students with condition code B on prompt

Number of students with condition code C on prompt

Number of students with condition code D on prompt or omit

Number of students with score of 1 point on prompt or selecting option A for MC items

Number of students with score of 2 points on prompt or selecting option B for MC items

Number of students with score of 3 points on prompt or selecting option C for MC items

Number of students with score of 4 points on prompt

Number of students who had multi-marked answer for MC item

Percent of students with condition code A for a CR item

Percent of students with condition code B for a CR item

Percent of students with condition code C for a CR item

Percent of students with condition code D for a CR item

Corrected point biserial correlation for option A for MC items

Corrected point biserial correlation for option B for MC items

Corrected point biserial correlation for option C for MC items

Corrected point biserial correlation for those scoring 1 point, CR items only

Corrected point biserial correlation for those scoring 2 points, CR items only

Corrected point biserial correlation for those scoring 3 points, CR items only

Corrected point biserial correlation for those scoring 4 points, CR items only

Corrected point biserial correlation for those with omitted MC item

Corrected point biserial correlation for "blank/refused to respond"

P-Value Flag: if an item's p-value was less than 0.33 or greater than 0.90

Item-Total Correlation Flag: if the point biserial was less than 0.25.

HumRRO matched all the variables (which were rounded to the second decimal place) at all grade/subject levels.

Calibration:

Calibration was done using Winsteps. For each grade/subject, HumRRO matched both Winsteps output files [IFILES (.ITM) and SFILES .ISF)] that were provided by Questar: all comparisons for each grade/subject were exact matches. An example of the verification of the calibration step is shown in Table 11.1.

Table 11.1
Example of Winsteps verification record provided to MDE

Match results? (HumRRO vs Questar)		
	Winsteps Output Files	
Subject/Grade	.ITM	.ISF
MA03	Yes	Yes
MA04	Yes	Yes
MA05	Yes	Yes
MA06	Yes	Yes
MA07	Yes	Yes
MA08	Yes	Yes

Equating:

HumRRO matched Questar's linking constants (LCs). Table 11.2 shows the LCs that were calculated and matched between HumRRO and Questar for Mathematics, Science, and Accessing Print.

Table 11.2
Linking Constant (LC) Comparison

Subject/ Grade	Questar LC	HumRRO LC
M03	0.104	0.104
M04	-0.039	-0.039
M05	-0.064	-0.064
M06	-0.048	-0.048
M07	0.055	0.055
M08	-0.166	-0.166
S05	0.010	0.010
S08	-0.054	-0.054
AP03	-0.112	-0.112
AP04	-0.023	-0.023
AP05	0.036	0.036
AP06	-0.072	-0.072
AP07	0.034	0.034
AP08	0.001	0.001

Raw Score to Scale Score Tables:

HumRRO used the LCs to calculate the RS–SS tables in an Excel spreadsheet. A separate spreadsheet program was developed for comparison purposes. Questar’s scale score results were copied and pasted into this spreadsheet and subtracted from the HumRRO-calculated scale score at each raw score point. There were no differences in any of the grade/subjects RS–SS conversion tables.

Final Item Analyses with Core and Field-Test Items:

HumRRO checked and matched field-test item data from Questar for all assessments and grades. The field-test analysis checks included verification of classical item statistics, calibration of field-test items, and calculation of differential item functioning (DIF) statistics.

Classical Statistics:

The classical statistics described above were computed for the Total population and the subgroups male/female, white/black, economically disadvantaged/non-economically disadvantaged, and accommodated/non-accommodated. HumRRO matched all variables for all comparison groups. An example of a verification table for Mathematics and Science is seen in Table 11.3. Similar verification tables were produced for AP.

Table 11.3

Example of HumRRO’s verification tables for classical statistics on field test items, by subgroup

Subject/Grade	Group	Match Results? (HumRRO vs Questar)	Number of Items (Core/FT)
MA11	All	Yes	40/20
	Male	Yes	40/20
	Female	Yes	40/20
	White	Yes	40/20
	Black	Yes	40/20
	Econ-Disadv	Yes	40/20
	Non-econ disadv	Yes	40/20
	Accommodated	Yes	40/20
	Non-accommo	Yes	40/20
SC11	All	Yes	45/20
	Male	Yes	45/20
	Female	Yes	45/20
	White	Yes	45/20
	Black	Yes	45/20
	Econ-Disadv	Yes	45/20
	Non-econ disadv	Yes	45/20
	Accommodated	Yes	45/20
	Non-accommo	Yes	45/20

Winsteps Output Files:

Field-test items were calibrated by anchoring the core items to the values obtained during Equating. HumRRO checked Winsteps output (using version 3.67) from the calibration of core and field test items. As seen in Table 11.4, all of HumRRO's output, for all grades, matched Questar's exactly for Mathematics, Accessing Print, and Science.

Table 11.4
Verification of matches of Winsteps output between HumRRO and Questar for
Mathematics, Accessing Print, and Science

Match results? (HumRRO vs Questar)		
Subject/Grade	Winsteps Output Files	
	.ITM	.ISF
MA03	Yes	Yes
MA04	Yes	Yes
MA05	Yes	Yes
MA06	Yes	Yes
MA07	Yes	Yes
MA08	Yes	Yes
MA11	Yes	Yes
AP03	Yes	Yes
AP04	Yes	Yes
AP05	Yes	Yes
AP06	Yes	Yes
AP07	Yes	Yes
AP08	Yes	Yes
AP11	Yes	Yes
SC05	Yes	Yes
SC08	Yes	Yes
SC11	Yes	Yes

Differential Item Functioning (DIF) Statistics:

HumRRO checked results of five DIF statistics for each of four sets of paired comparison groups. These five statistics were Mantel-Haenszel chi-square (MH-CHI), Delta, DIF category, Favored Group, and standardized mean difference (SMD). A program called DIFAS (version 4.0) was used to run all of the DIF analyses except SMD. HumRRO developed a custom program in SAS® to do the SMD calculation. The four sets of comparison groups that were checked for DIF were female(focal)/male, black(focal)/white, economically disadvantaged(focal)/non-economically disadvantaged, and accommodated(focal)/non-accommodated. An example of the verification of DIF results is shown in Table 11.5.

Table 11.5
Verification of the DIF analyses output between HumRRO and Questar for
Mathematics, Accessing Print, and Science

Match results? (HumRRO vs Questar)					
Subject/grade	MH CHI	Delta	DIF Category*	SMD	Favored Group
MA03	Yes	Yes	Yes	Yes	Yes
MA04	Yes	Yes	Yes	Yes	Yes
MA05	Yes	Yes	Yes	Yes	Yes
MA06	Yes	Yes	Yes	Yes	Yes
MA07	Yes	Yes	Yes	Yes	Yes
MA08	Yes	Yes	Yes	Yes	Yes
MA11	Yes	Yes	Yes	Yes	Yes
AP03	Yes	Yes	Yes	Yes	Yes
AP04	Yes	Yes	Yes	Yes	Yes
AP05	Yes	Yes	Yes	Yes	Yes
AP06	Yes	Yes	Yes	Yes	Yes
AP07	Yes	Yes	Yes	Yes	Yes
AP08	Yes	Yes	Yes	Yes	Yes
AP11	Yes	Yes	Yes	Yes	Yes
SC05	Yes	Yes	Yes	Yes	Yes
SC08	Yes	Yes	Yes	Yes	Yes
SC11	Yes	Yes	Yes	Yes	Yes

*The DIF categories of A, B, and C were used for dichotomous items (ETS) and AA, BB, CC were used for constructed response items (NAEP).

The following emails were sent by HumRRO to the Michigan Department of Education, Office of Educational Assessment and Accountability to announce when verification of a particular assessment had been made:

December 17, 2010 Verification of Fall Calibration Steps for MI-Access: Functional Independence. [Equating Results for Functional Independence Mathematics grades 3 – 8, Accessing Print grades 3 – 8, Science grades 5 and 8].

February 11, 2011 FI — Core and FT Verification for MI-Access [Final Item Analyses for Mathematics grades 3 – 8, Accessing Print grades 3 – 8, Science grades 5 and 8].

April 27, 2011 Verification of FI (MI-Access) Grade 11 subjects [Accessing Print, Math, Science— Functional Independence Equating Results].

May 31, 2011 FI — Core and FT Verification for Grade 11 MA & SC [Core and FT Analysis for MI-Access: Functional Independence].

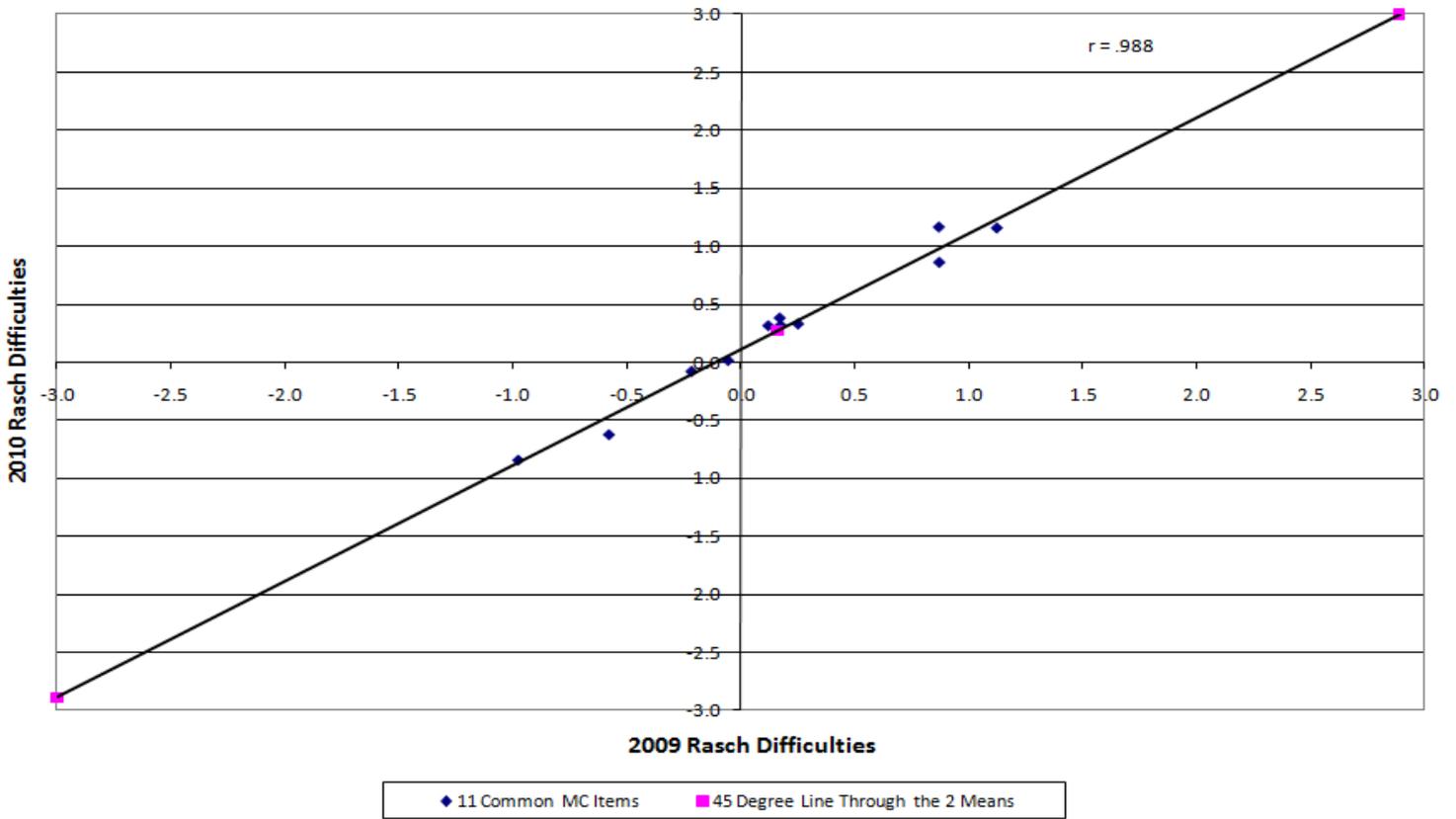
July 12, 2011 Verification of FI (MI-Access) Grade 11 — Accessing Print [Core and FT Analysis for Functional Independence].

REFERENCES

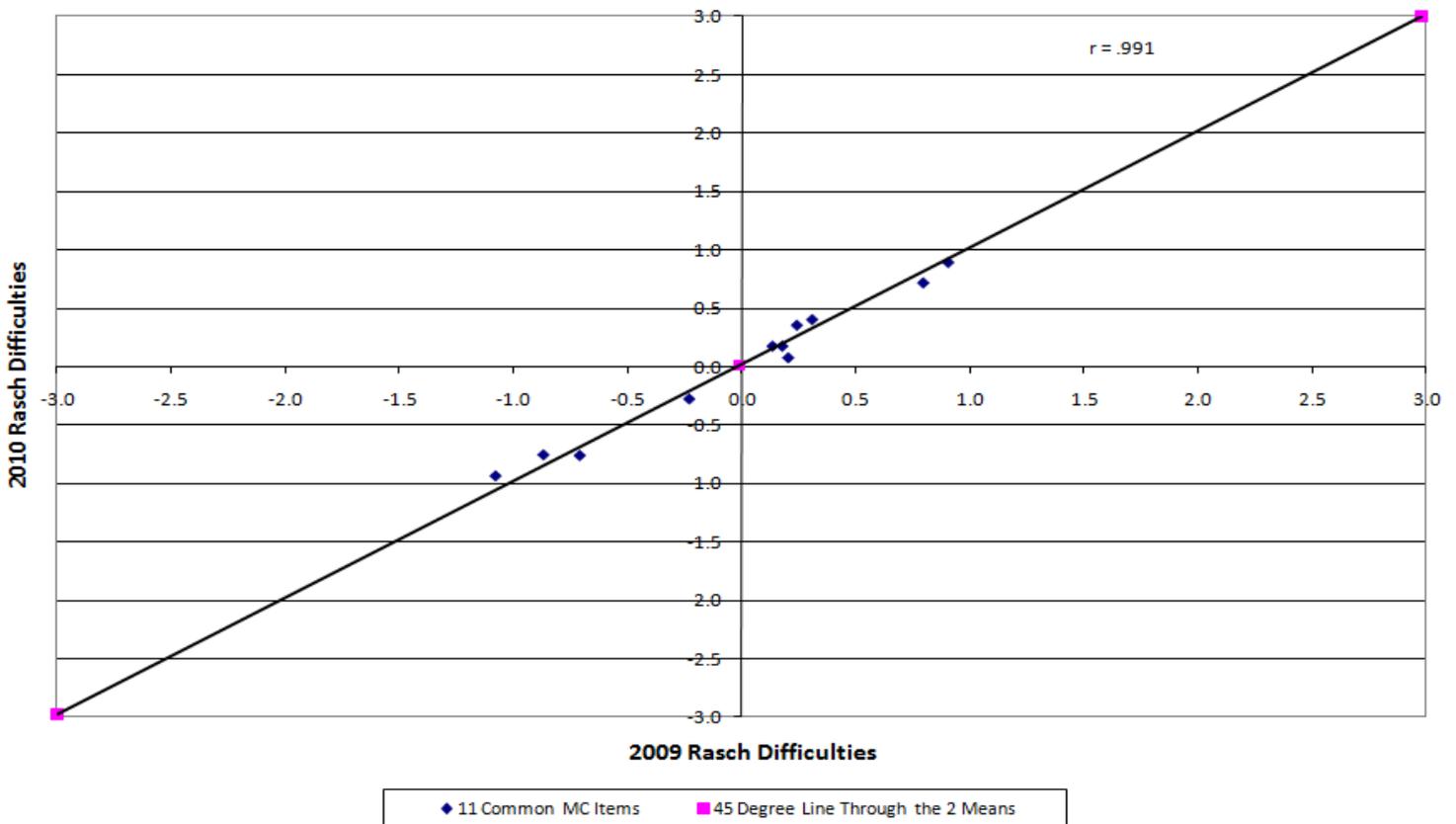
- Brennan, R. L. (2004). BB-CLASS: Beta-Binomial Classification Consistency and Accuracy, version 1.1. CASMA, University of Iowa, Iowa City, IA: Author.
- Linacre, J. M. (2006). WINSTEPS Rasch measurement computer program. Chicago: Winsteps.com
- Livingston, S. A. & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179–197.
- Penfield, R. D. (2007). DIFAS 4.0 Differential item functioning analysis system. Thousand Oaks, CA: Sage Publications.
- Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 337–347). Hillsdale, NJ: Erlbaum.

APPENDIX A: ANCHOR ITEM PLOTS

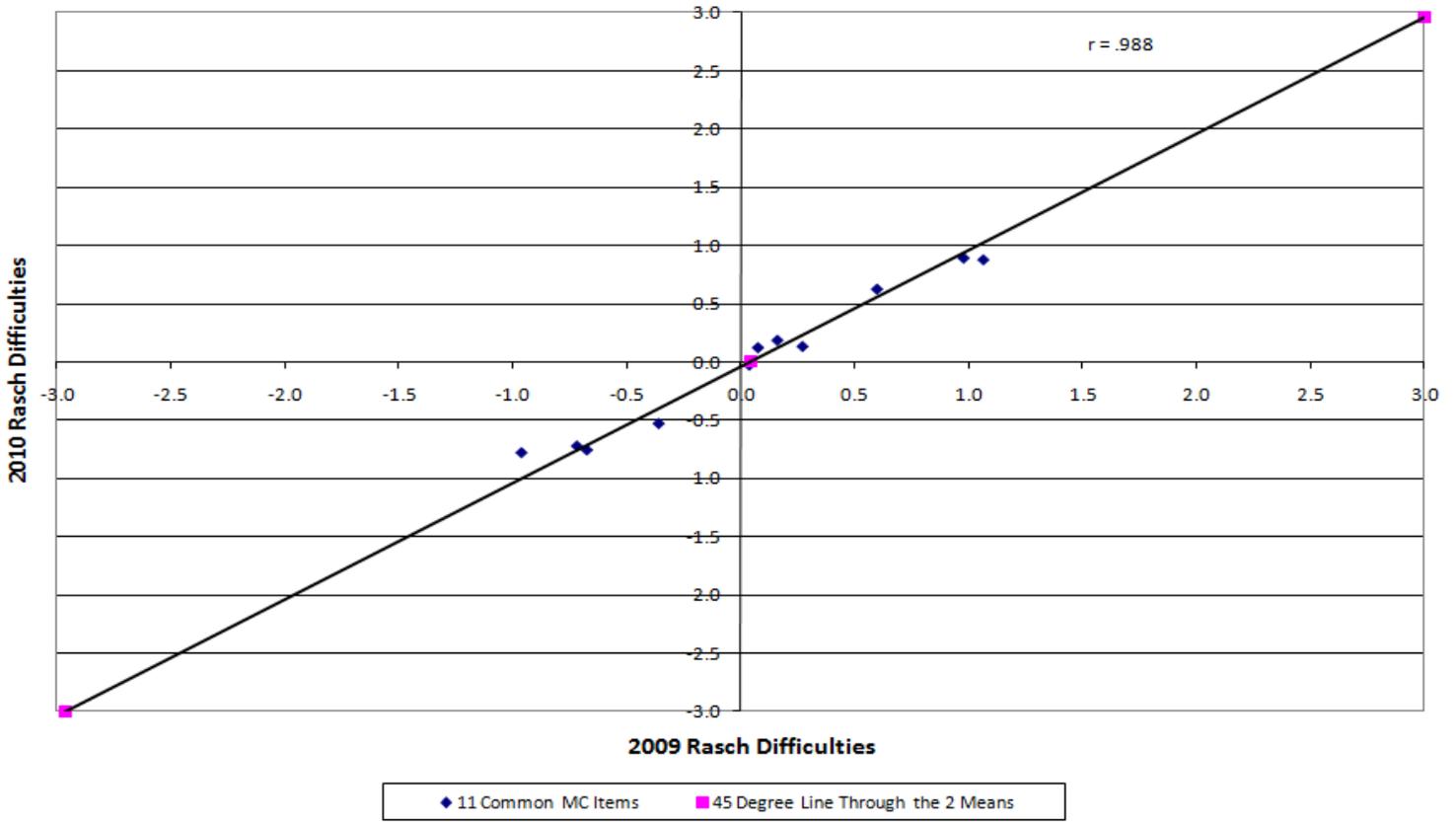
Fall 2010 Accessing Print Grade 3 Anchor Items



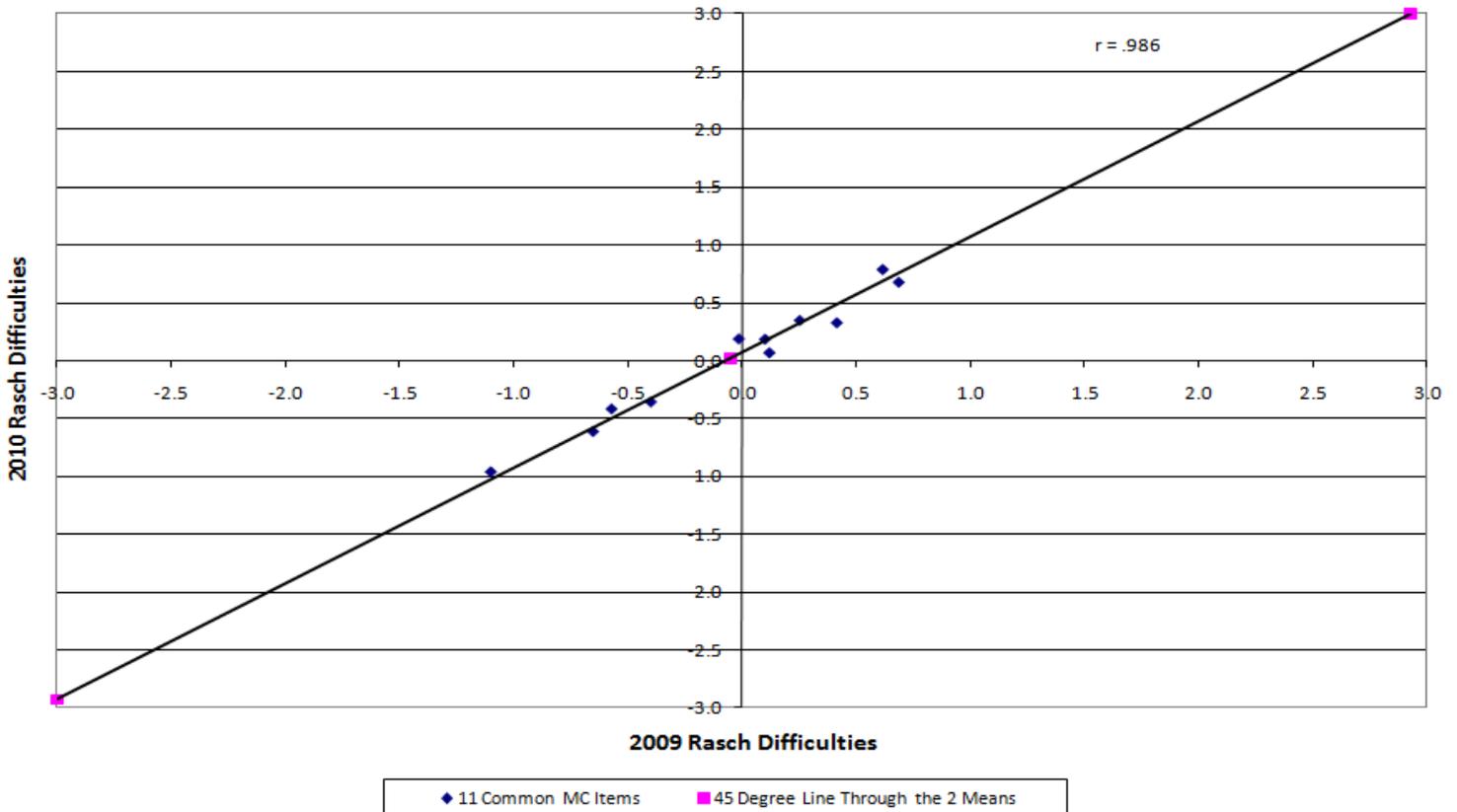
Fall 2010 Accessing Print Grade 4 Anchor Items



