MI Access™
Michigan's Alternate Assessment Program

Participation & Supported Independence

ELA, Mathematics, and Science

Michigan's Alternate Assessment Program

Michigan Department of Education
Bureau of Assessment and Accountability

Technical Report
2011-2012 Addendum

October 2012

Compiled by
Adam E. Wyse

TABLE OF CONTENTS

LIST OF TABLES iii

INTRODUCTION 1

SECTION 1: FORM DESIGN 3

SECTION 2: PARTICIPATION IN ASSESSMENTS AND SUBGROUP ANALYSIS 6

SECTION 3: SCORE RELIABILITY AND SUMMARY STATISTICS 23

SECTION 4: RATER CONSISTENCY 26

SECTION 5: CLASSIFICATION ACCURACY AND CONSISTENCY 30

SECTION 6: INTERRELATIONS AMONG STRANDS 35

SECTION 7: SUMMARY OF ITEMS FLAGGED FOR DIFFICULTY AND DISCRIMINATION 49

SECTION 8: SPRING ONLINE PILOT ANALYSES 51

SECTION 9: VERIFICATION OF PSYCHOMETRIC PROCEDURES 55

SECTION 10: VALIDITY ARGUMENT 59

REFERENCES 63

LIST OF TABLES

TABLE 1.1: OPERATIONAL ELA TEST BLUEPRINT: GRADES 3 – 5, GRADES 6 – 8, AND GRADE 11 3

TABLE 1.2: OPERATIONAL MATHEMATICS TEST BLUEPRINT: GRADES 3 – 5 AND 6 – 8 4

TABLE 1.3: OPERATIONAL MATHEMATICS TEST BLUEPRINT: GRADE 11 4

TABLE 1.4: OPERATIONAL SCIENCE TEST BLUEPRINT: GRADES 5, 8, AND 11 5

TABLE 2.1: 2011–2012 N-COUNTS AND PERCENTS BY GENDER AND GRADE FOR ELA PARTICIPATION 6

TABLE 2.2: 2011–2012 N-COUNTS AND PERCENTS BY GENDER AND GRADE FOR ELA SUPPORTED INDEPENDENCE 6

TABLE 2.3: 2011–2012 N-COUNTS AND PERCENTS BY GENDER AND GRADE FOR MATHEMATICS PARTICIPATION 7

TABLE 2.4: 2011–2012 N-COUNTS AND PERCENTS BY GENDER AND GRADE FOR MATHEMATICS SUPPORTED INDEPENDENCE 7

TABLE 2.5: 2011–2012 N-COUNTS AND PERCENTS BY GENDER AND GRADE FOR SCIENCE PARTICIPATION .. 7

TABLE 2.6: 2011–2012 N-COUNTS AND PERCENTS BY GENDER AND GRADE FOR SCIENCE SUPPORTED INDEPENDENCE 8

TABLE 2.7: 2011–2012 N-COUNTS AND PERCENTS BY ETHNICITY AND GRADE FOR ELA PARTICIPATION 8

TABLE 2.8: 2011–2012 N-COUNTS AND PERCENTS BY ETHNICITY AND GRADE FOR ELA SUPPORTED INDEPENDENCE 9

TABLE 2.9: 2011–2012 N-COUNTS AND PERCENTS BY ETHNICITY AND GRADE FOR MATHEMATICS PARTICIPATION 9

TABLE 2.10: 2011–2012 N-COUNTS AND PERCENTS BY ETHNICITY AND GRADE FOR MATHEMATICS SUPPORTED INDEPENDENCE 10

TABLE 2.11: 2011–2012 N-COUNTS AND PERCENTS BY ETHNICITY AND GRADE FOR SCIENCE PARTICIPATION 10

TABLE 2.12: 2011–2012 N-COUNTS AND PERCENTS BY ETHNICITY AND GRADE FOR SCIENCE SUPPORTED INDEPENDENCE 11

TABLE 2.13: 2011–2012 MATERIAL ORDERS VERSUS TEST FORMS USED 12

TABLE 2.14: 2011–2012 PERCENT OF STUDENTS BY SUBGROUP AND FORM ELA PARTICIPATION 13

TABLE 2.15: 2011–2012 PERCENT OF STUDENTS BY SUBGROUP AND FORM MATHEMATICS PARTICIPATION 14

TABLE 2.16: 2011–2012 PERCENT OF STUDENTS BY SUBGROUP AND FORM SCIENCE PARTICIPATION 15

TABLE 2.17: 2011–2012 PERCENT OF STUDENTS BY SUBGROUP AND FORM ELA SUPPORTED INDEPENDENCE 16

TABLE 2.18: 2011–2012 PERCENT OF STUDENTS BY SUBGROUP AND FORM MATHEMATICS SUPPORTED INDEPENDENCE 17

TABLE 2.19: 2011–2012 PERCENT OF STUDENTS BY SUBGROUP AND FORM SCIENCE SUPPORTED INDEPENDENCE 18

TABLE 2.20: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY GENDER AND GRADE – ELA PARTICIPATION 19

TABLE 2.21: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY GENDER AND GRADE – ELA SUPPORTED INDEPENDENCE 19

TABLE 2.22: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY GENDER AND GRADE – MATHEMATICS PARTICIPATION 19

TABLE 2.23: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY GENDER AND GRADE – MATHEMATICS SUPPORTED INDEPENDENCE 20