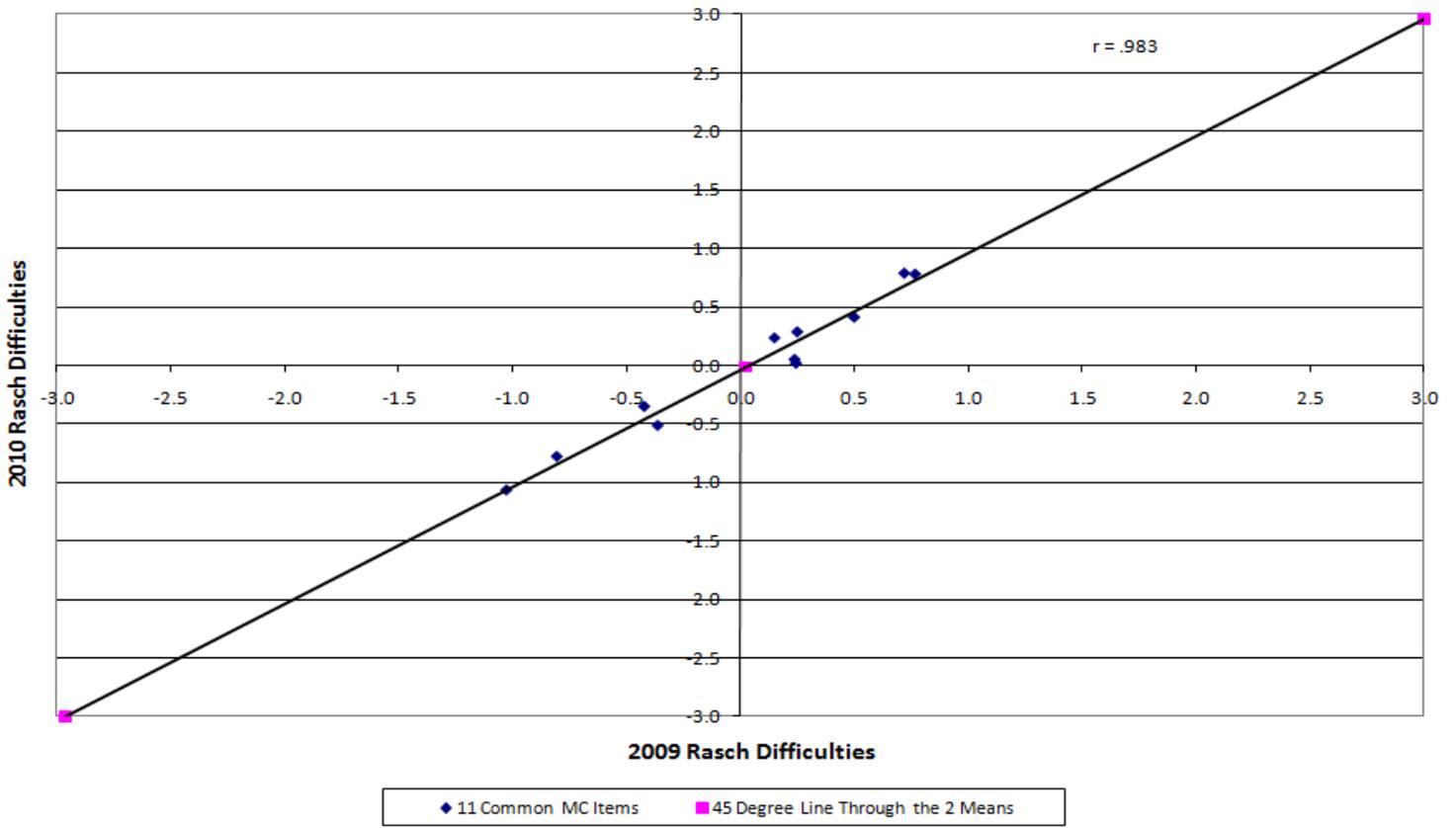
Fall 2010 Accessing Print Grade 5 Anchor Items



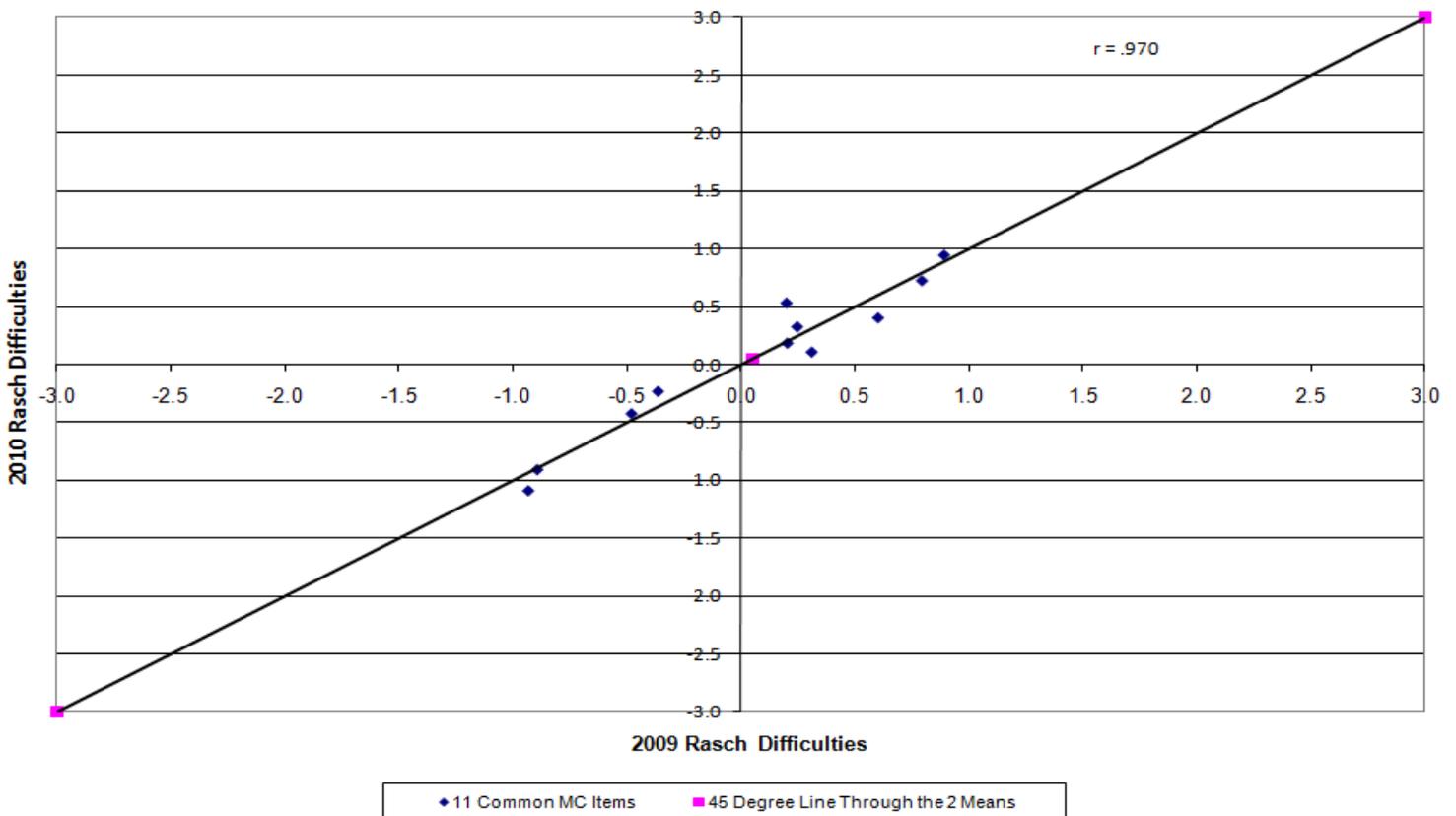
Fall 2010 Accessing Print Grade 6 Anchor Items



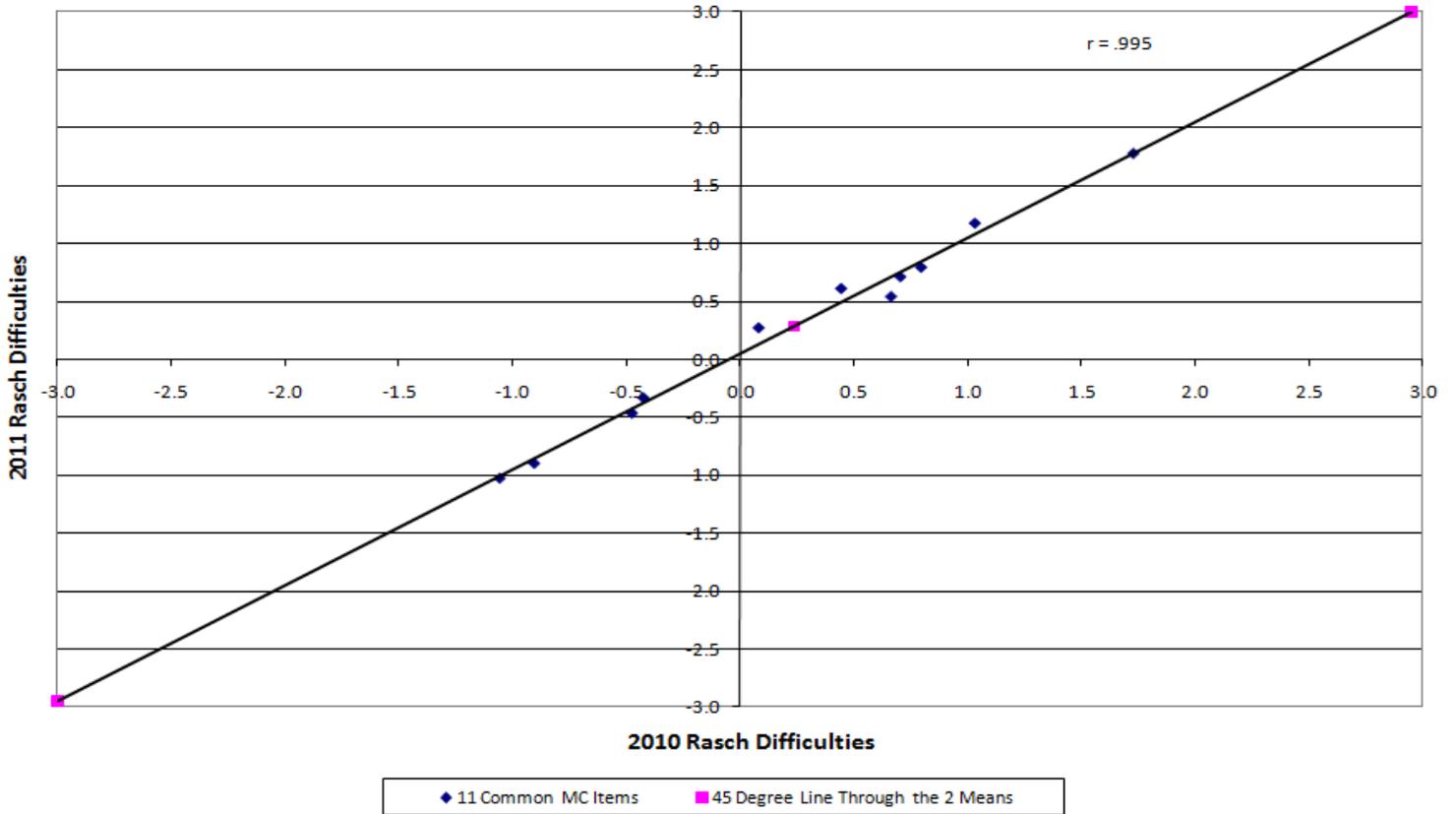
MI-ACCESS ELA, Mathematics, and Science Addendum | 2010-2011
Fall 2010 Accessing Print Grade 7 Anchor Items



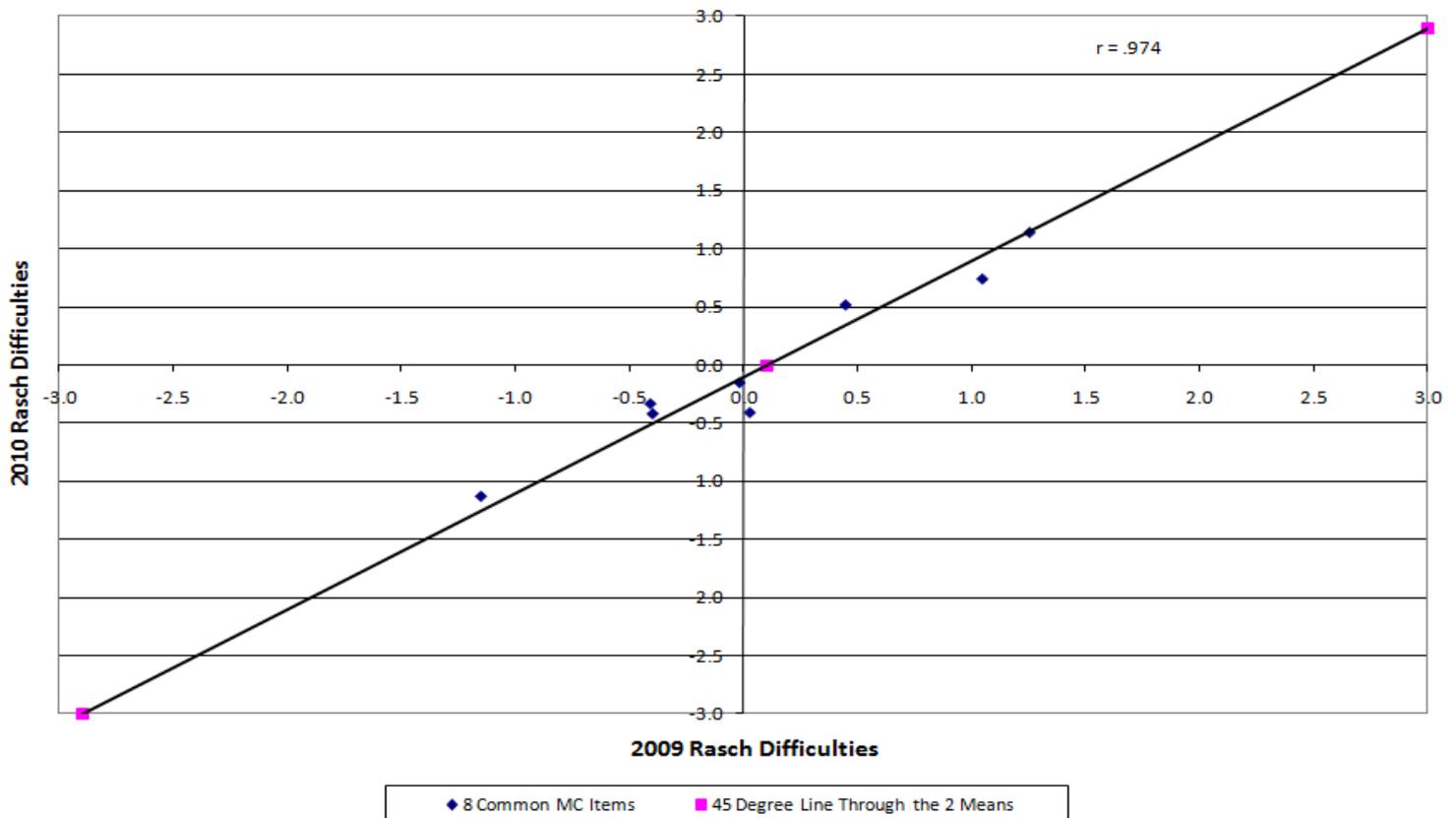
Fall 2010 Accessing Print Grade 8 Anchor Items



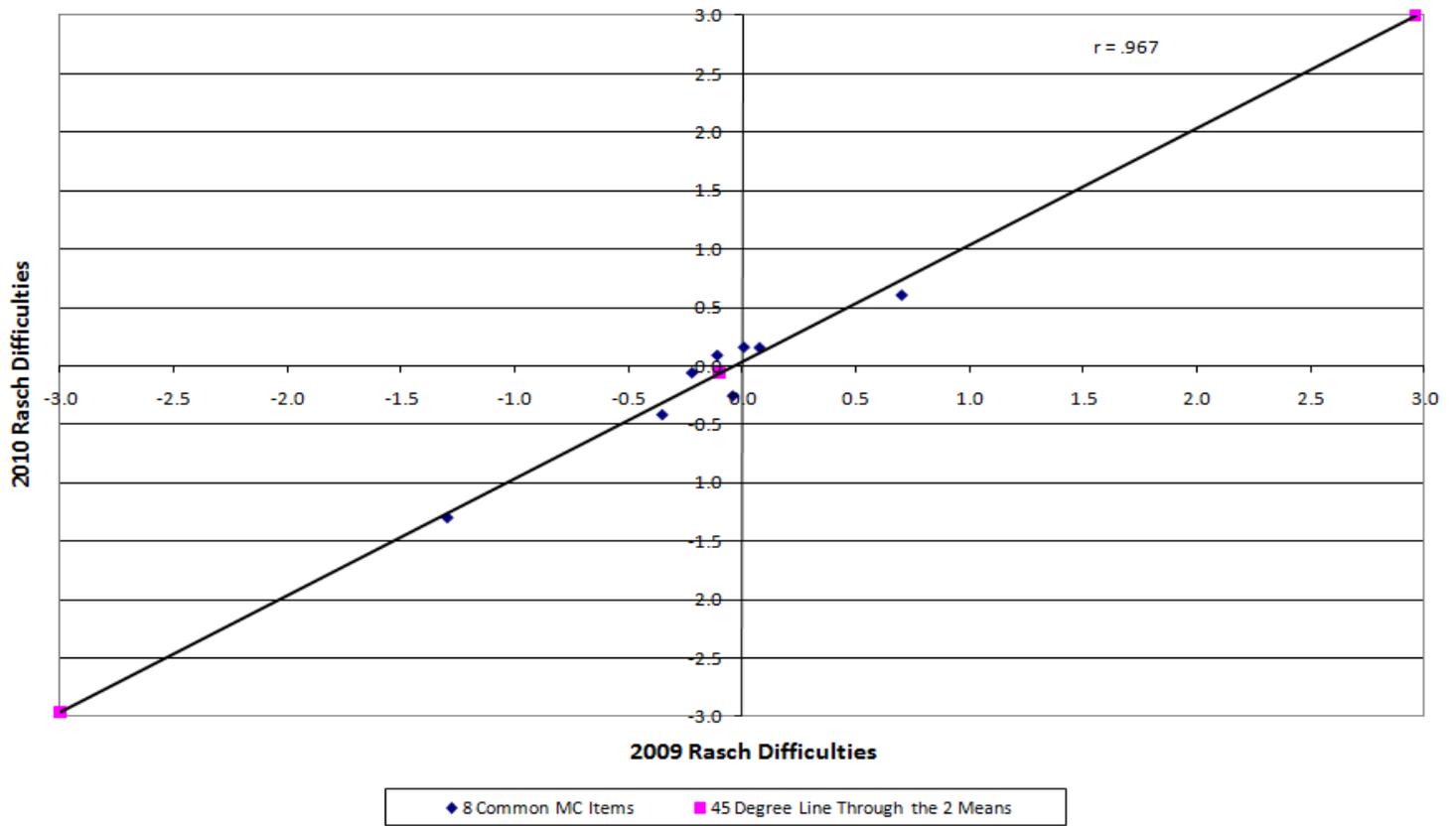
MI-ACCESS ELA, Mathematics, and Science Addendum | 2010-2011
Spring 2011 Accessing Print Grade 11 Anchor Items



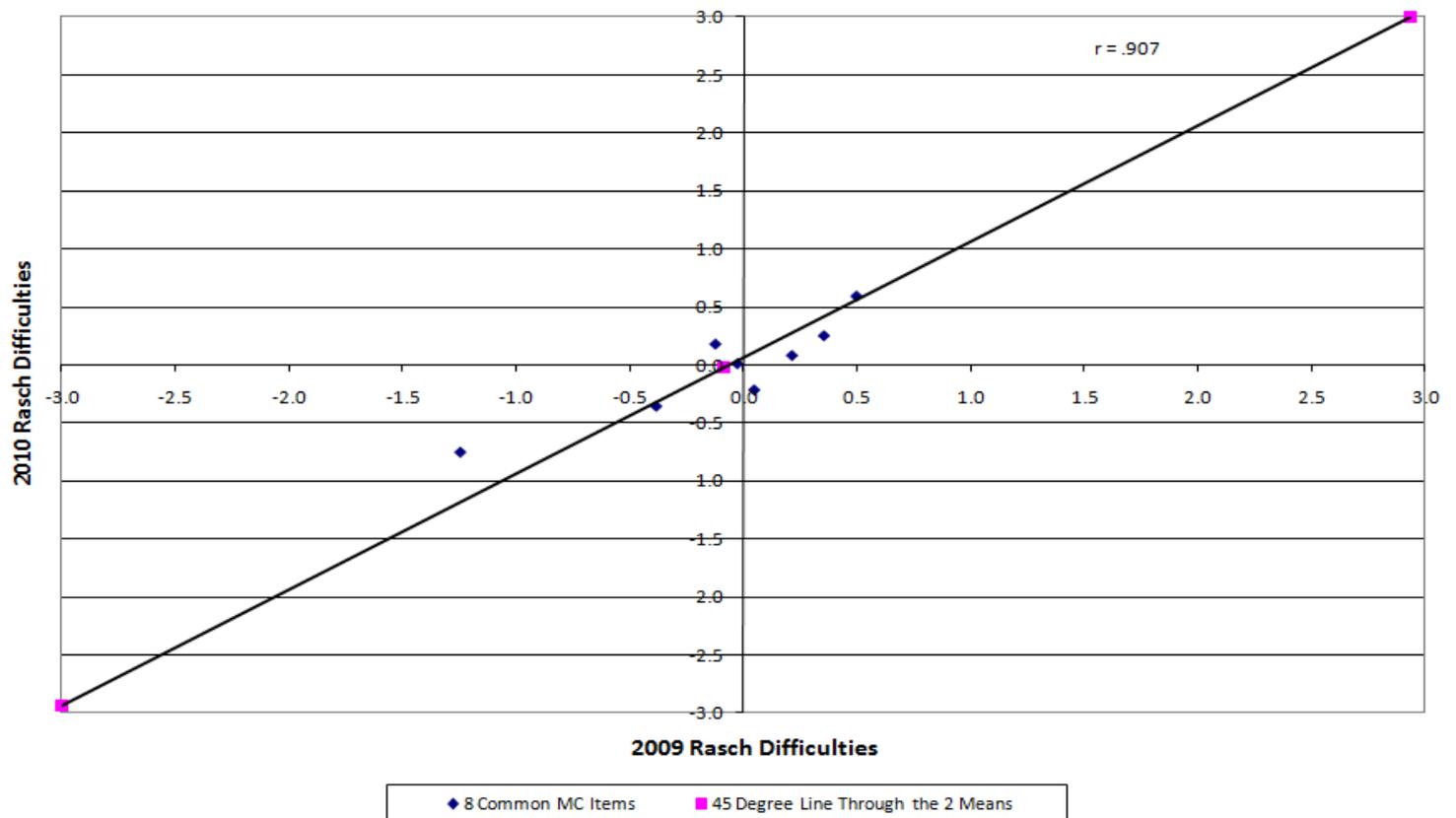
Fall 2010 Mathematics Grade 3 Anchor Items



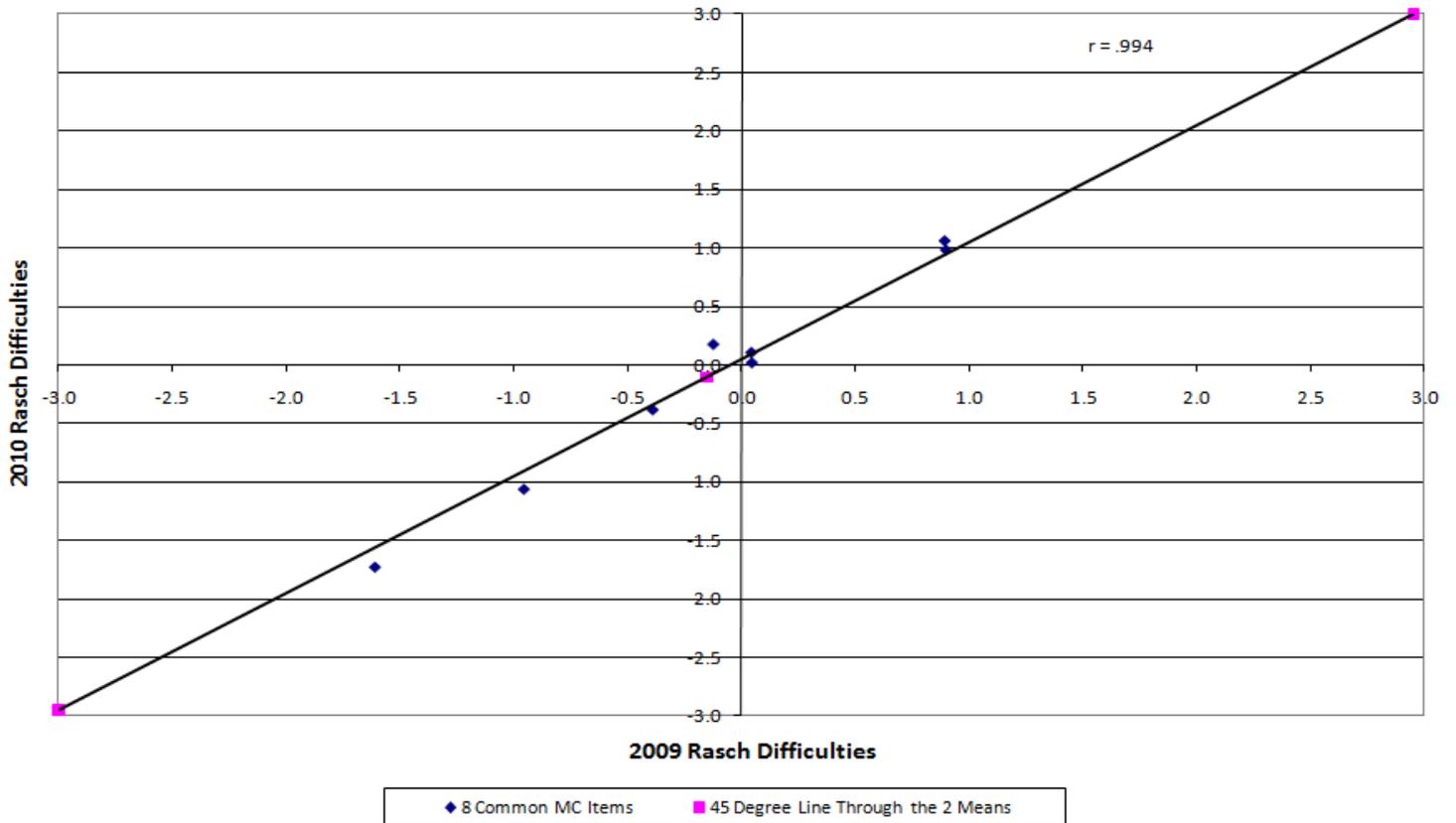
Fall 2010 Mathematics Grade 4 Anchor Items



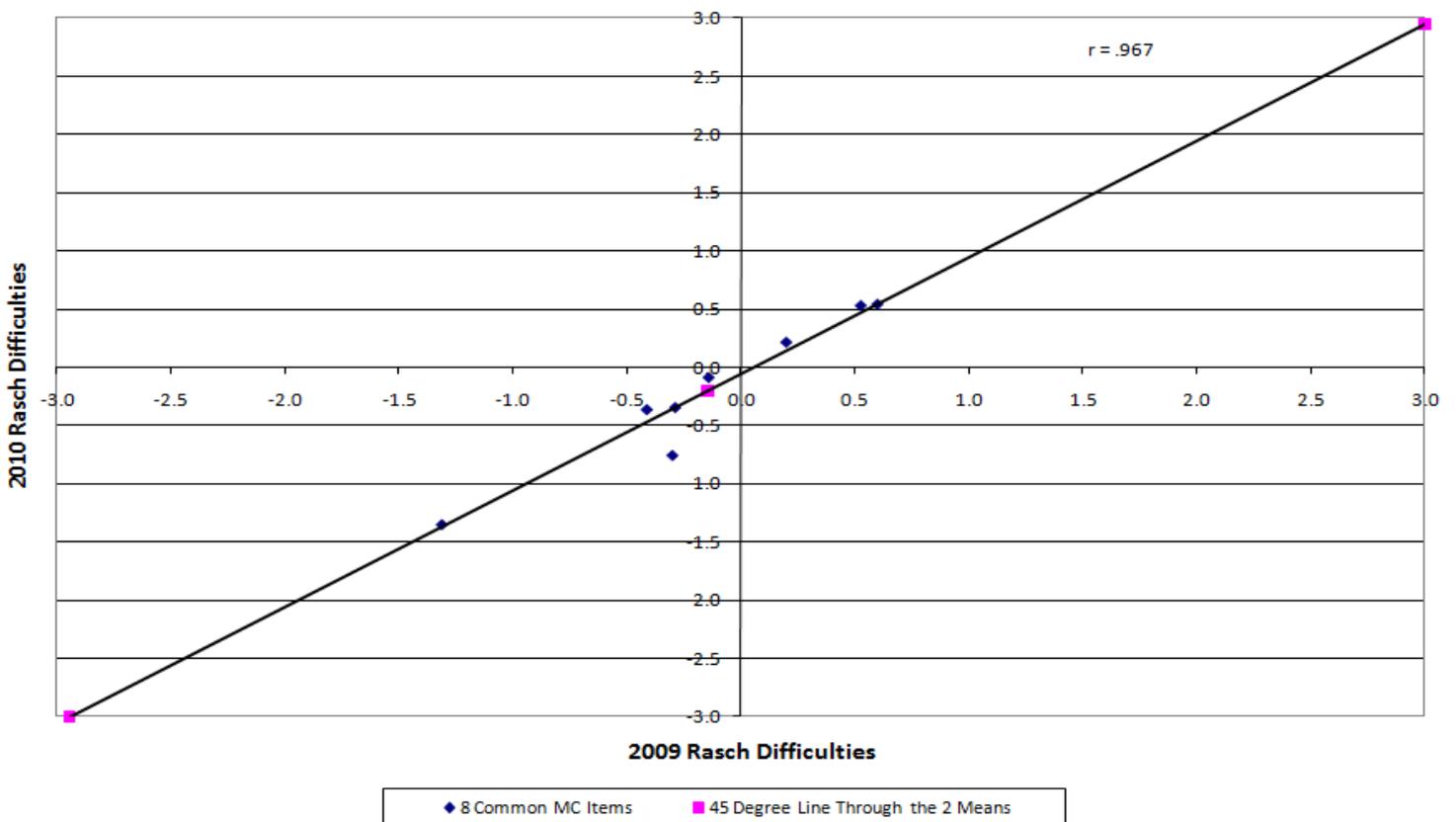
Fall 2010 Mathematics Grade 5 Anchor Items



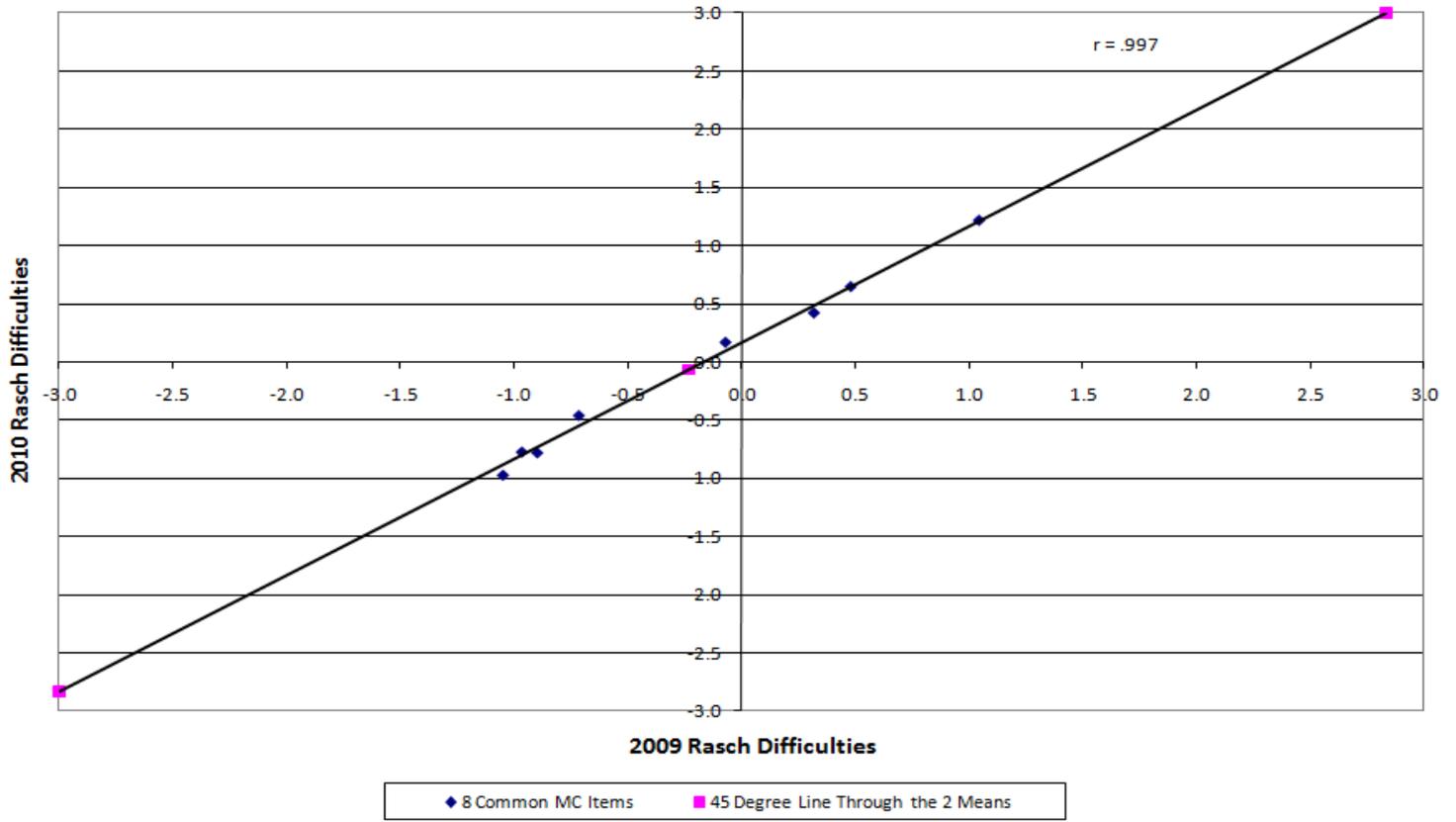
Fall 2010 Mathematics Grade 6 Anchor Items



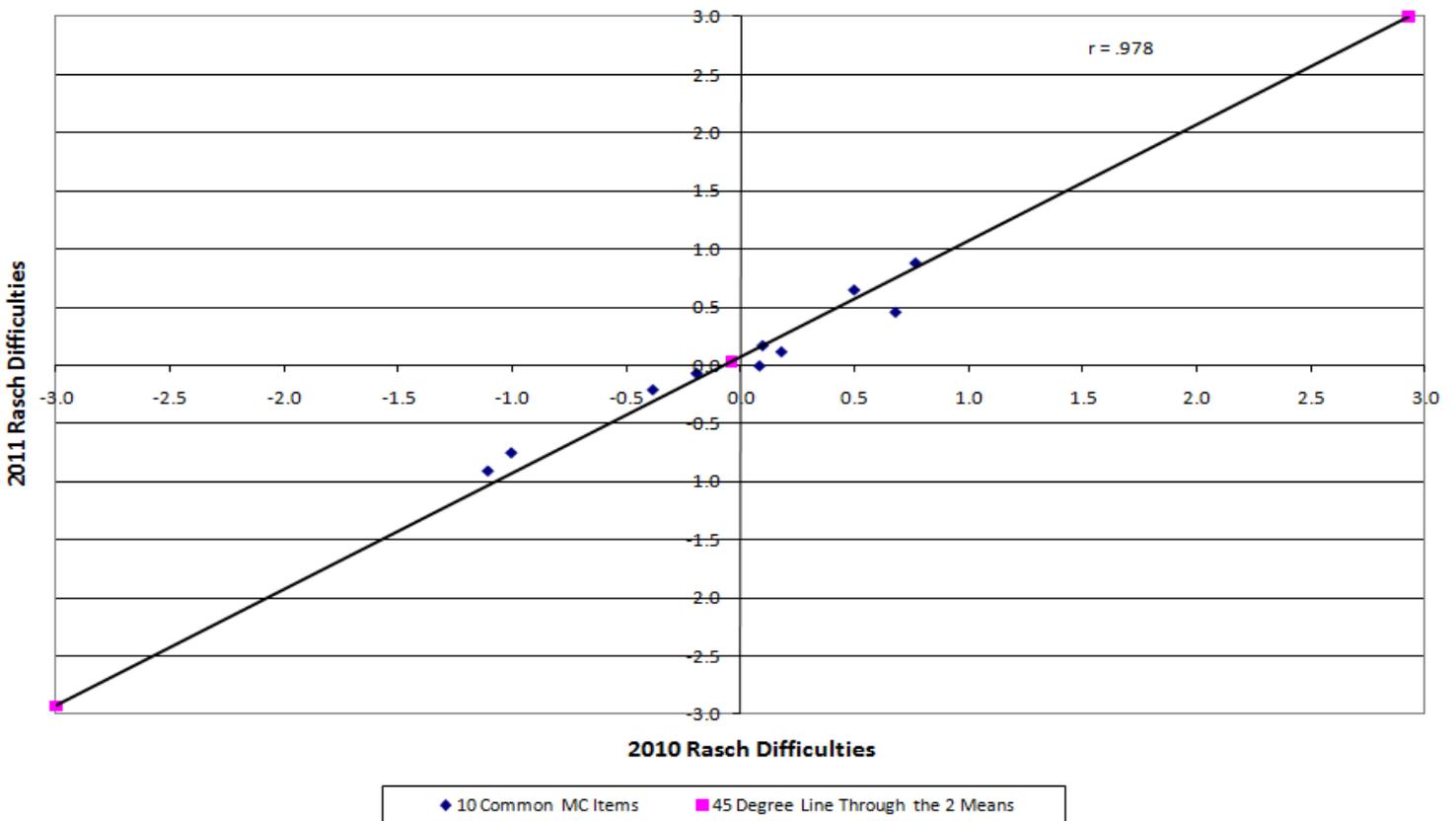
Fall 2010 Mathematics Grade 7 Anchor Items



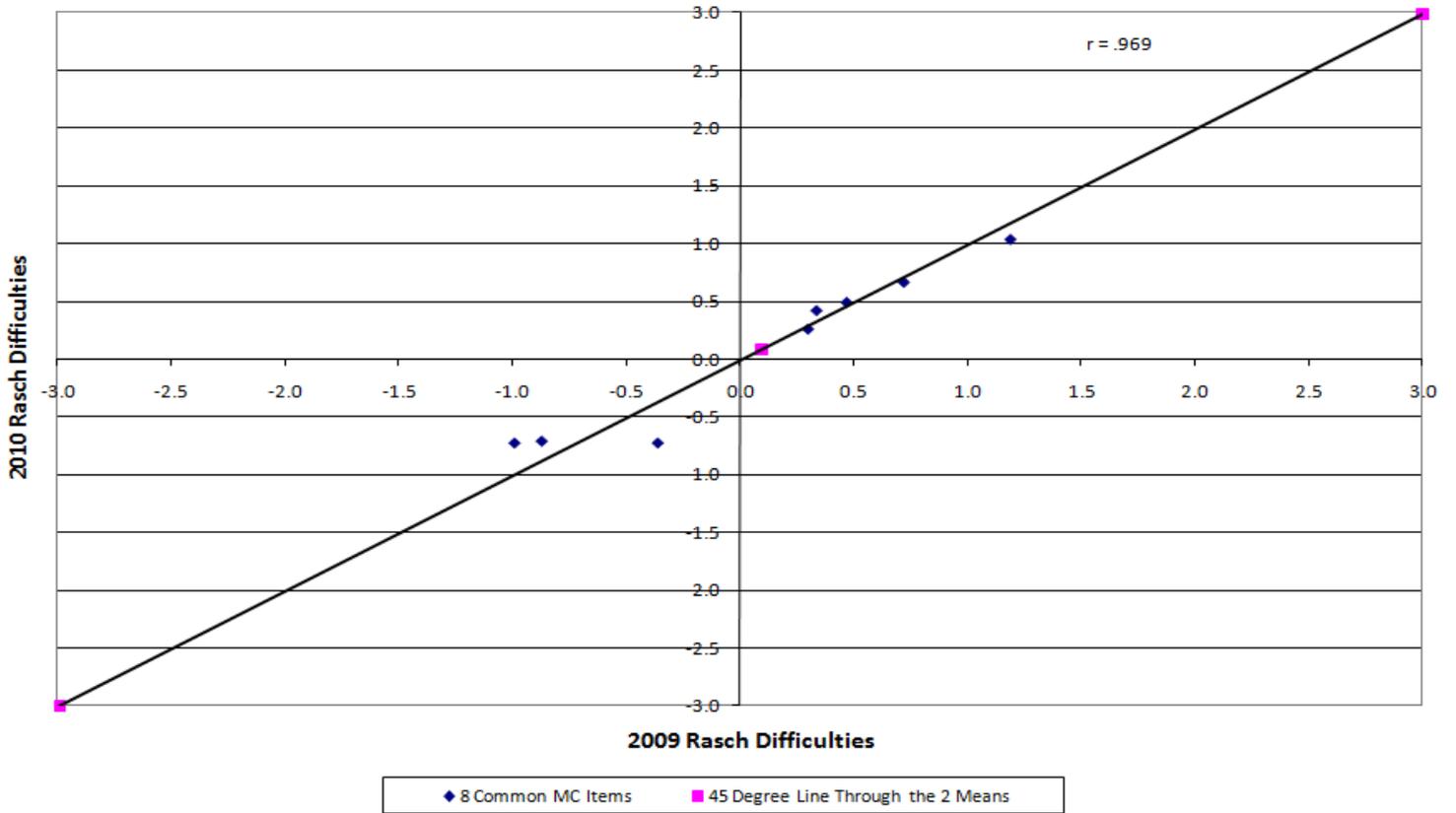
Fall 2010 Mathematics Grade 8 Anchor Items



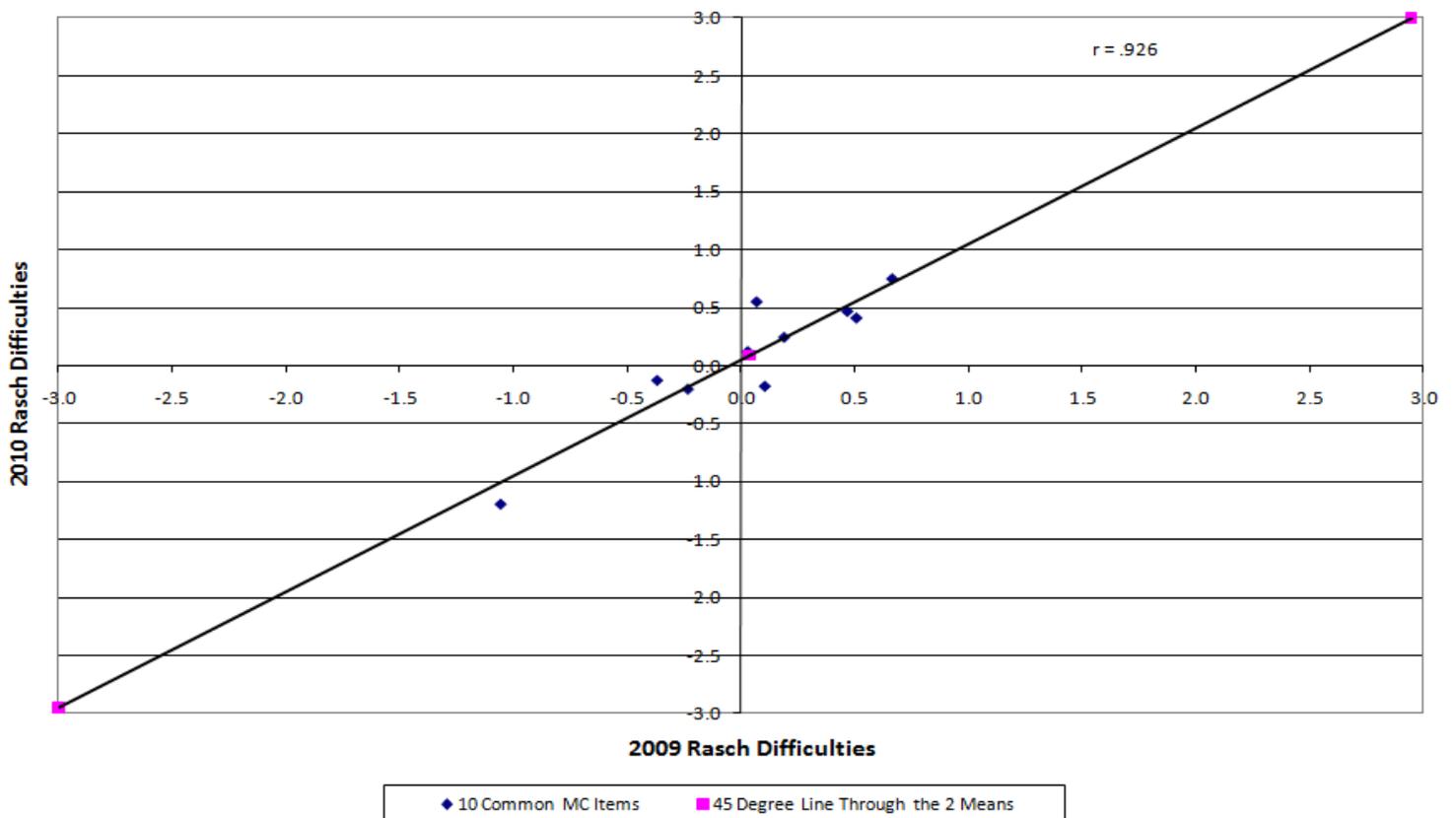
Spring 2011 Mathematics Grade 11 Anchor Items