TABLE 2.24: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY GENDER AND GRADE – SCIENCE PARTICIPATION 20

TABLE 2.25: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY GENDER AND GRADE – SCIENCE SUPPORTED INDEPENDENCE 20

TABLE 2.26: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY ETHNICITY AND GRADE – ELA PARTICIPATION 21

TABLE 2.27: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY ETHNICITY AND GRADE – ELA SUPPORTED INDEPENDENCE 21

TABLE 2.28: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY ETHNICITY AND GRADE – MATHEMATICS PARTICIPATION 21

TABLE 2.29: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY ETHNICITY AND GRADE – MATHEMATICS SUPPORTED INDEPENDENCE 22

TABLE 2.30: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY ETHNICITY AND GRADE – SCIENCE PARTICIPATION 22

TABLE 2.31: 2011–2012 OPERATIONAL FORM SCORE SUMMARIES BY ETHNICITY AND GRADE – SCIENCE SUPPORTED INDEPENDENCE 22

TABLE 3.1: 2011–2012 OPERATIONAL FORM SUMMARIES, INCLUDING SAMPLE SIZE, SCORE STATISTICS, COEFFICIENT ALPHA, AND SEM BY GRADE – ELA PARTICIPATION 23

TABLE 3.2: 2011–2012 OPERATIONAL FORM SUMMARIES, INCLUDING SAMPLE SIZE, SCORE STATISTICS, COEFFICIENT ALPHA, AND SEM BY GRADE – ELA SUPPORTED INDEPENDENCE 24

TABLE 3.3: 2011–2012 OPERATIONAL FORM SUMMARIES, INCLUDING SAMPLE SIZE, SCORE STATISTICS, COEFFICIENT ALPHA, AND SEM BY GRADE – MATHEMATICS PARTICIPATION 24

TABLE 3.4: 2011–2012 OPERATIONAL FORM SUMMARIES, INCLUDING SAMPLE SIZE, SCORE STATISTICS, COEFFICIENT ALPHA, AND SEM BY GRADE – MATHEMATICS SUPPORTED INDEPENDENCE 25

TABLE 3.5: 2011–2012 OPERATIONAL FORM SUMMARIES, INCLUDING SAMPLE SIZE, SCORE STATISTICS, COEFFICIENT ALPHA, AND SEM BY GRADE – SCIENCE PARTICIPATION 25

TABLE 3.6: 2011–2012 OPERATIONAL FORM SUMMARIES, INCLUDING SAMPLE SIZE, SCORE STATISTICS, COEFFICIENT ALPHA, AND SEM BY GRADE – SCIENCE SUPPORTED INDEPENDENCE 25

TABLE 4.1: 2011–2012 OPERATIONAL FORM PERCENT PERFECT INTERRATER AGREEMENT RATES PARTICIPATION ELA 26

TABLE 4.2: 2011–2012 OPERATIONAL FORM PERCENT PERFECT INTERRATER AGREEMENT RATES PARTICIPATION MATHEMATICS 27

TABLE 4.3: 2011–2012 OPERATIONAL FORM PERCENT PERFECT INTERRATER AGREEMENT RATES PARTICIPATION SCIENCE 27

TABLE 4.4: 2011–2012 OPERATIONAL FORM PERCENT PERFECT INTERRATER AGREEMENT RATES SUPPORTED INDEPENDENCE ELA 28

TABLE 4.5: 2011–2012 OPERATIONAL FORM PERCENT PERFECT INTERRATER AGREEMENT RATES SUPPORTED INDEPENDENCE MATHEMATICS 28

TABLE 4.6: 2011–2012 OPERATIONAL FORM PERCENT PERFECT INTERRATER AGREEMENT RATES SUPPORTED INDEPENDENCE SCIENCE 29

TABLE 5.1: PERCENT OF STUDENTS BY PROFICIENCY LEVEL – PARTICIPATION 30

TABLE 5.2: PERCENT OF STUDENTS BY PROFICIENCY LEVEL – SUPPORTED INDEPENDENCE 31

TABLE 5.3: ESTIMATED CLASSIFICATION ACCURACY AND CONSISTENCY BY CONTENT AREA AND GRADE PARTICIPATION 32

TABLE 5.4: ESTIMATED CLASSIFICATION ACCURACY AND CONSISTENCY BY SUBJECT AND GRADE SUPPORTED INDEPENDENCE 33

TABLE 6.1: 2011–2012 OPERATIONAL FORM INTERCORRELATIONS BETWEEN SECTION SCORES BY GRADE – ELA PARTICIPATION 36

TABLE 6.2: 2011–2012 OPERATIONAL FORM INTERCORRELATIONS BETWEEN SECTION SCORES BY GRADE – ELA SUPPORTED INDEPENDENCE 37

TABLE 6.3: FALL 2011 OPERATIONAL FORM INTERCORRELATIONS BETWEEN SECTION SCORES BY GRADE – MATHEMATICS PARTICIPATION 38

TABLE 6.4: FALL 2011 OPERATIONAL FORM INTERCORRELATIONS BETWEEN SECTION SCORES BY GRADE – MATHEMATICS SUPPORTED INDEPENDENCE 39