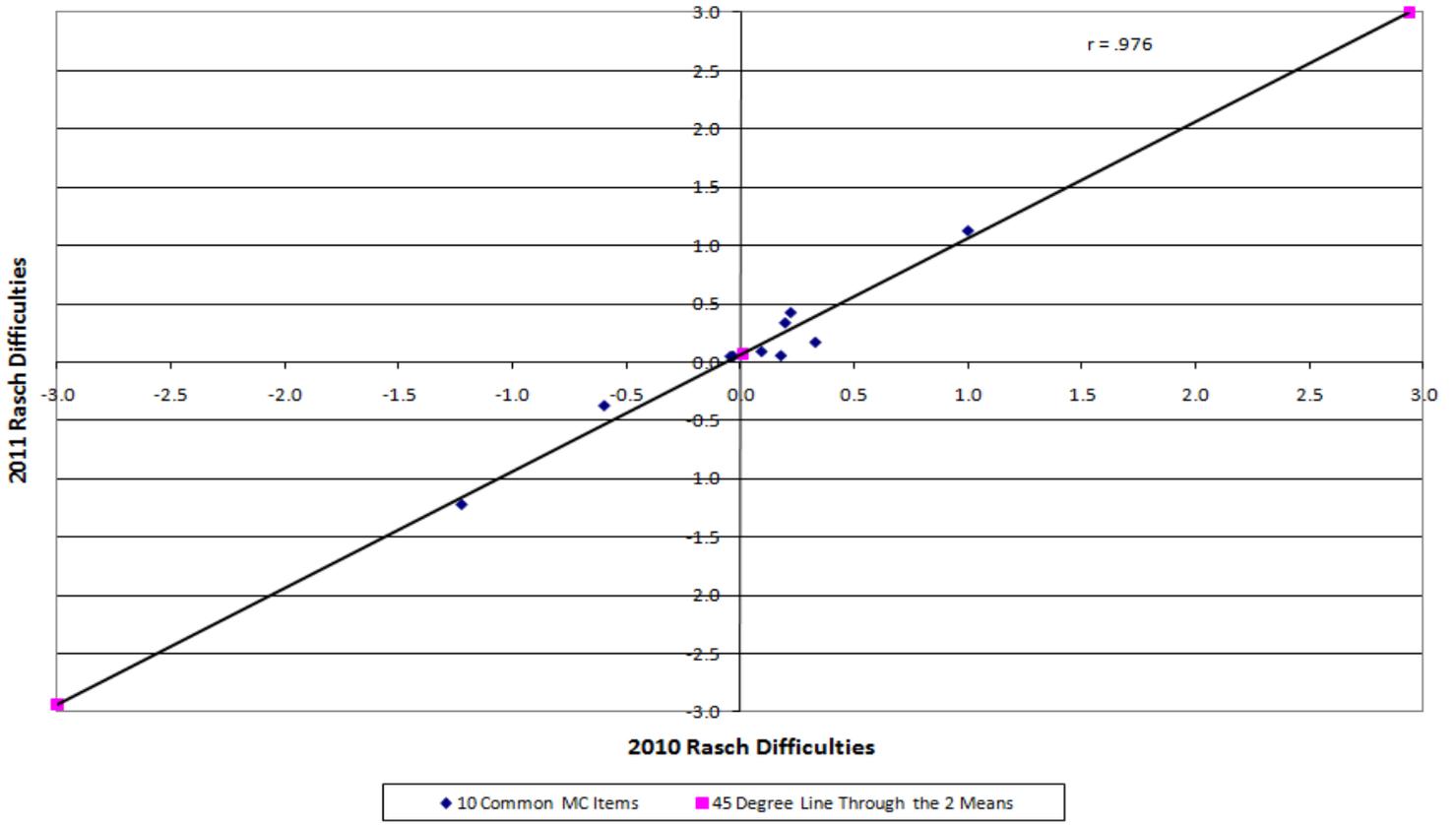
Fall 2010 Science Grade 5 Anchor Items



Fall 2010 Science Grade 8 Anchor Items

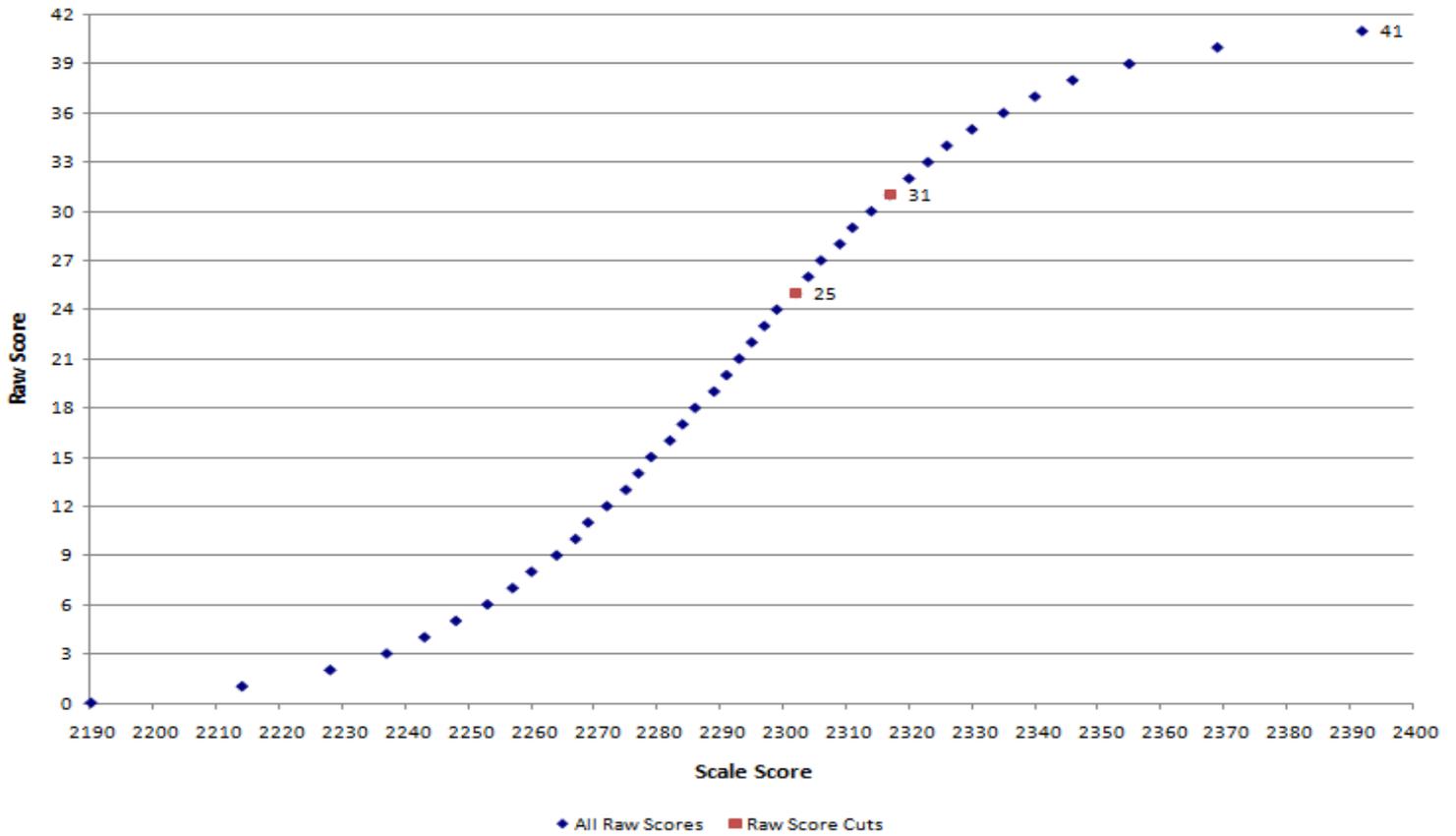


Spring 2011 Science Grade 11 Anchor Items

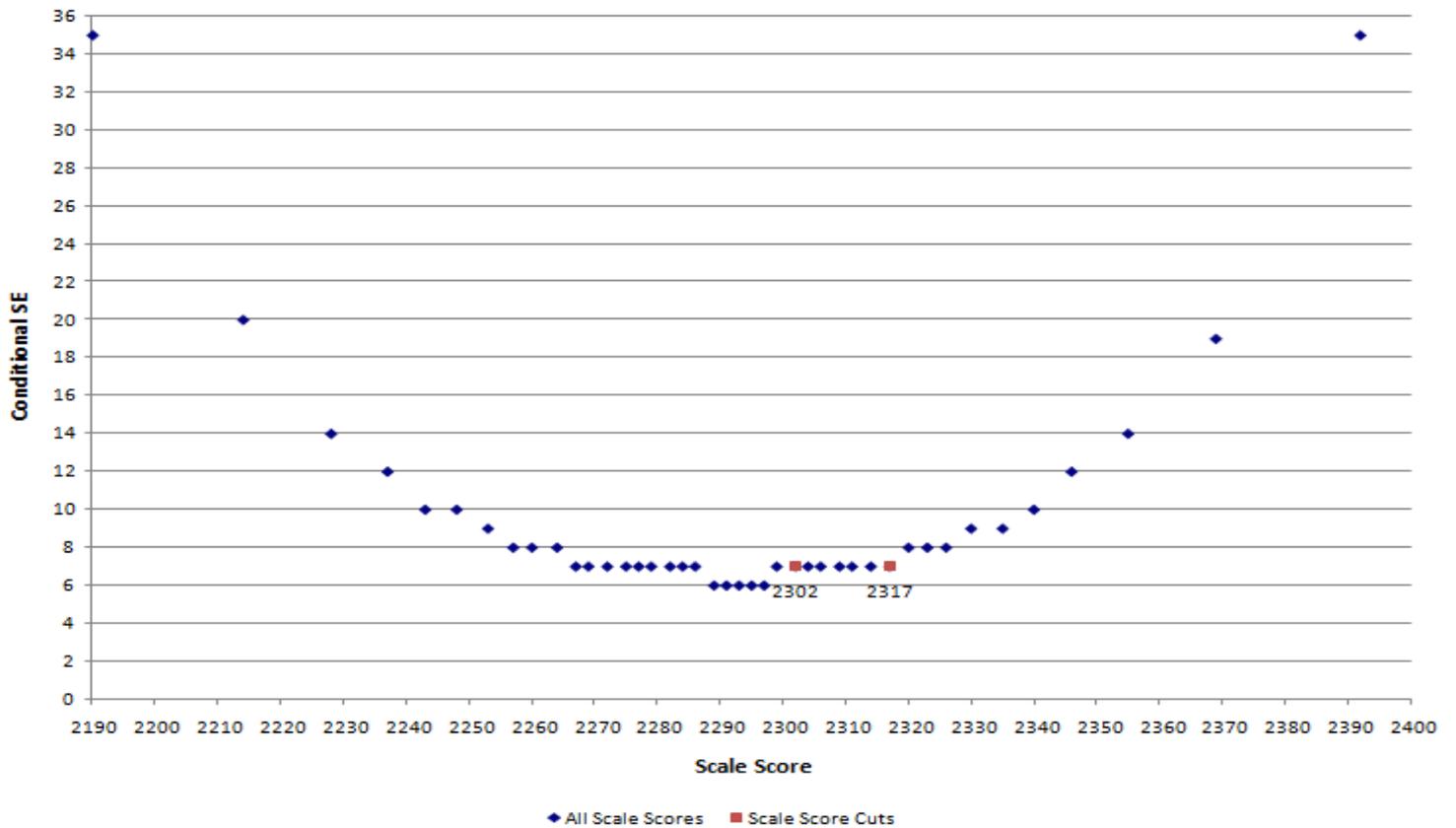


**APPENDIX B:
TEST CHARACTERISTIC CURVES AND STANDARD ERROR CURVES**

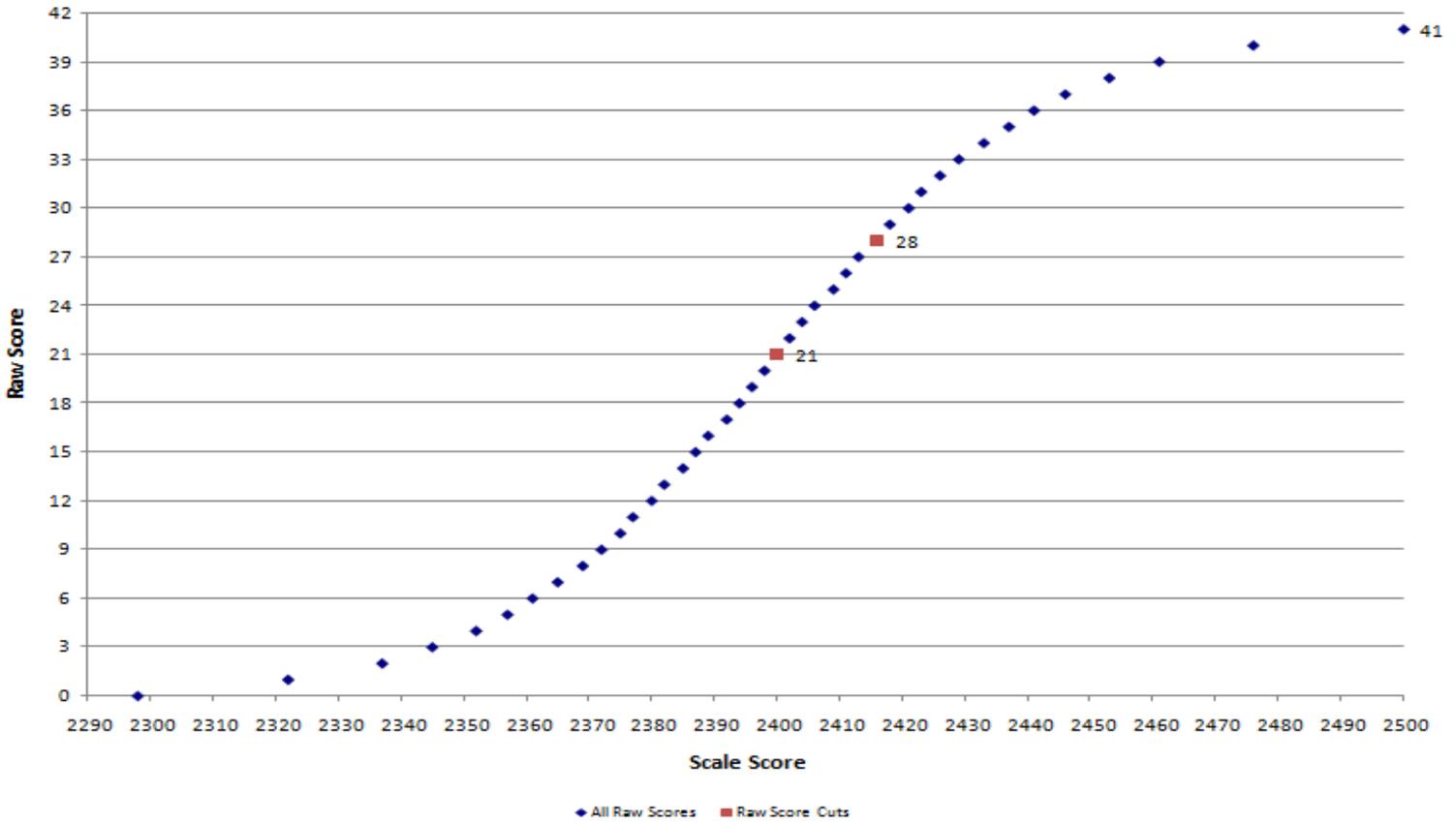
Fall 2010 Accessing Print Grade 3 Test Characteristic Curve



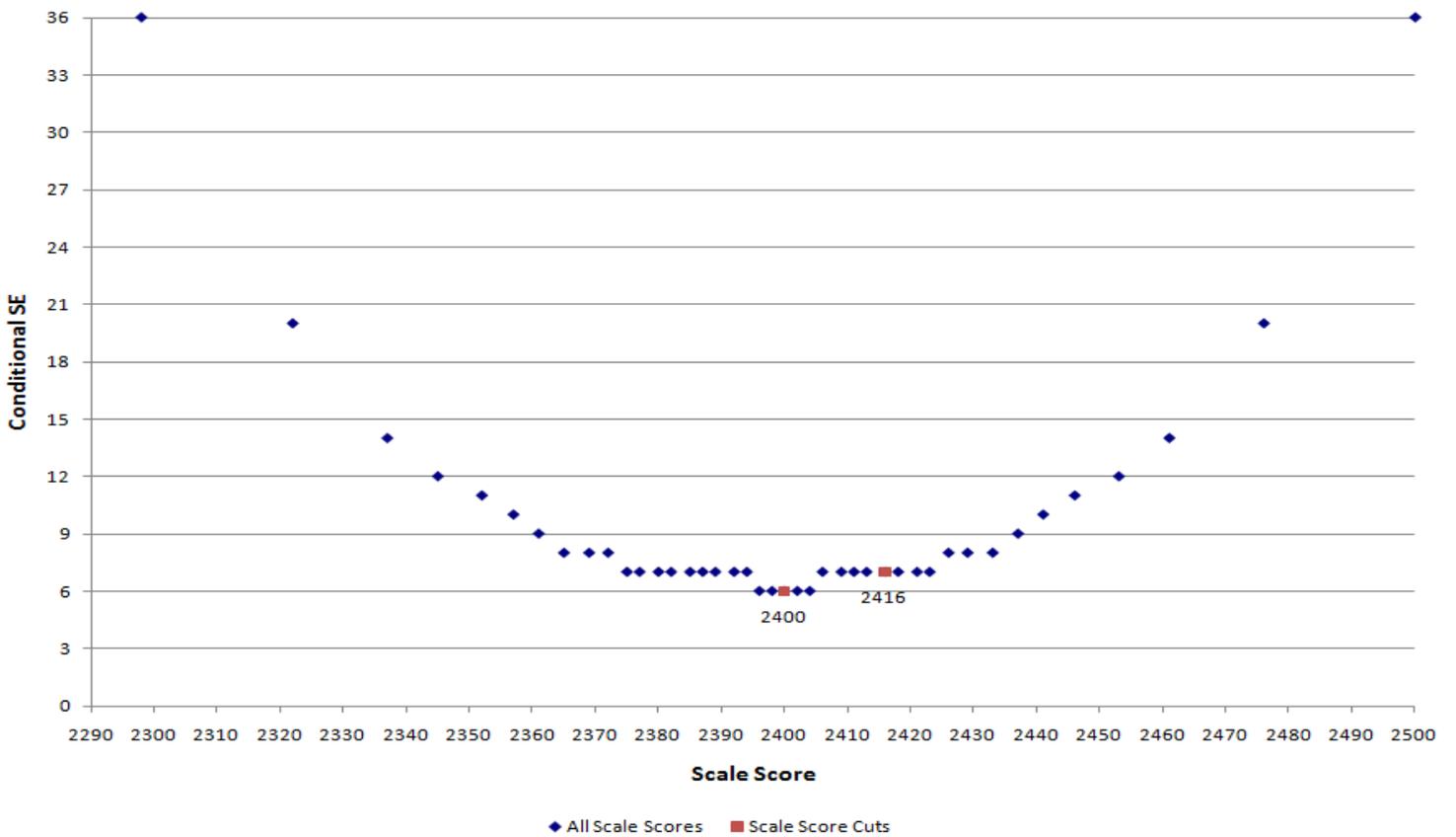
Fall 2010 Accessing Print Grade 3 Standard Error Curve



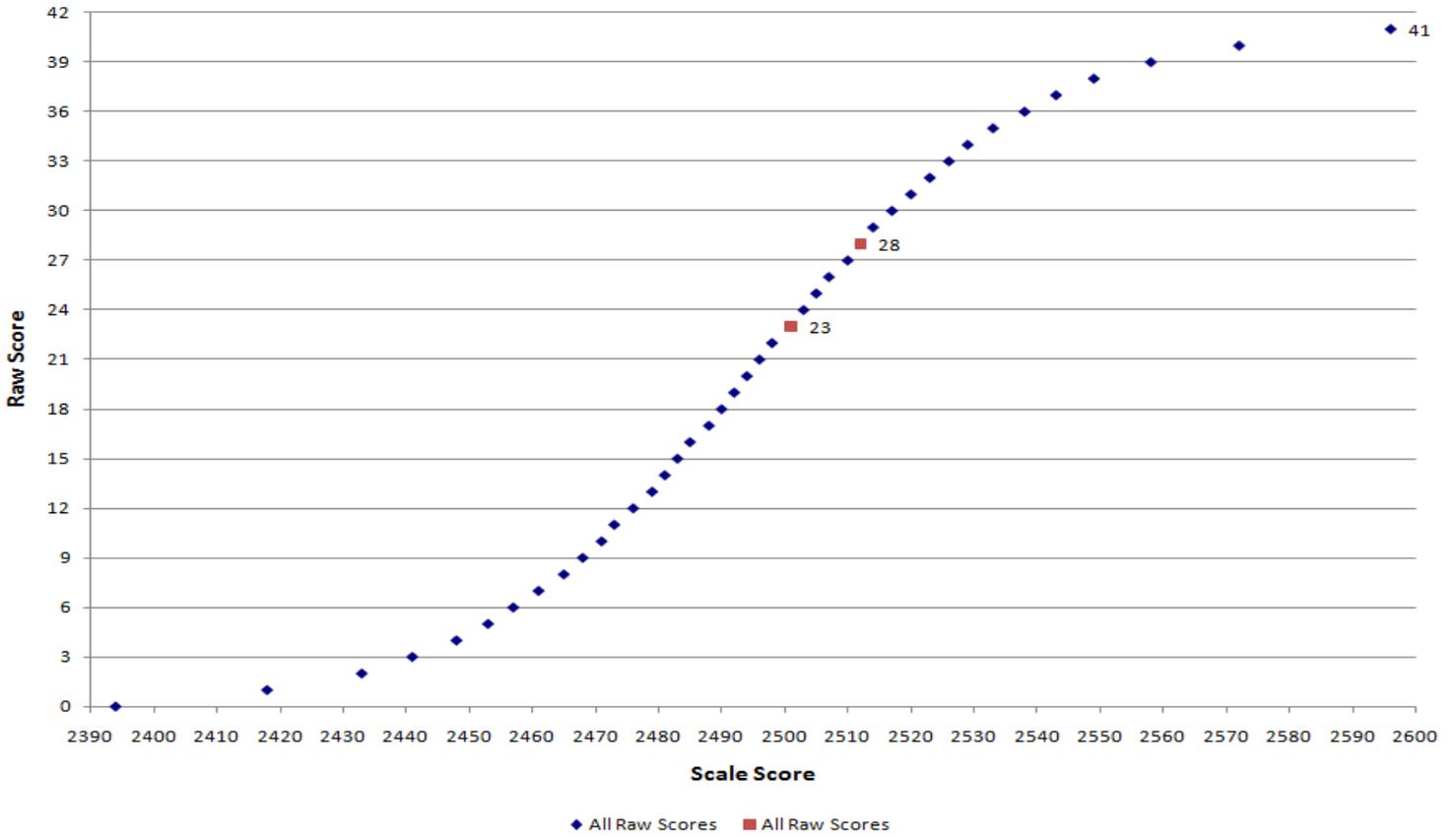
Fall 2010 Accessing Print Grade 4 Test Characteristic Curve



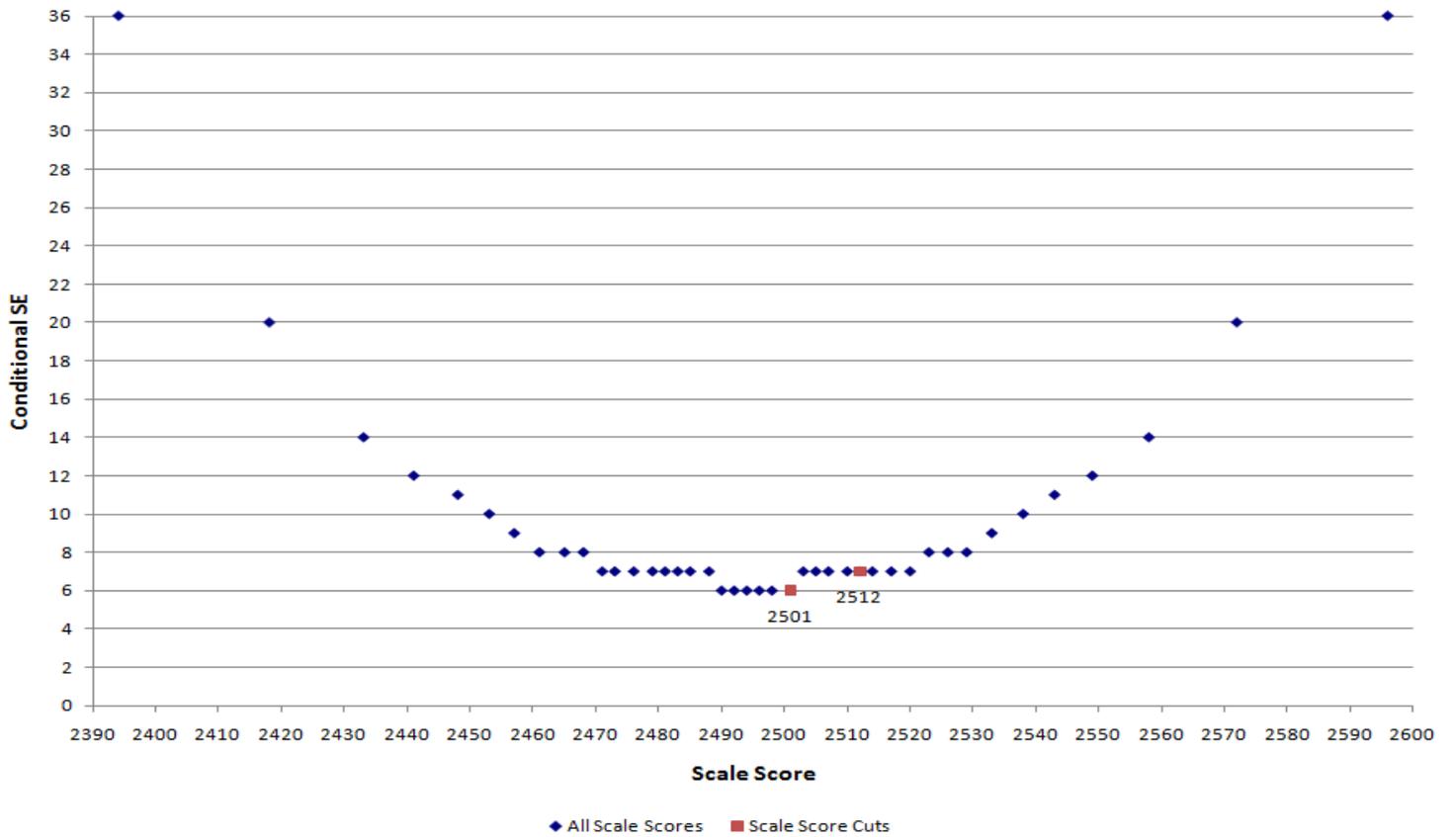
Fall 2010 Accessing Print Grade 4 Standard Error Curve



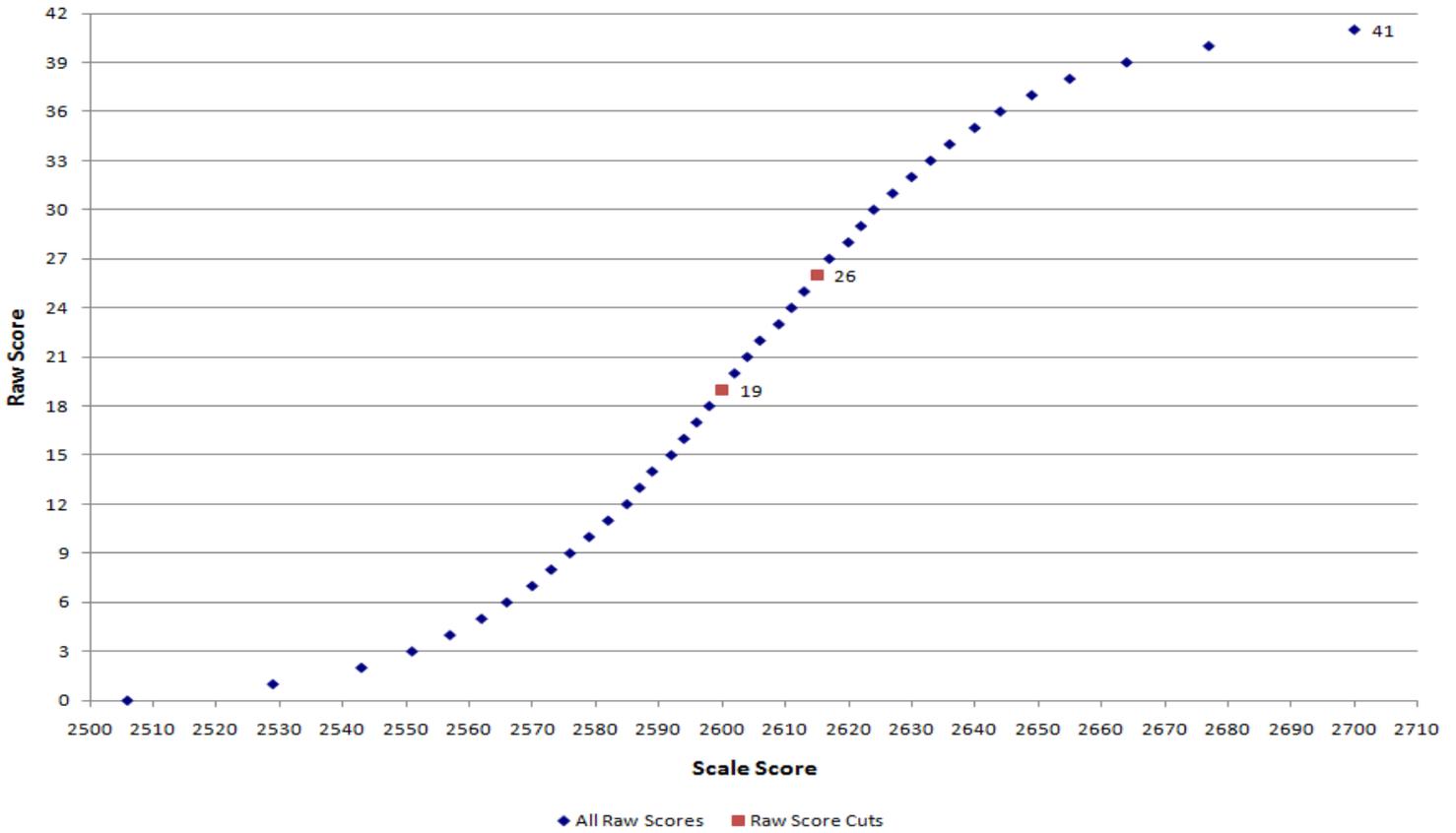
Fall 2010 Accessing Print Grade 5 Test Characteristic Curve



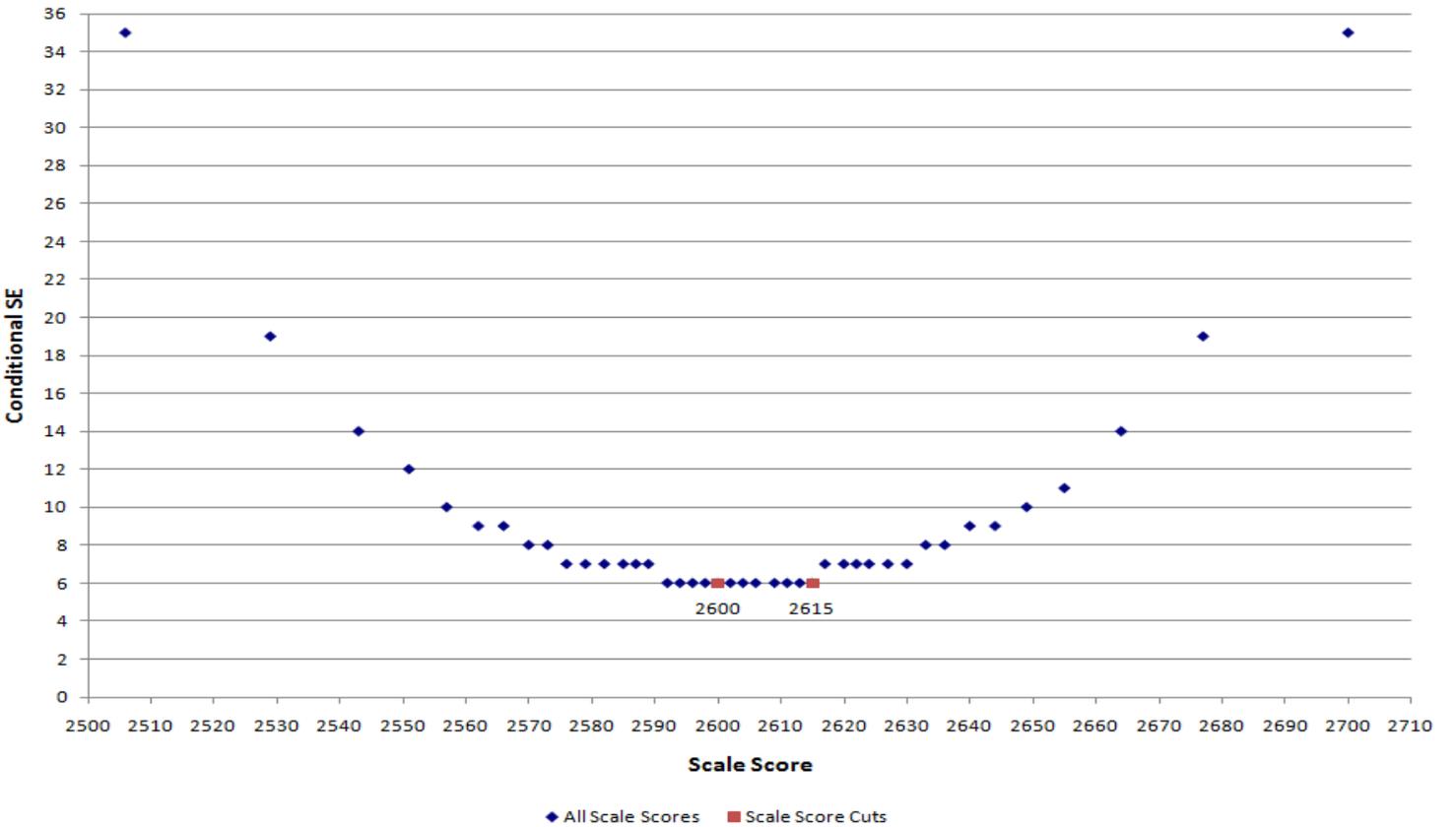
Fall 2010 Accessing Print Grade 5 Standard Error Curve



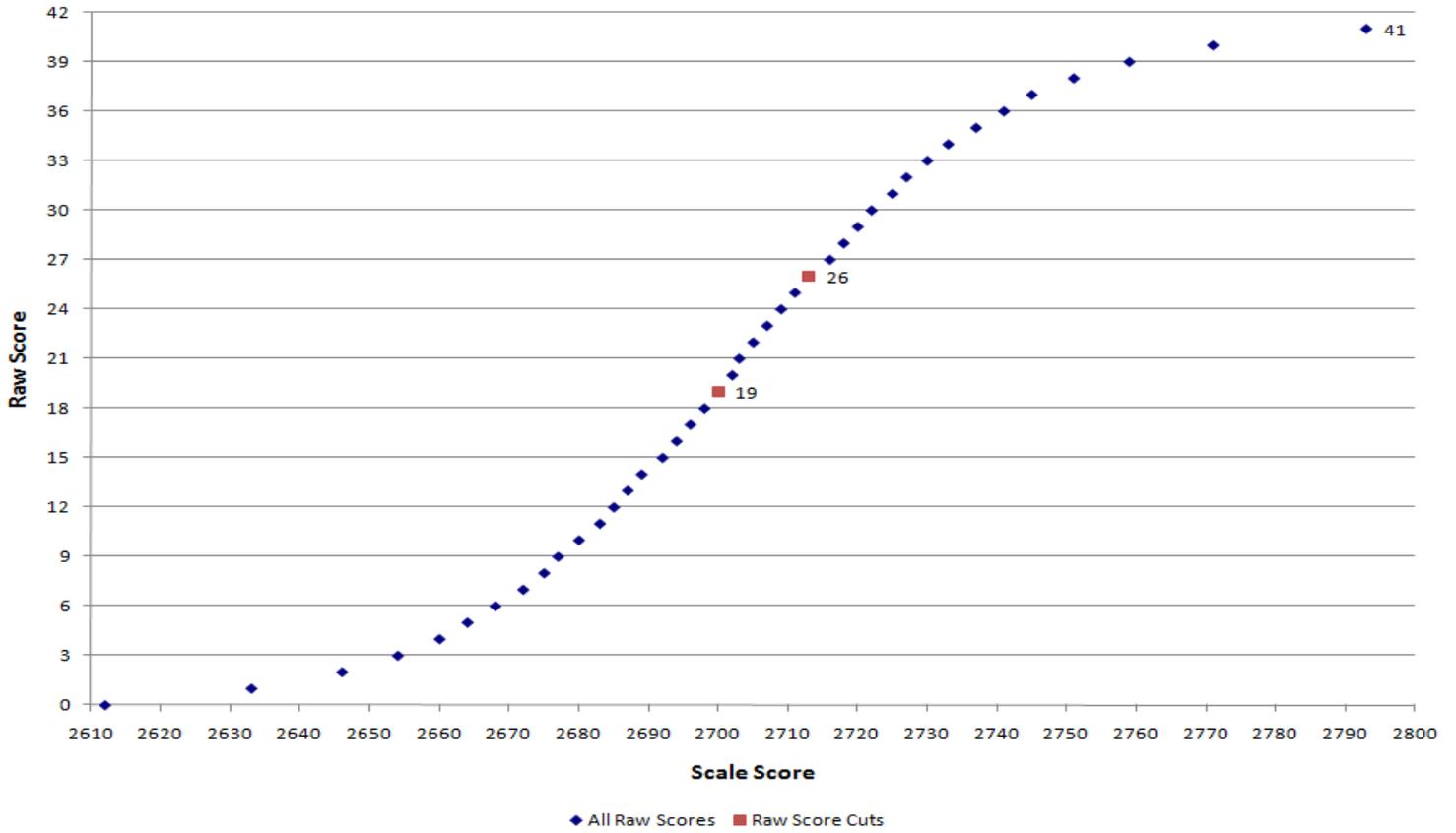
Fall 2010 Accessing Print Grade 6 Test Characteristic Curve



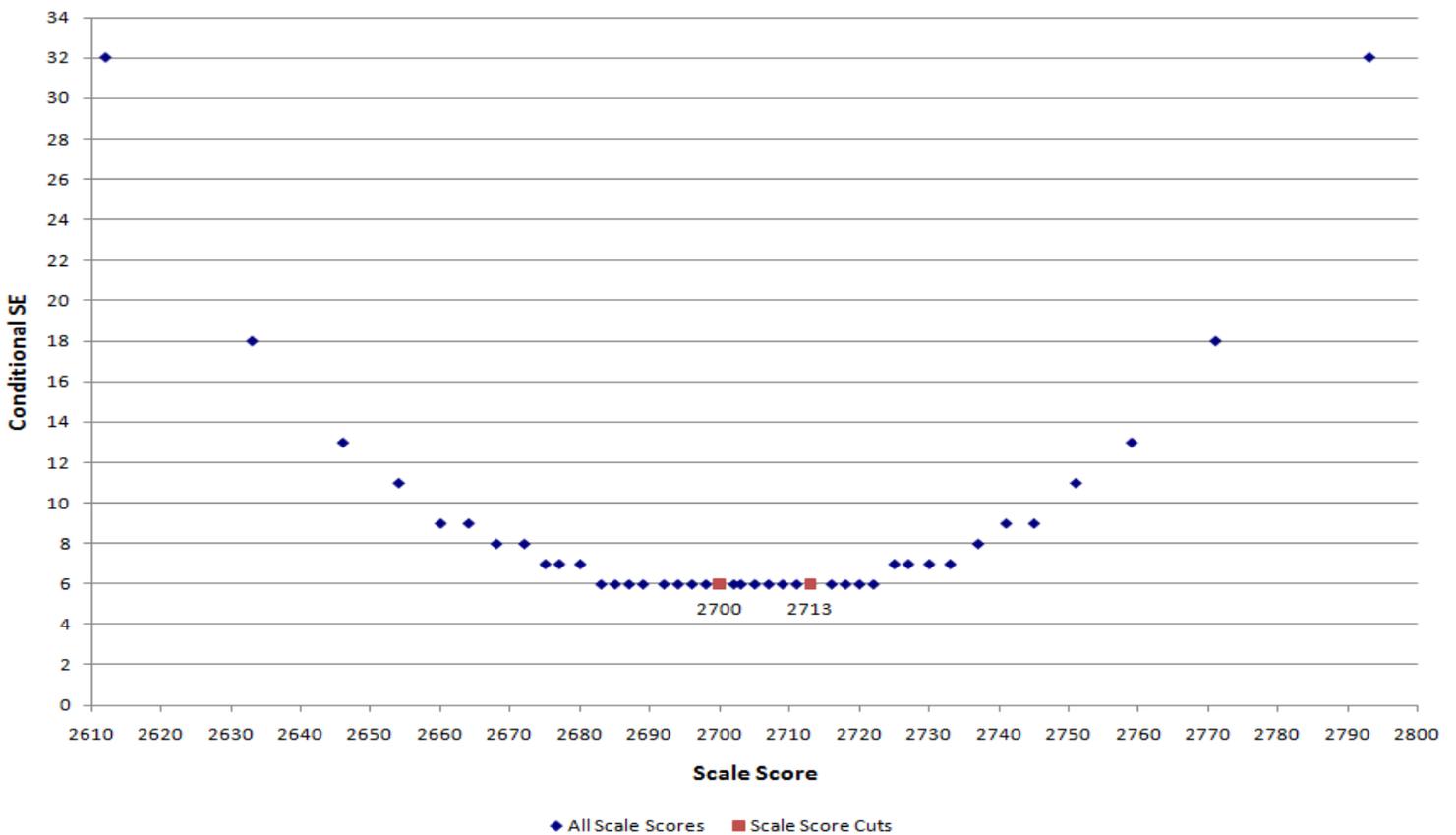
Fall 2010 Accessing Print Grade 6 Standard Error Curve



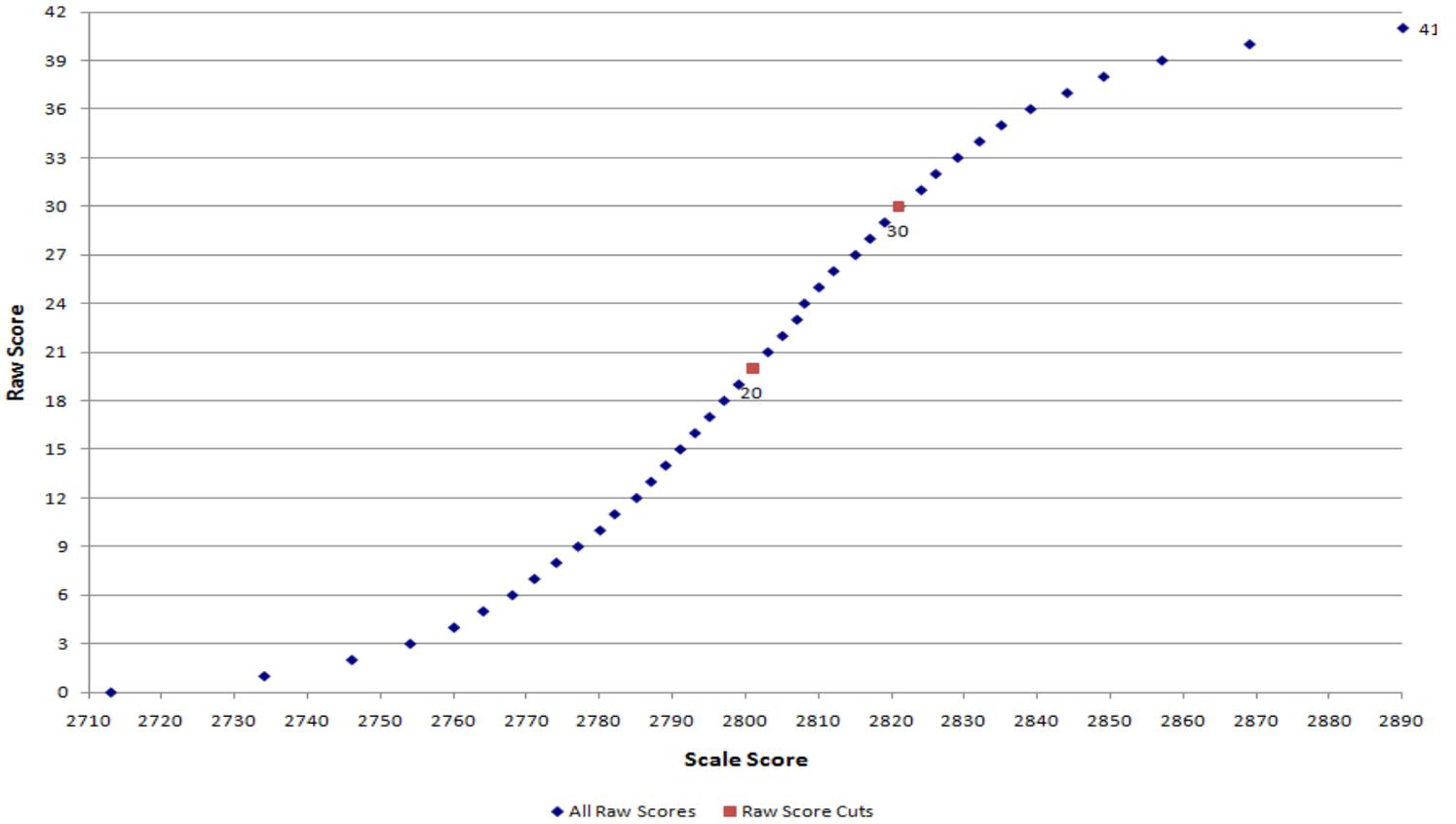
Fall 2010 Accessing Print Grade 7 Test Characteristic Curve



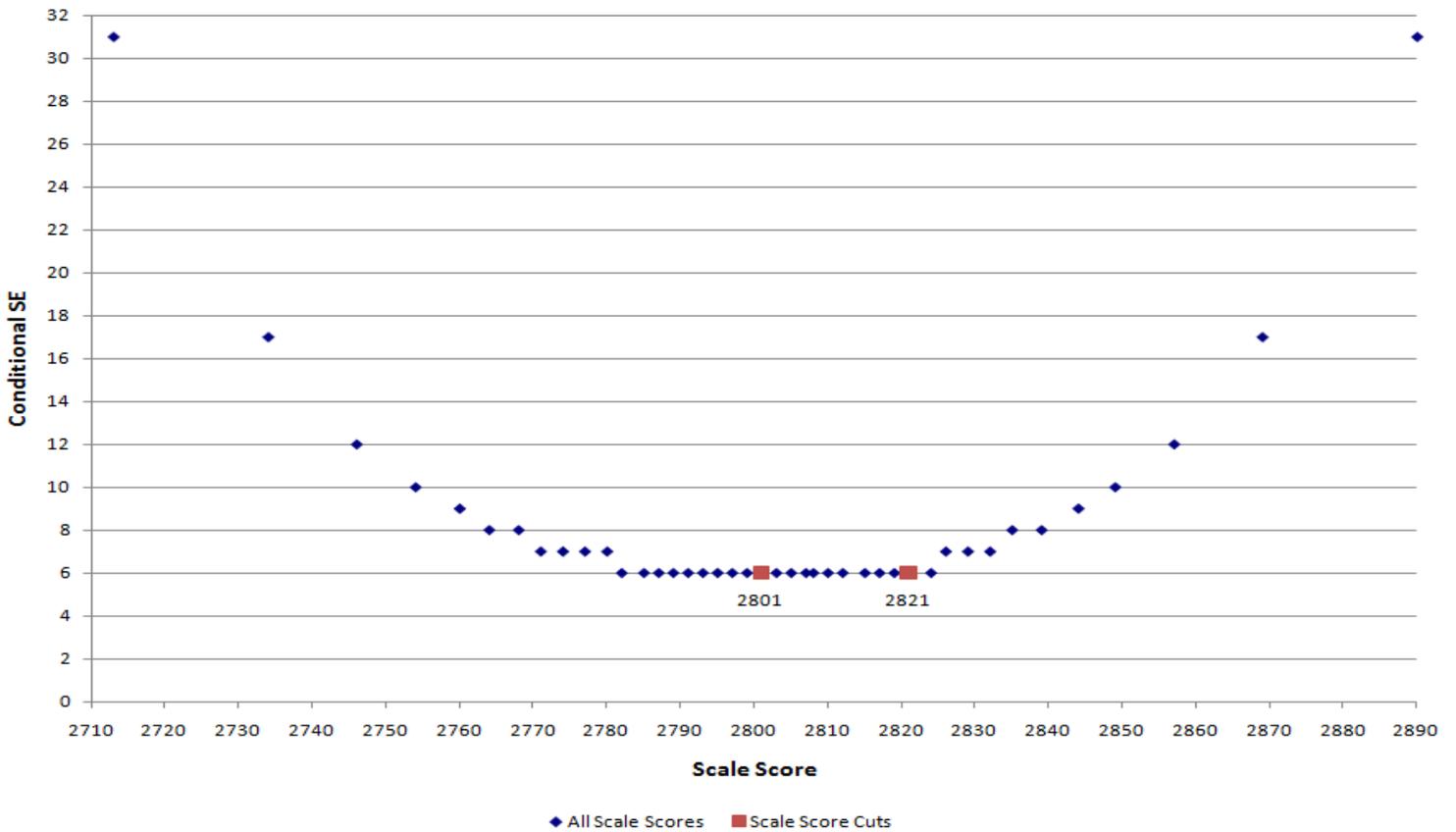
Fall 2010 Accessing Print Grade 7 Standard Error Curve



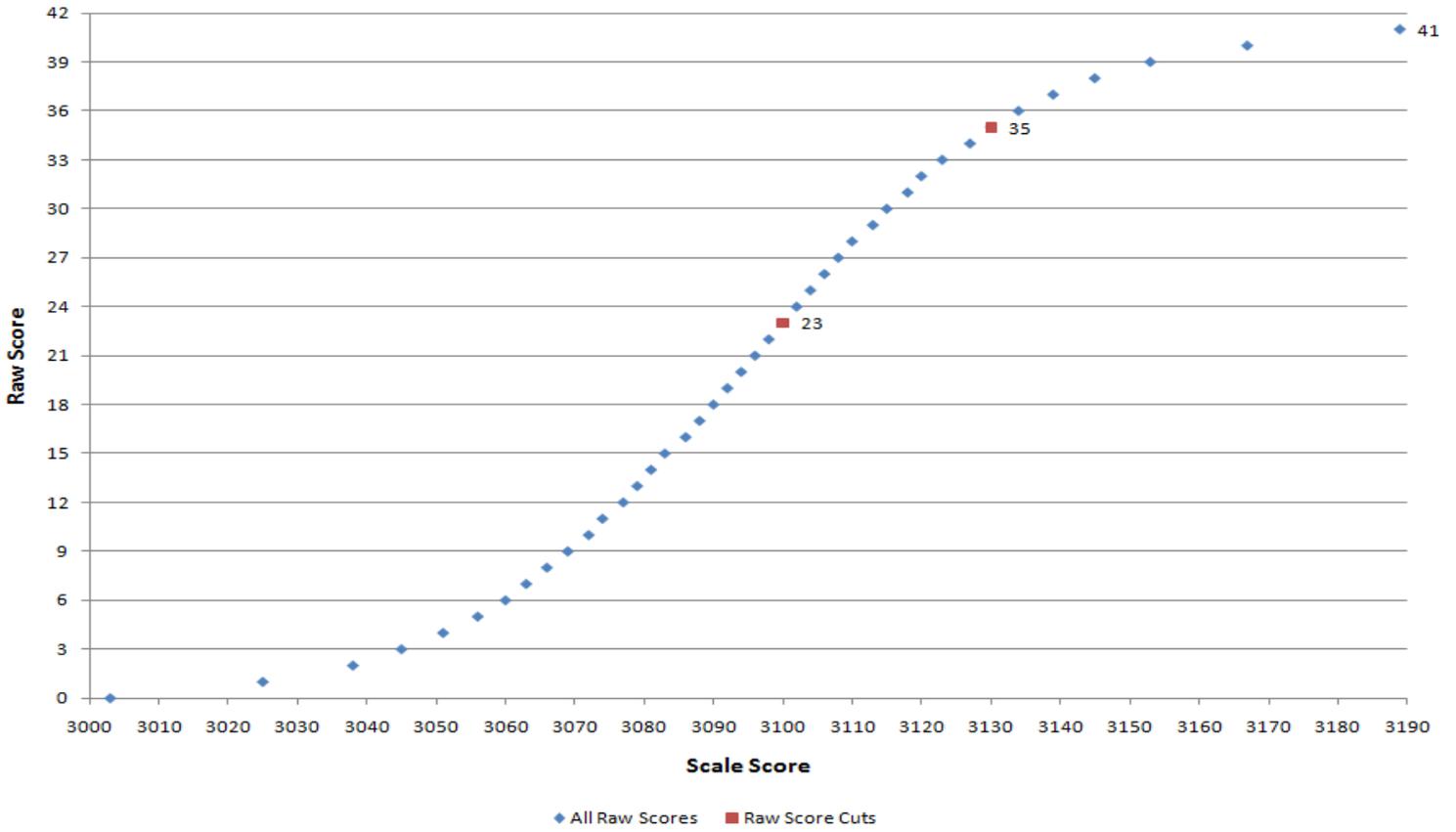
Fall 2010 Accessing Print Grade 8 Test Characteristic Curve



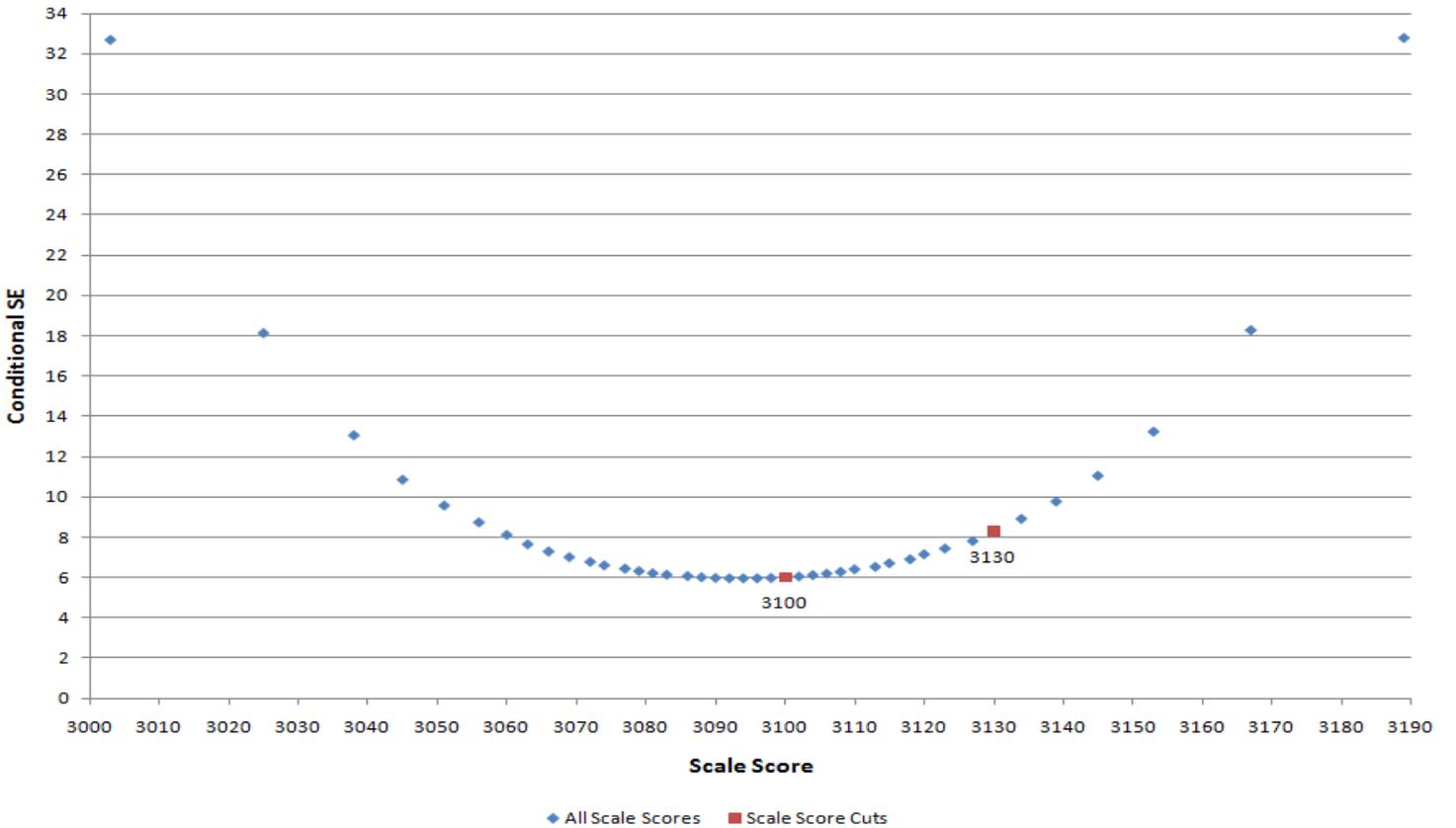
Fall 2010 Accessing Print Grade 8 Standard Error Curve



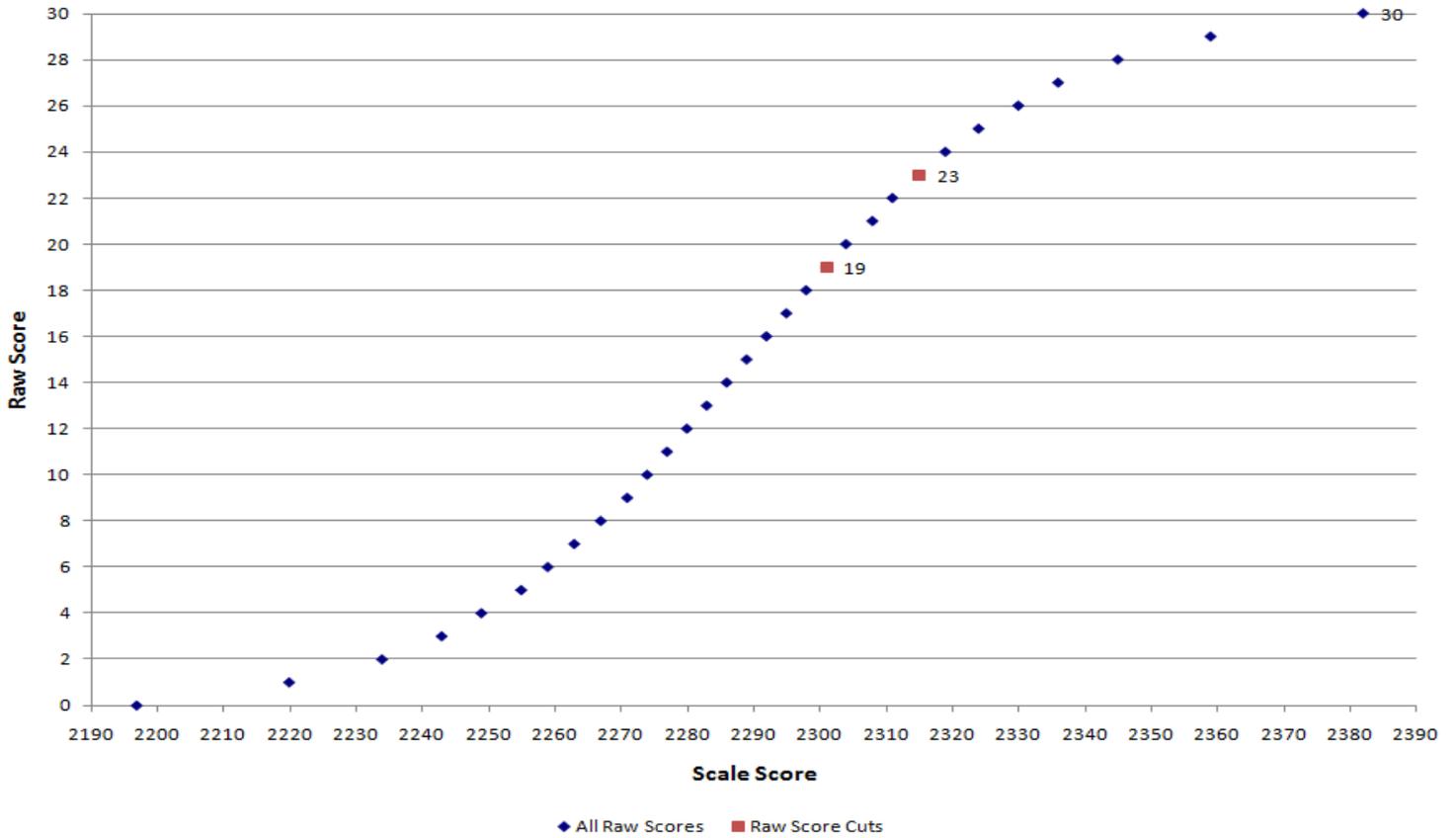
Spring 2011 Accessing Print Grade 11 Test Characteristic Curve



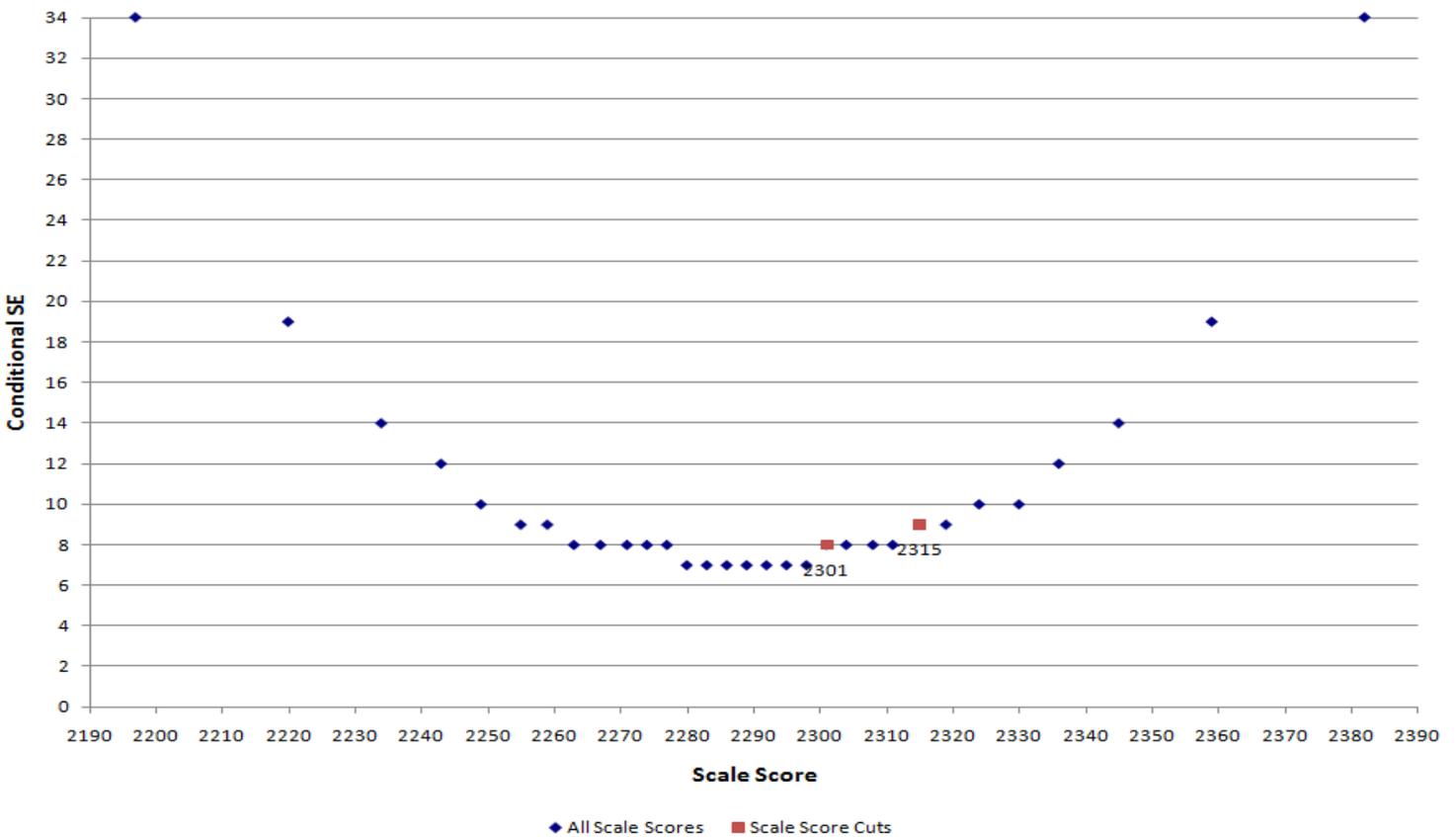
Spring 2011 Accessing Print Grade 11 Standard Error Curve



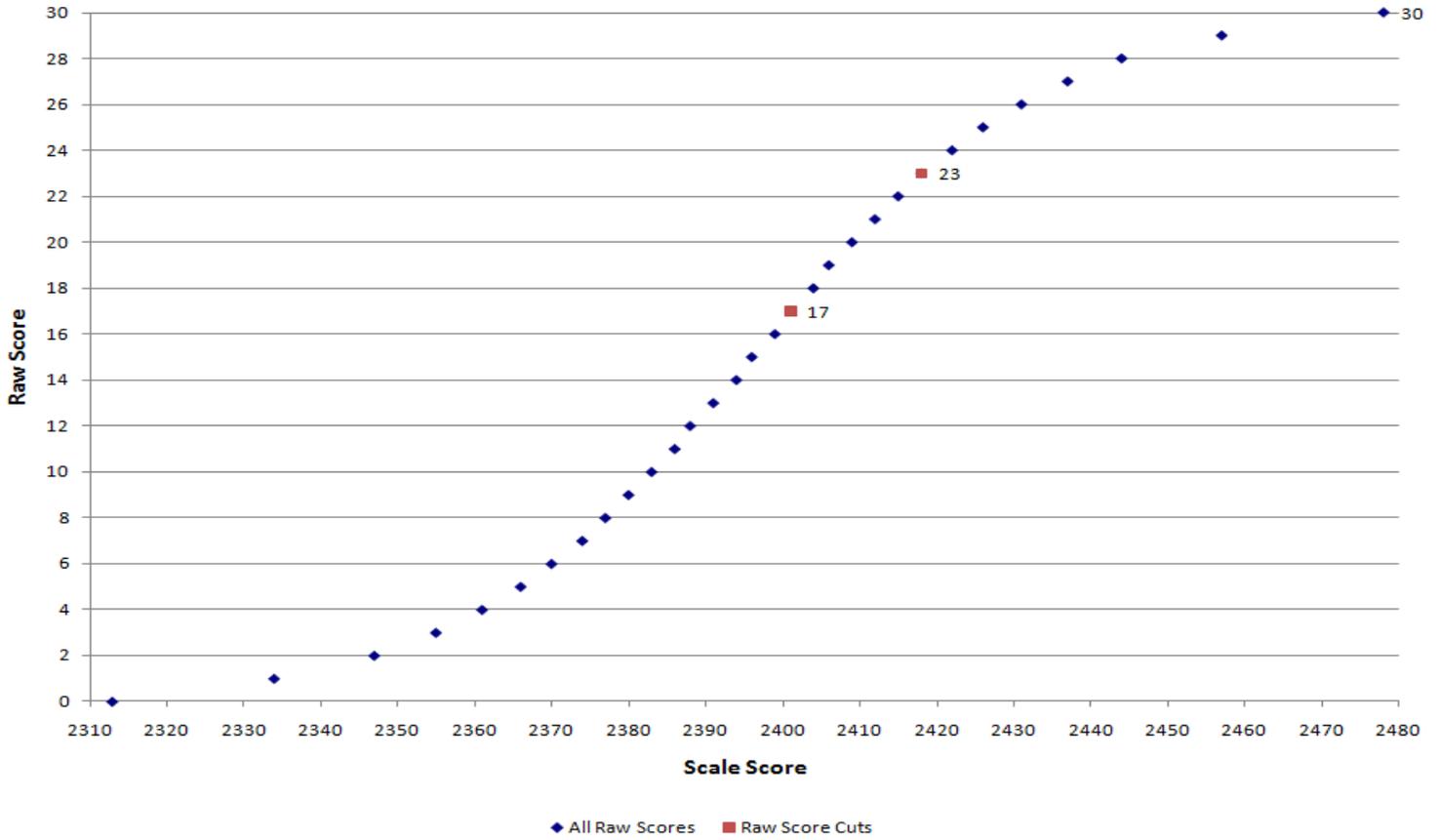
Fall 2010 Mathematics Grade 3 Test Characteristic Curve



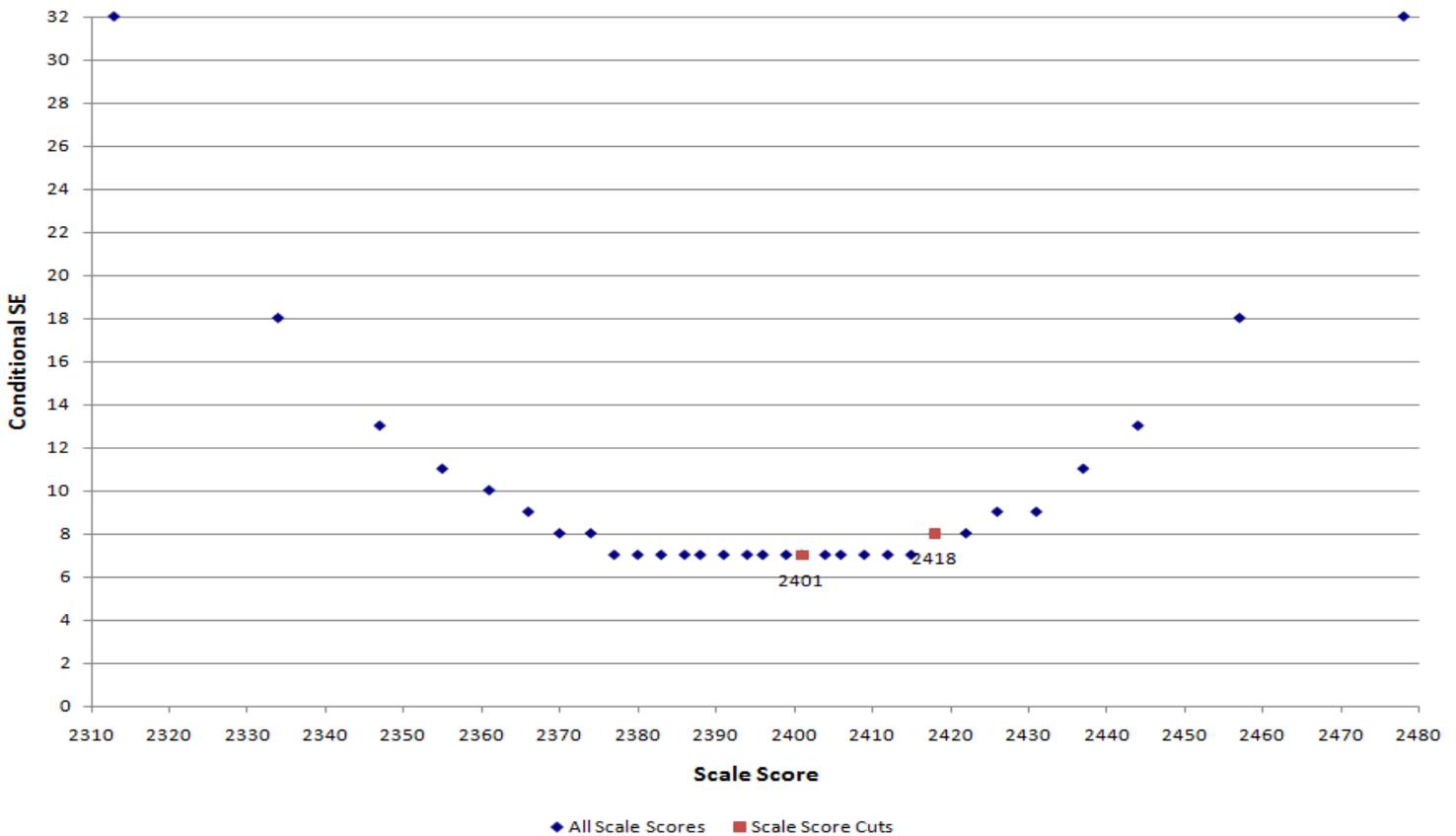
Fall 2010 Mathematics Grade 3 Standard Error Curve



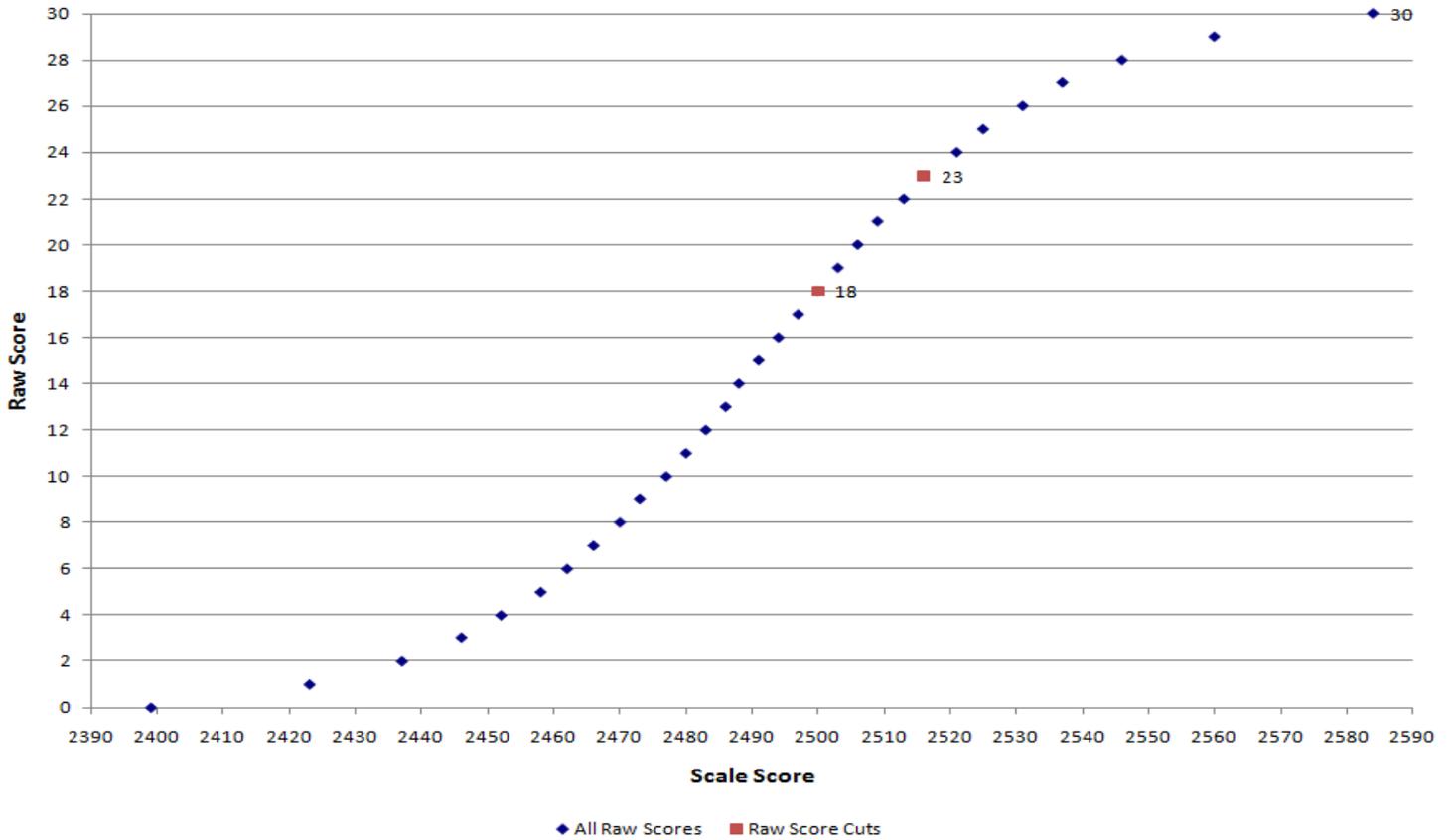
Fall 2010 Mathematics Grade 4 Test Characteristic Curve



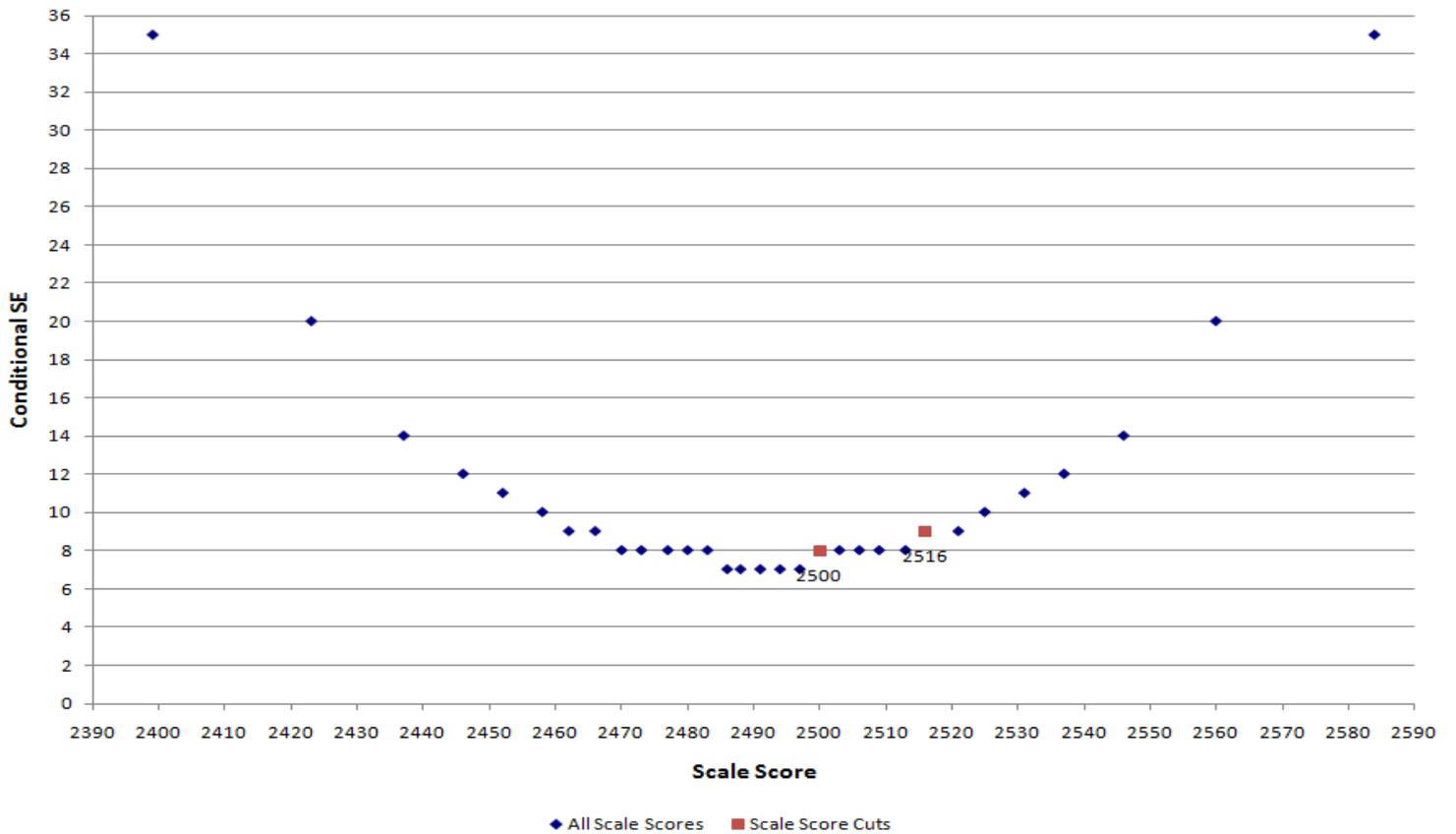
Fall 2010 Mathematics Grade 4 Standard Error Curve



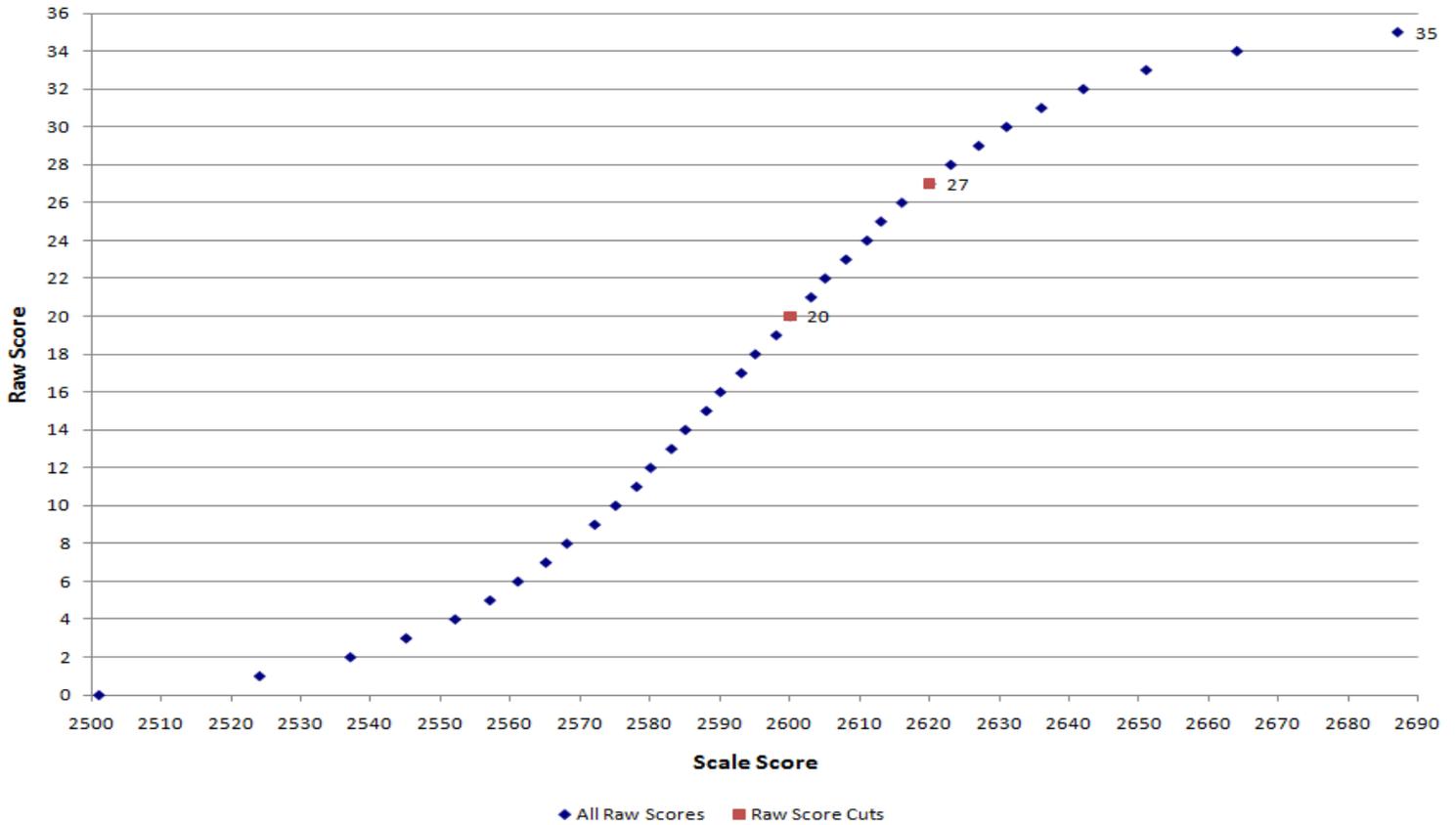
Fall 2010 Mathematics Grade 5 Test Characteristic Curve



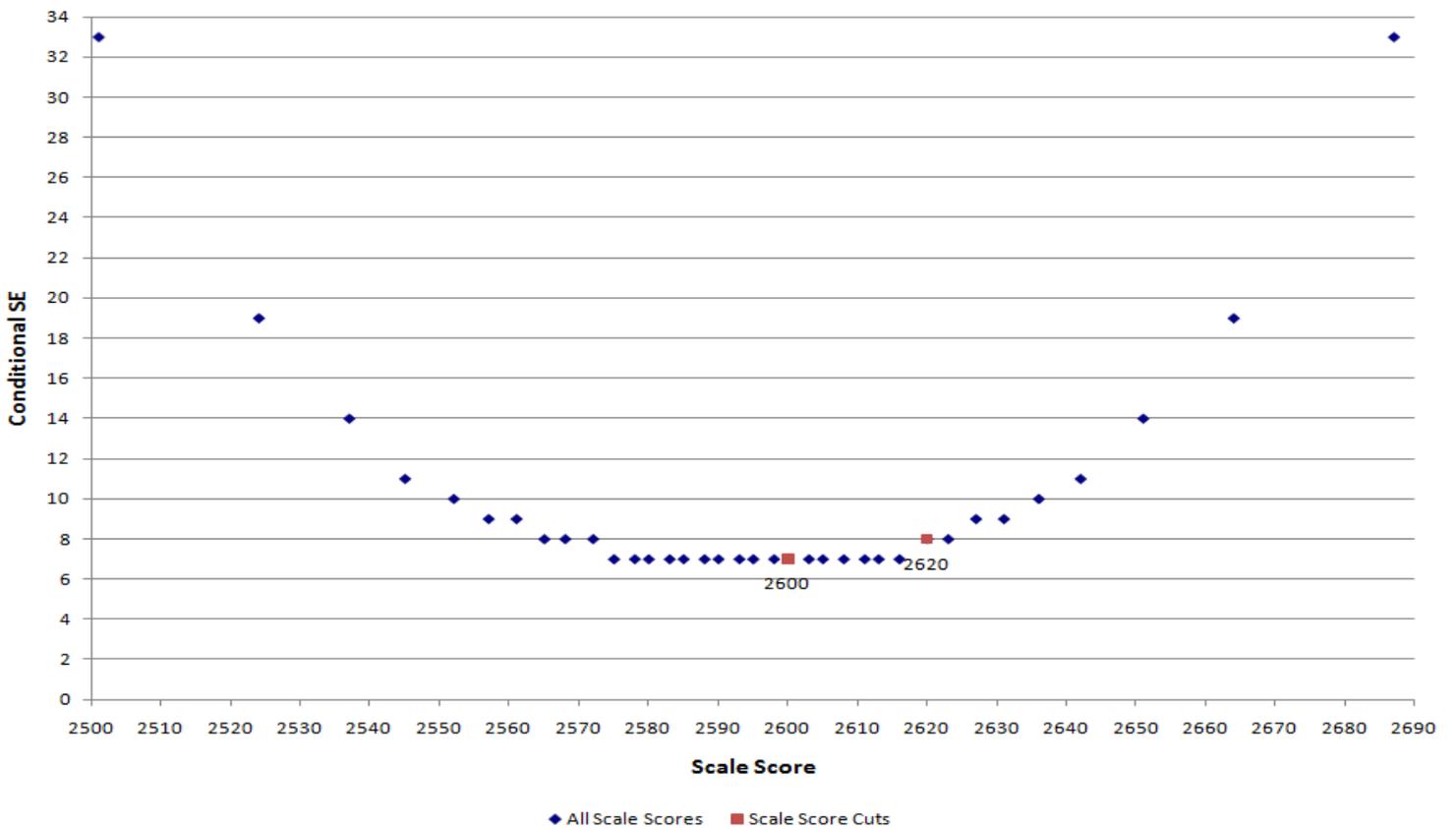
Fall 2010 Mathematics Grade 5 Standard Error Curve



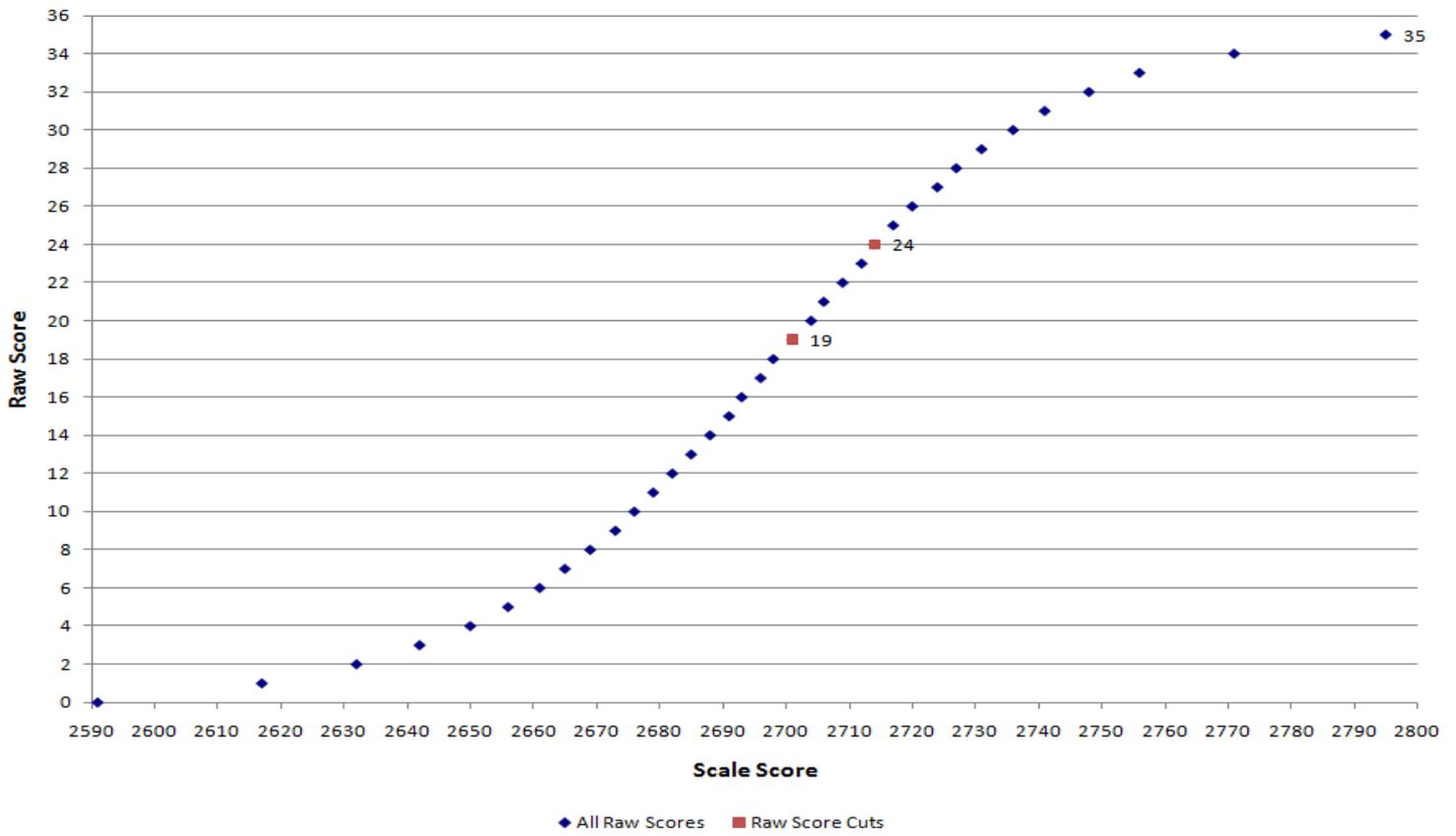
Fall 2010 Mathematics Grade 6 Test Characteristic Curve



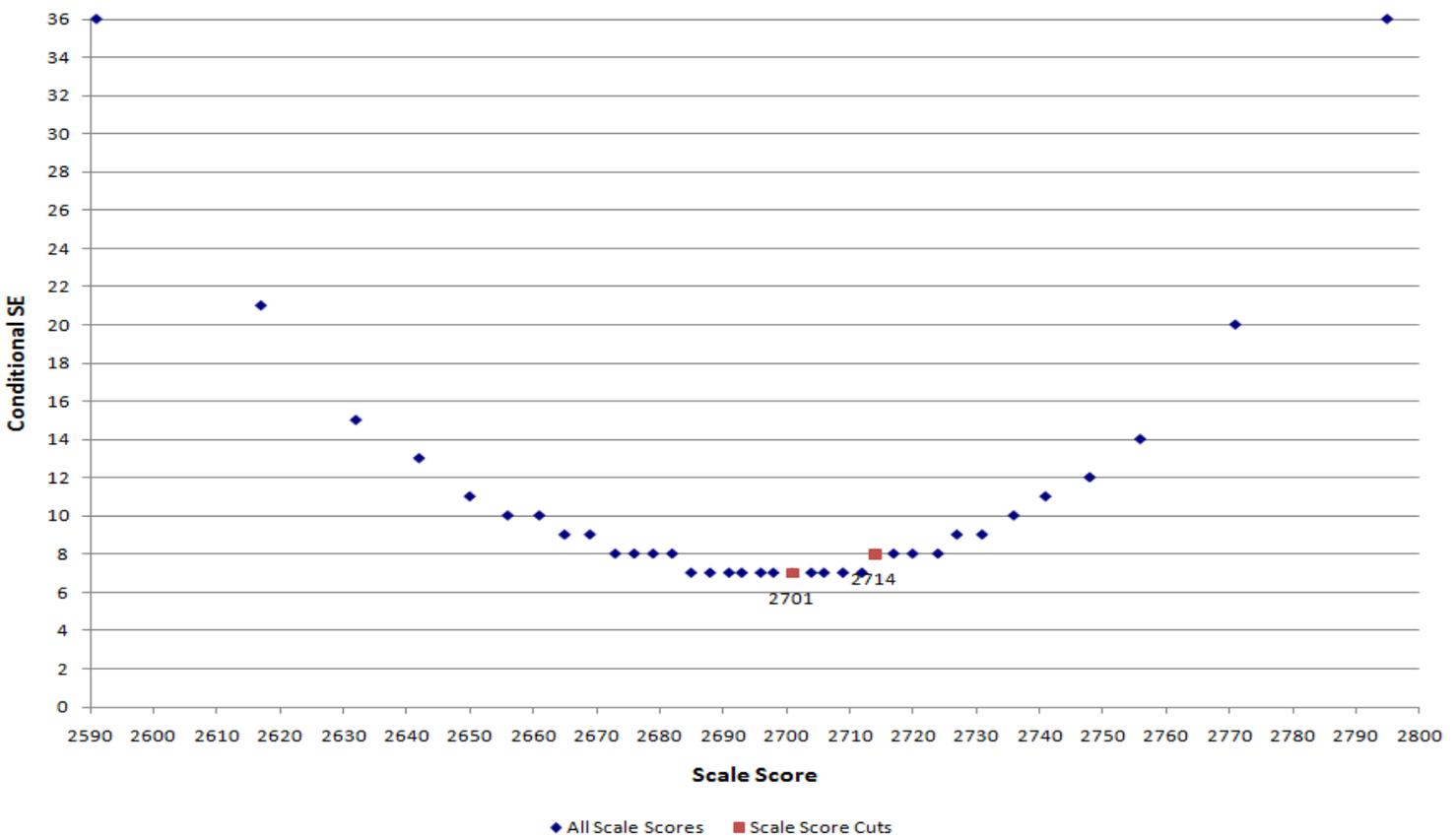
Fall 2010 Mathematics Grade 6 Standard Error Curve



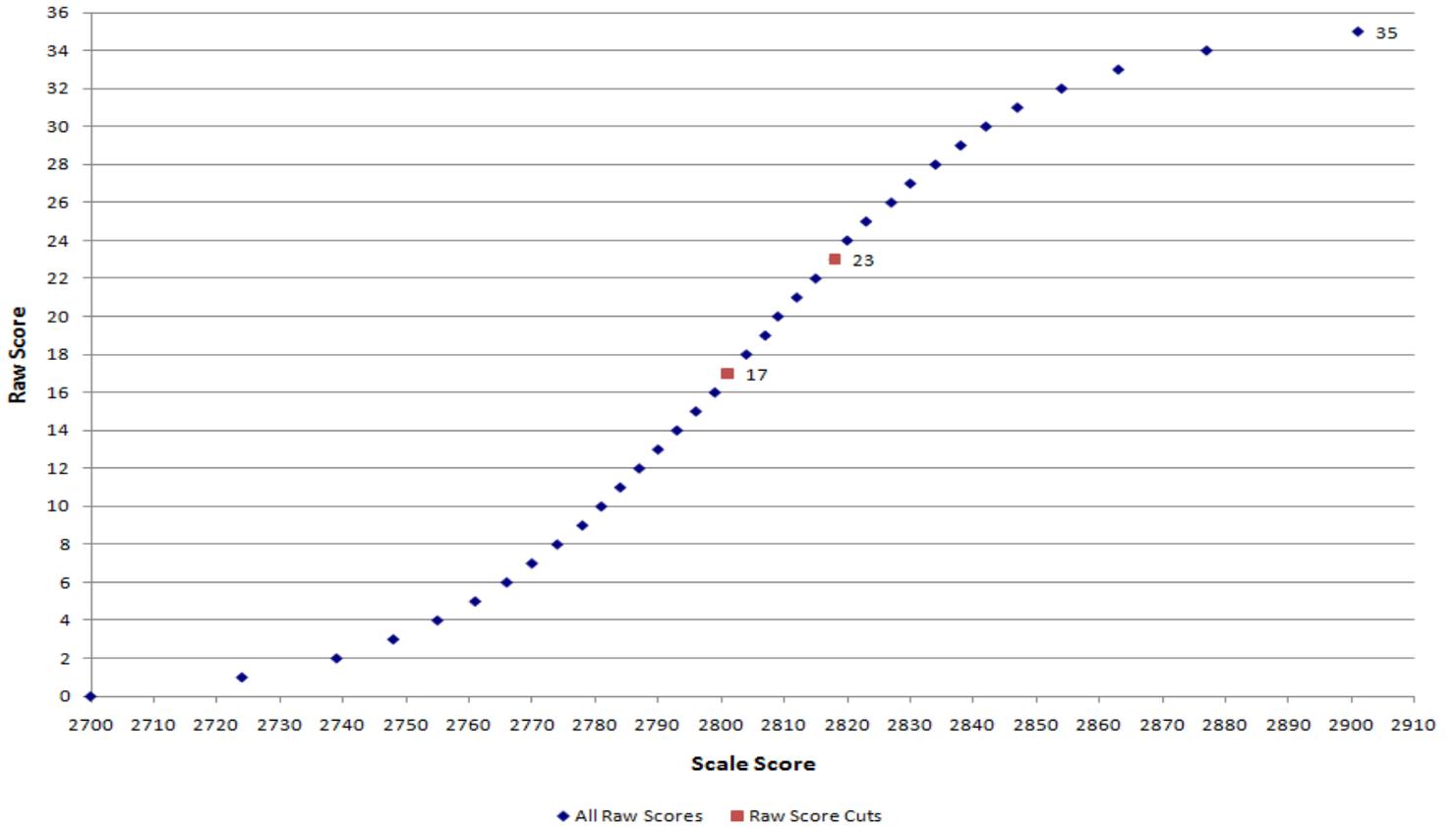
Fall 2010 Mathematics Grade 7 Test Characteristic Curve



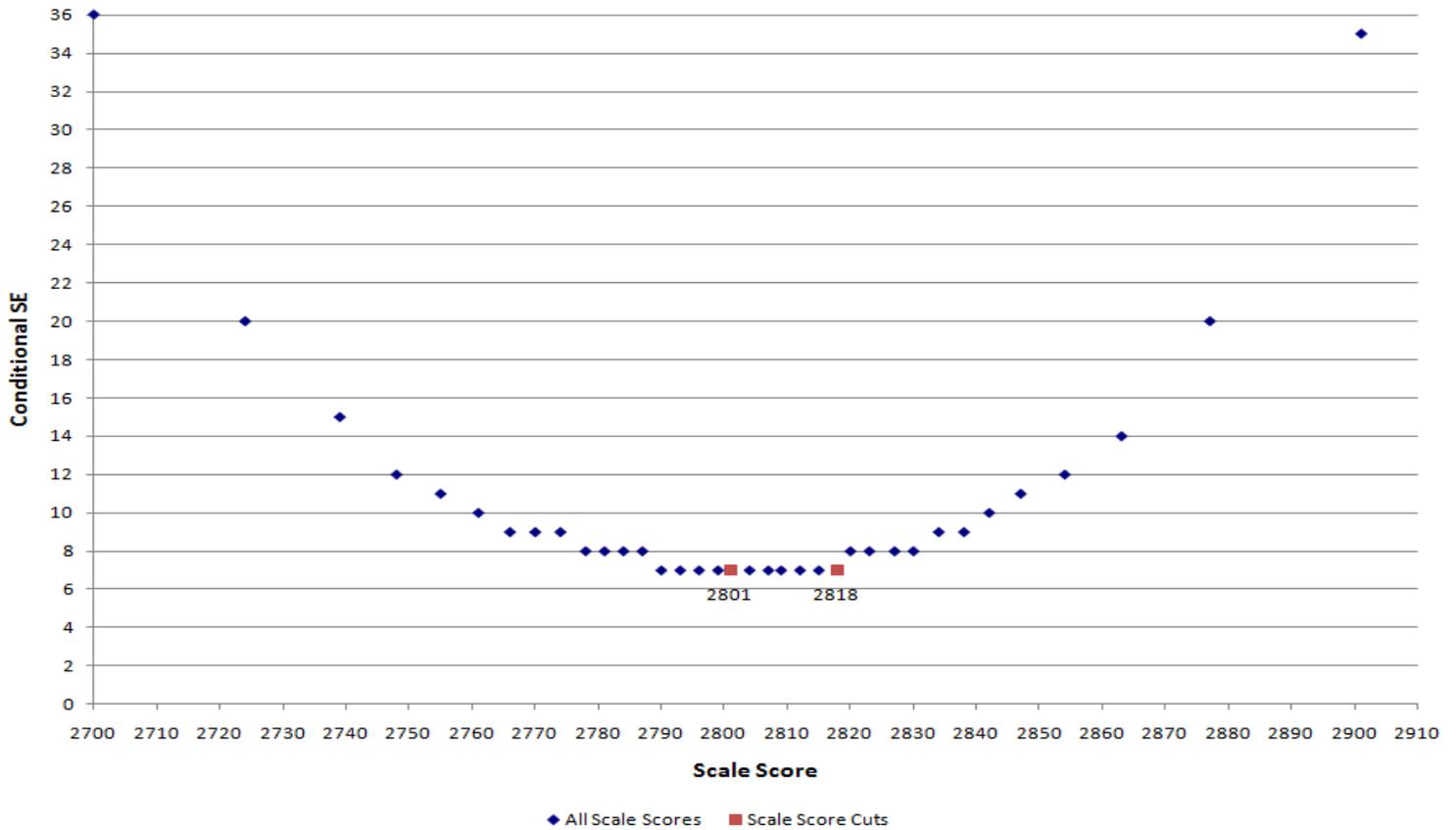
Fall 2010 Mathematics Grade 7 Standard Error Curve



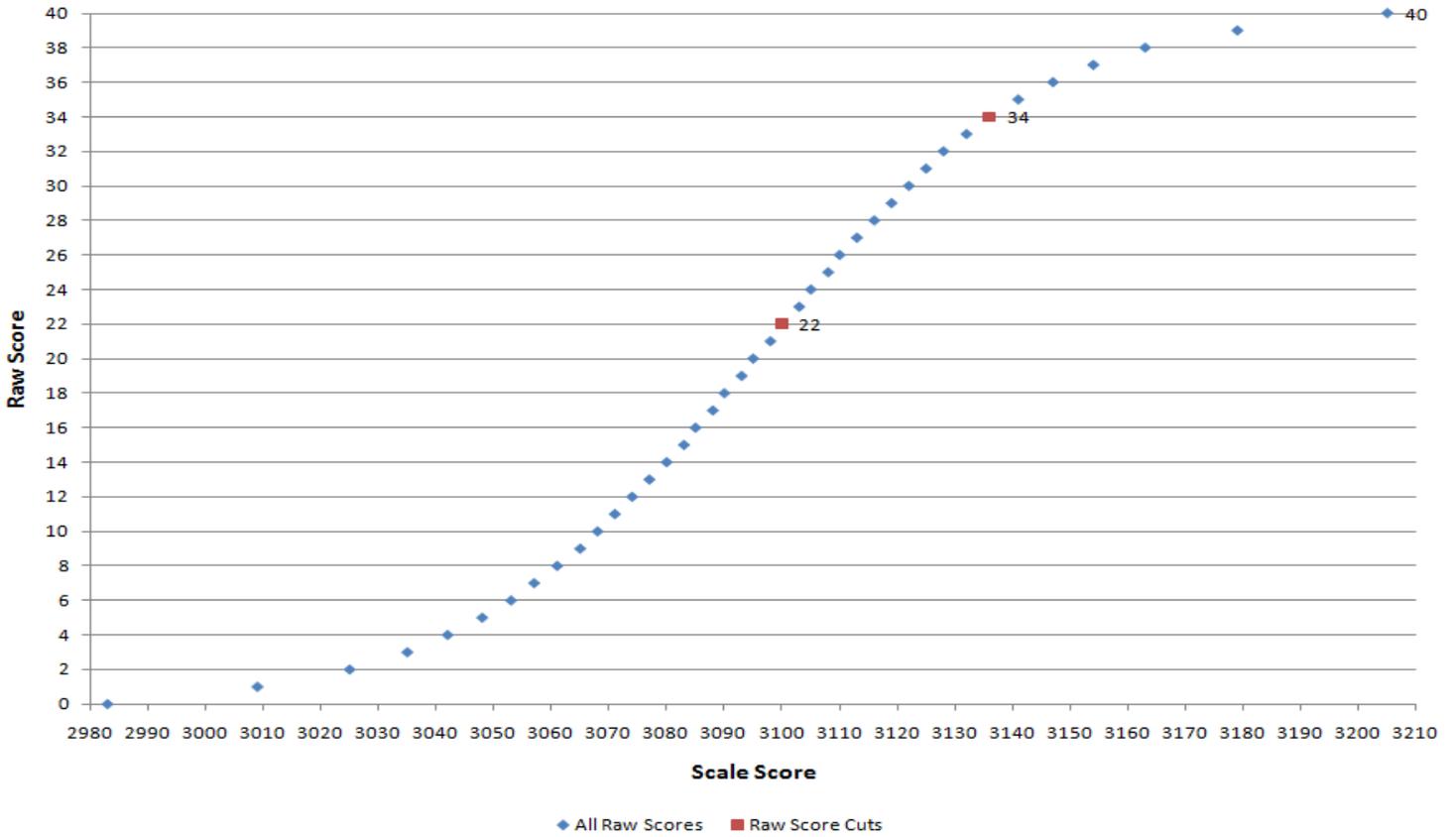
Fall 2010 Mathematics Grade 8 Test Characteristic Curve



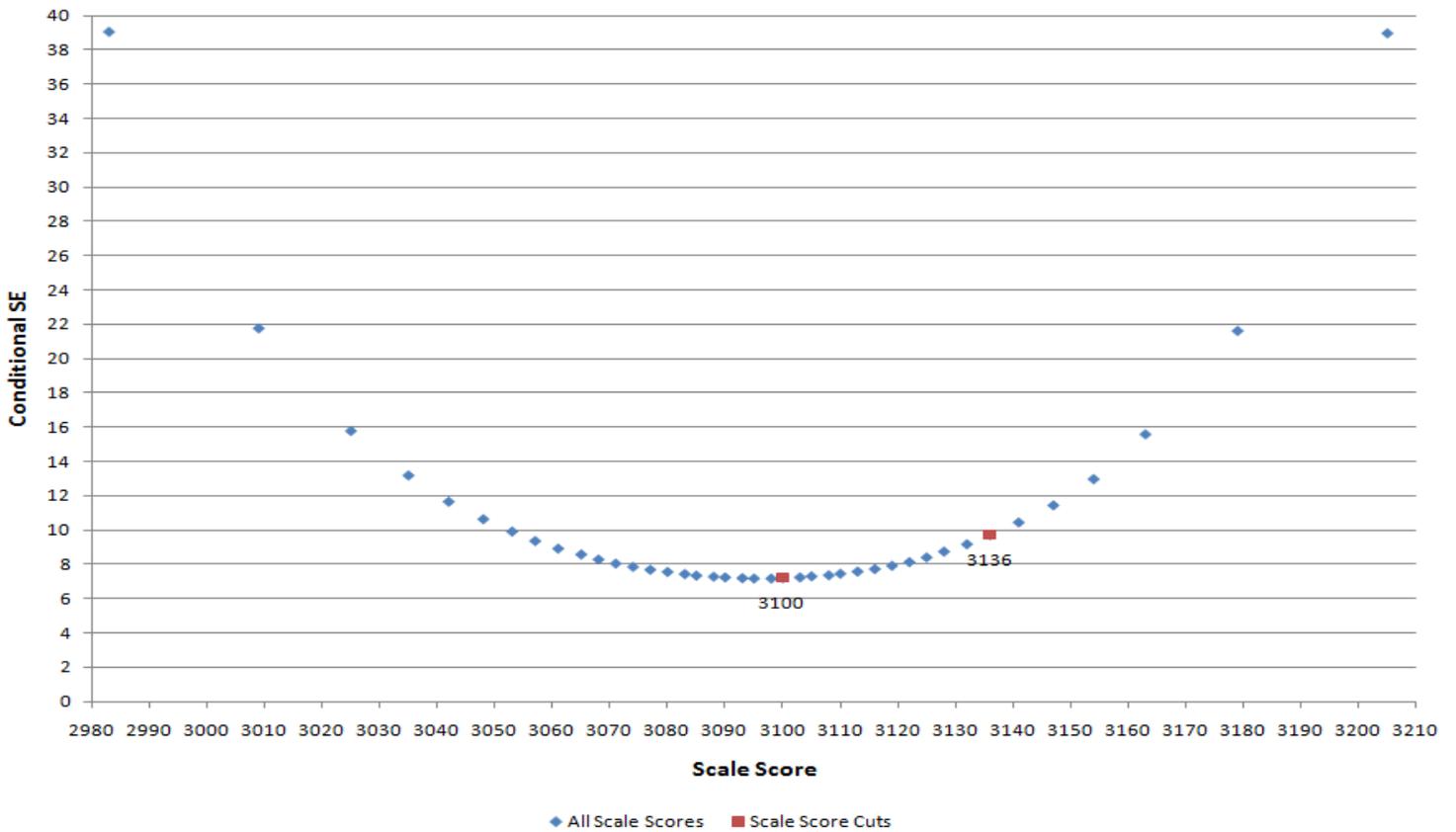
Fall 2010 Mathematics Grade 8 Standard Error Curve



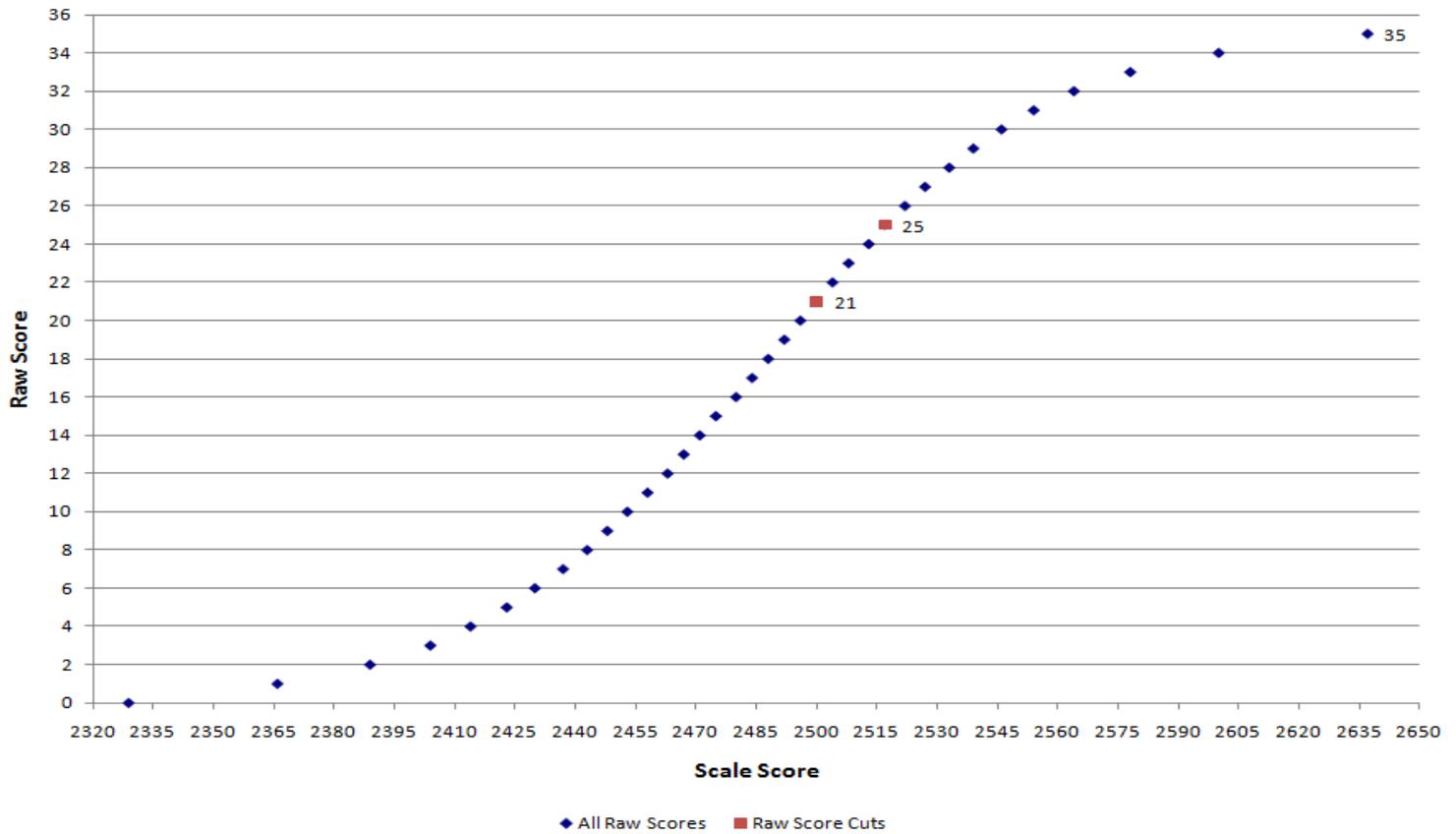
Spring 2011 Mathematics Grade 11 Test Characteristic Curve



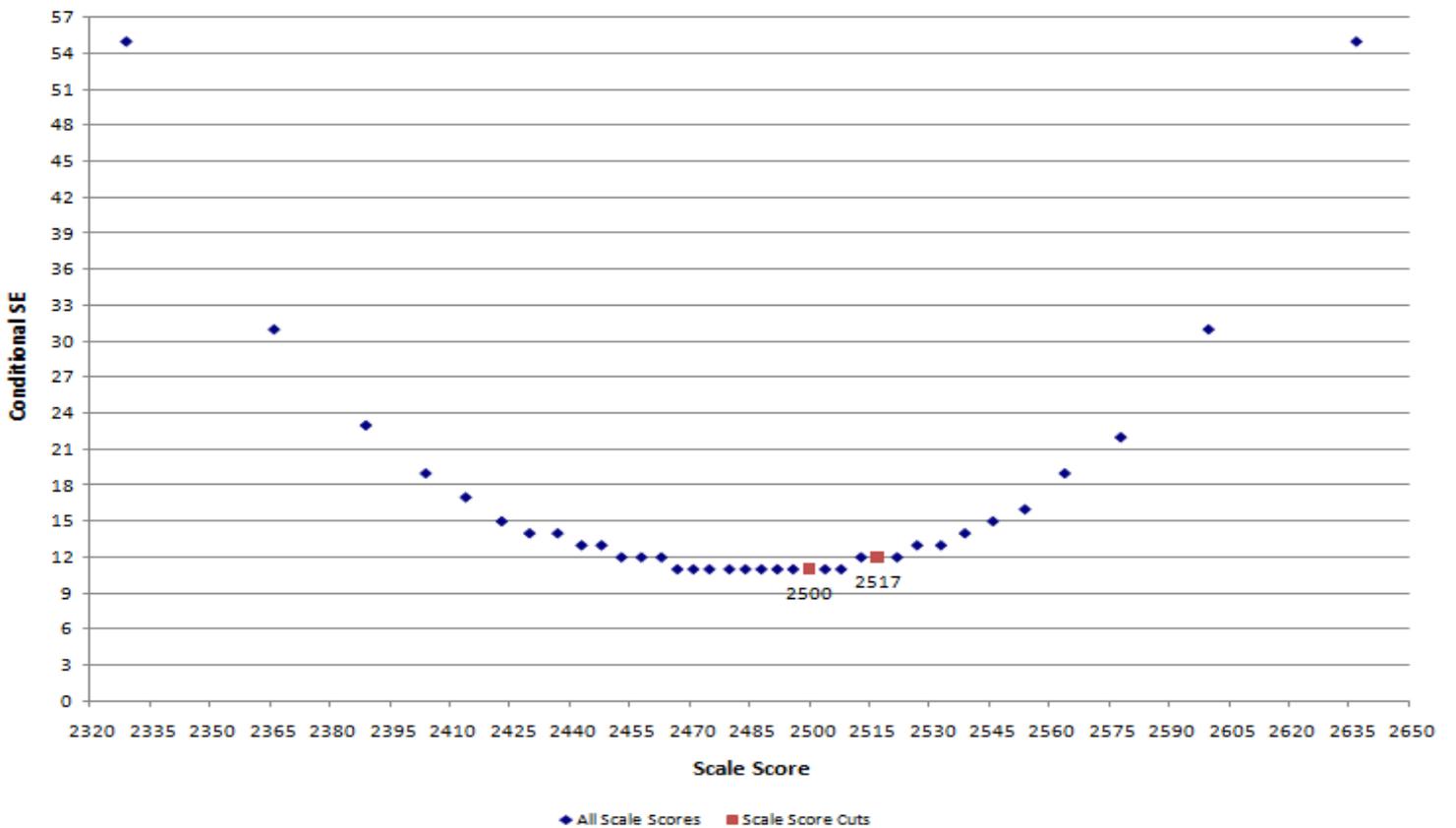
Spring 2011 Mathematics Grade 11 Standard Error Curve



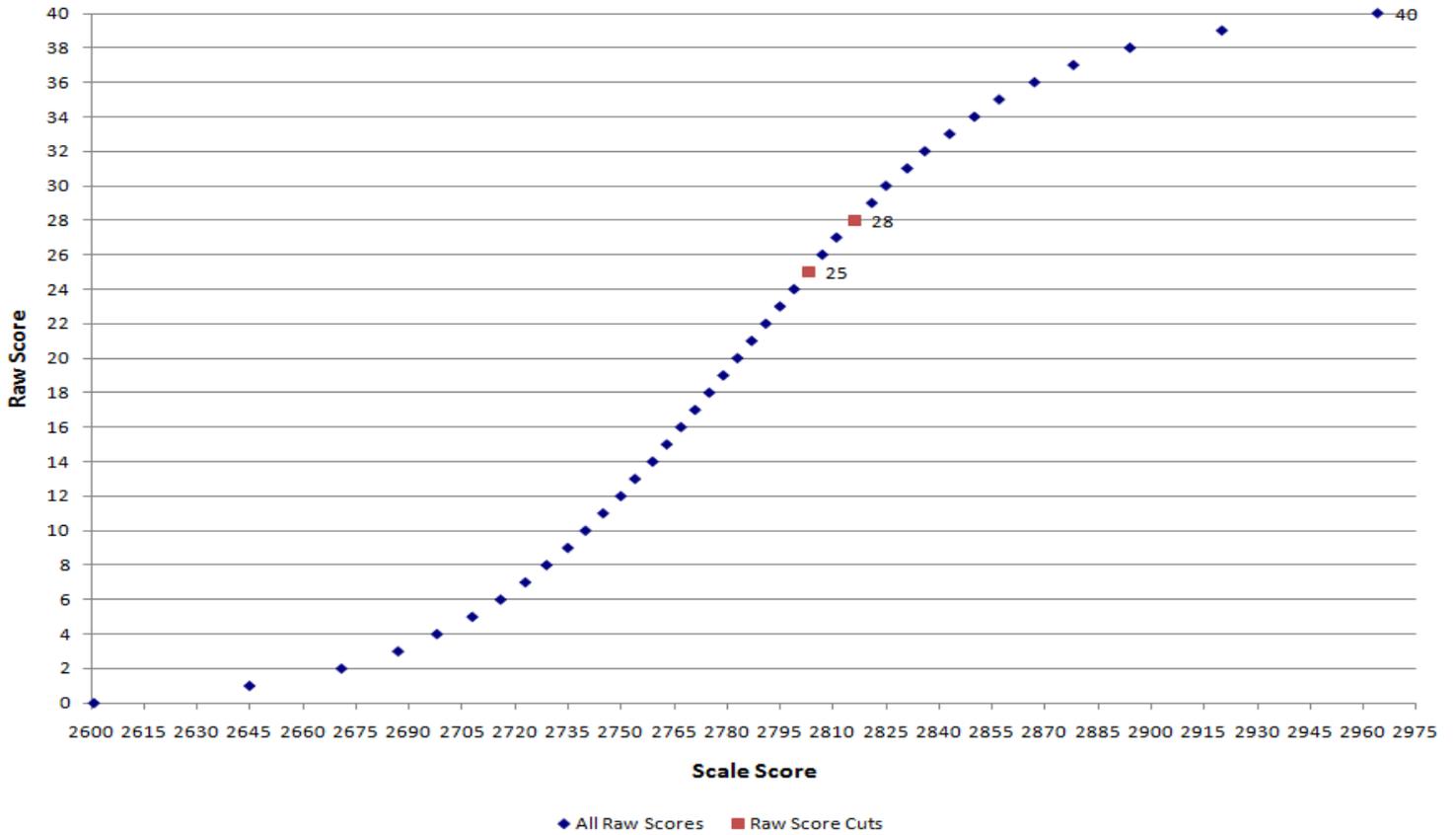
Fall 2010 Science Grade 5 Test Characteristic Curve



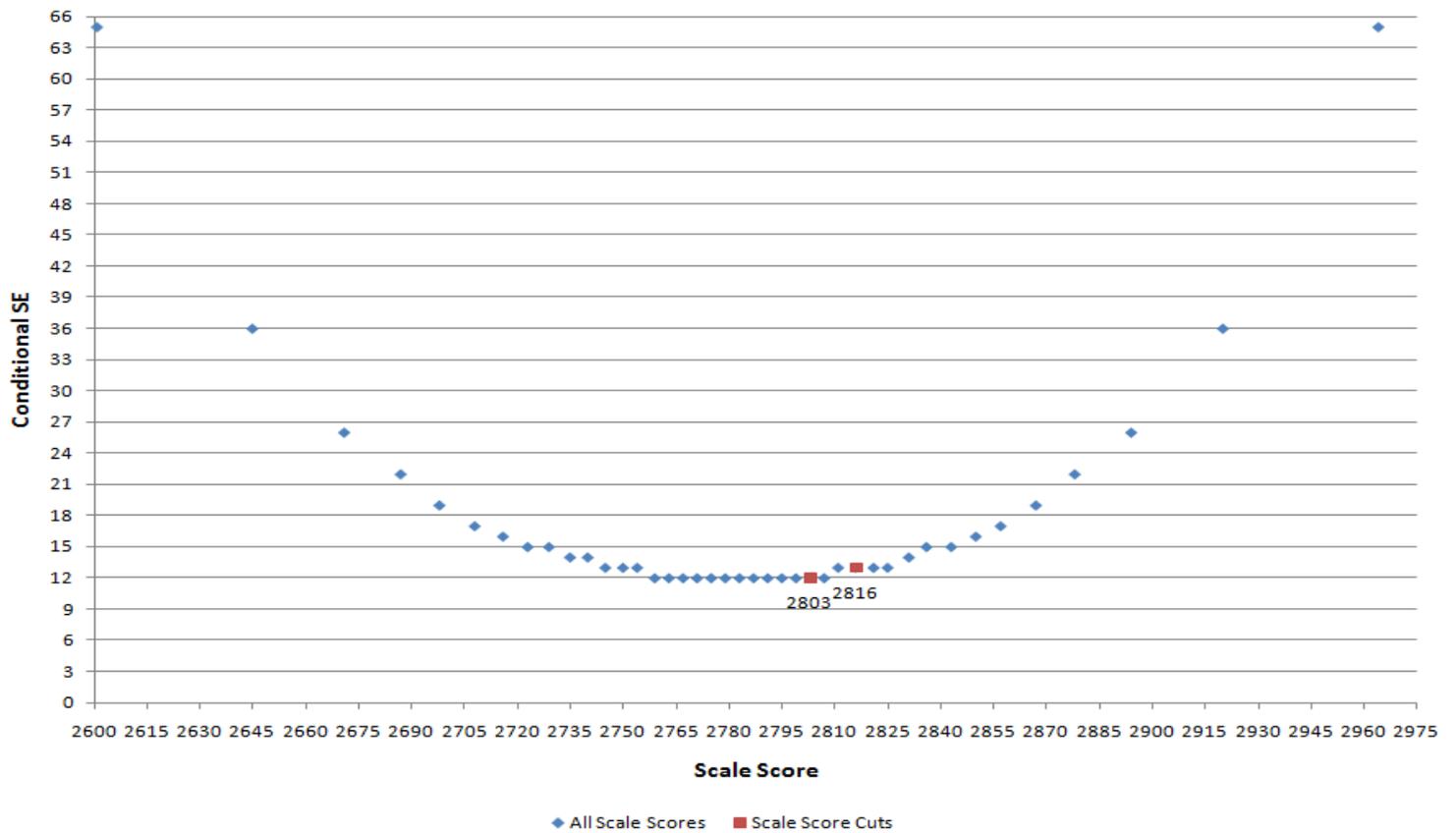
Fall 2010 Science Grade 5 Standard Error Curve



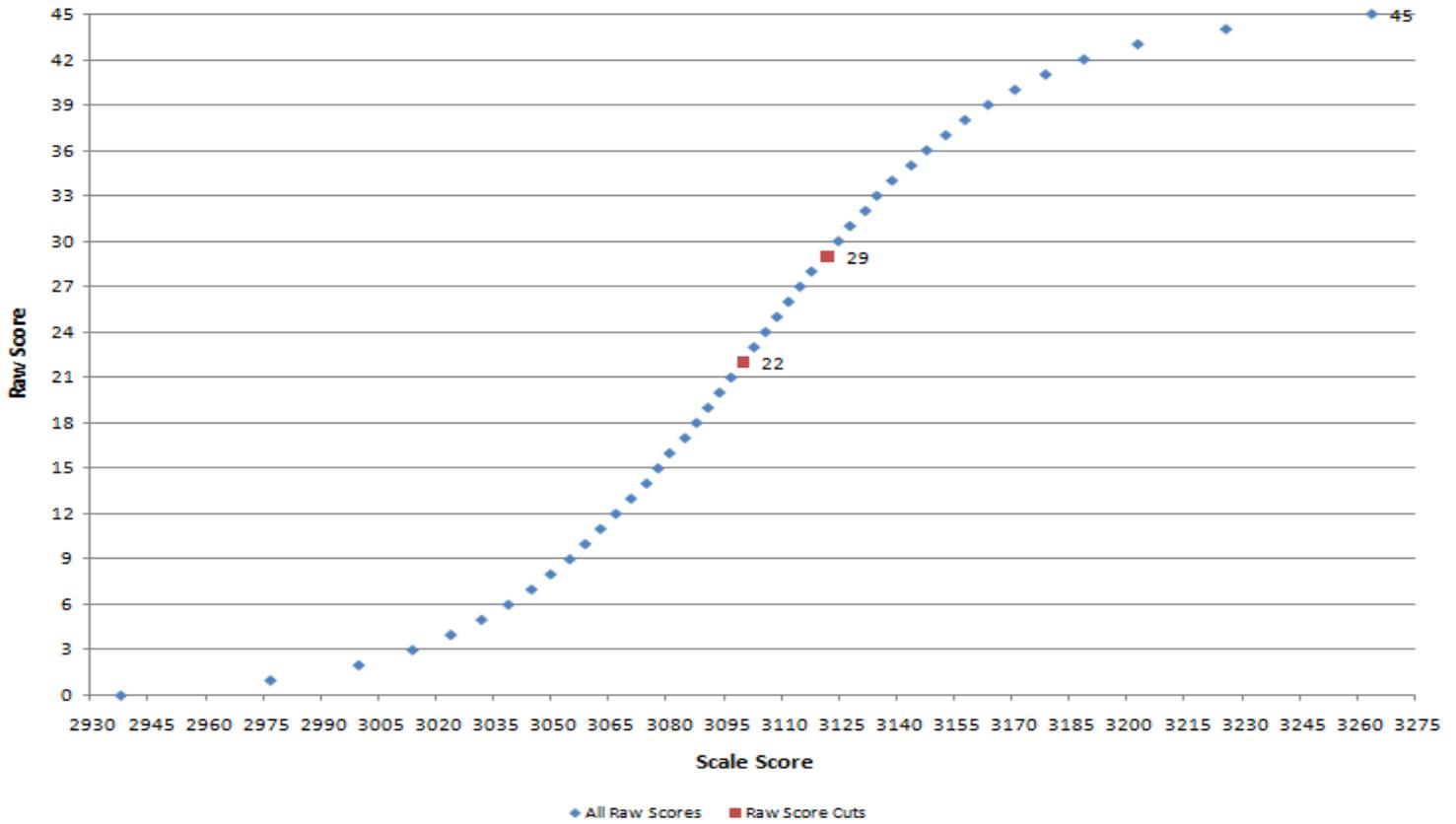
Fall 2010 Science Grade 8 Test Characteristic Curve



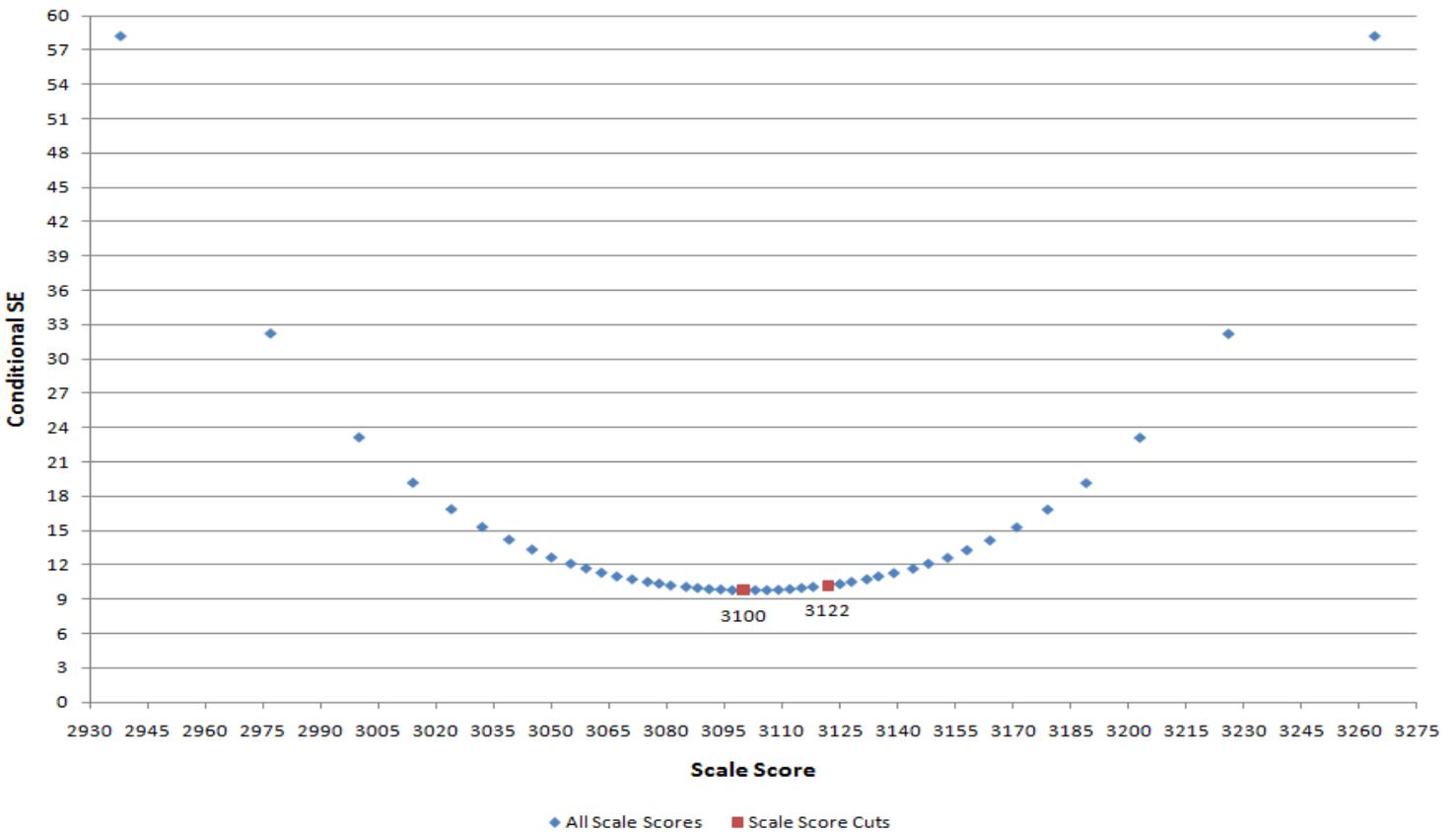
Fall 2010 Science Grade 8 Standard Error Curve



Spring 2011 Science Grade 11 Test Characteristic Curve



Spring 2011 Science Grade 11 Standard Error Curve





Michigan Department of Education
Bureau of Assessment and Accountability
608 West Allegan Street
P.O. Box 30008
Lansing, MI 48909
(877) 560-8378
www.mi.gov/mi-access