TABLE 6.5: SPRING 2012 OPERATIONAL FORM INTERCORRELATIONS BETWEEN SECTION SCORES MATHEMATICS PARTICIPATION 40

TABLE 6.6: SPRING 2012 OPERATIONAL FORM INTERCORRELATIONS BETWEEN SECTION SCORES MATHEMATICS SUPPORTED INDEPENDENCE 40

TABLE 6.7: 2011–2012 OPERATIONAL FORM INTERCORRELATIONS BETWEEN SECTION SCORES BY GRADE – SCIENCE PARTICIPATION 40

TABLE 6.8: 2011–2012 OPERATIONAL FORM INTERCORRELATIONS BETWEEN SECTION SCORES BY GRADE – SCIENCE SUPPORTED INDEPENDENCE 41

TABLE 6.9: 2011–2012 OPERATIONAL FORM SUMMARY STATISTICS FOR SECTION SCORES BY GRADE – ELA PARTICIPATION 42

TABLE 6.10: 2011–2012 OPERATIONAL FORM SUMMARY STATISTICS FOR SECTION SCORES BY GRADE – ELA SUPPORTED INDEPENDENCE 43

TABLE 6.11: FALL 2011 OPERATIONAL FORM SUMMARY STATISTICS FOR SECTION SCORES BY GRADE – MATHEMATICS PARTICIPATION 44

TABLE 6.12: FALL 2011 OPERATIONAL FORM SUMMARY STATISTICS FOR SECTION SCORES BY GRADE – MATHEMATICS SUPPORTED INDEPENDENCE 45

TABLE 6.13: SPRING 2012 OPERATIONAL FORM SUMMARY STATISTICS FOR SECTION SCORES MATHEMATICS PARTICIPATION 46

TABLE 6.14: SPRING 2012 OPERATIONAL FORM SUMMARY STATISTICS FOR SECTION SCORES MATHEMATICS SUPPORTED INDEPENDENCE 46

TABLE 6.15: 2011–2012 OPERATIONAL FORM SUMMARY STATISTICS FOR SECTION SCORES BY GRADE – SCIENCE PARTICIPATION 47

TABLE 6.16: 2011–2012 OPERATIONAL FORM SUMMARY STATISTICS FOR SECTION SCORES BY GRADE – SCIENCE SUPPORTED INDEPENDENCE 48

TABLE 7.1: NUMBER OF ITEMS FLAGGED FOR DIFFICULTY AND DISCRIMINATION 50

TABLE 8.1: 2011-2012 PERCENT OF STUDENTS BY SUBGROUP AND CONTENT AREA FOR ONLINE AND PAPER FORMS FOR PARTICIPATION 51

TABLE 8.2: 2011-2012 PERCENT OF STUDENTS BY SUBGROUP AND CONTENT AREA FOR ONLINE AND PAPER FORMS FOR SUPPORTED INDEPENDENCE 52

TABLE 8.3: 2011-2012 SCORE SUMMARIES BY ONLINE AND PAPER – PARTICIPATION 52

TABLE 8.4: 2011-2012 SCORE SUMMARIES BY ONLINE AND PAPER – SUPPORTED INDEPENDENCE 53

TABLE 8.5: 2011-2012 PERCENT EXACT RATER AGREEMENT FOR PRIMARY AND SHADOW RATERS FOR ONLINE AND PAPER – PARTICIPATION OPERATIONAL ITEM RATINGS 53

TABLE 8.6: 2011-2012 PERCENT EXACT RATER AGREEMENT FOR PRIMARY AND SHADOW RATERS FOR ONLINE AND PAPER – SUPPORTED INDEPENDENCE OPERATIONAL ITEM RATINGS 54

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 (AERA, APA, & NCME, 1999, p. 67). Technical Reports have been developed for the Participation and Supported Independence assessments and the Functional Independence assessments.

The following Technical Reports have been developed:

Functional Independence English Languages Arts and Mathematics, March 2007
Participation and Supported Independence English Language Arts and Mathematics, June 2007
Participation, Supported Independence, and 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 fourth annual addendum and includes the Participation and Supported Independence ELA, Mathematics, and Science assessments administered in the 2011 – 2012 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, 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 scaling, and reporting.

1. Form Design

The fall 2011 grades 3 to 8 and spring 2012 grade 11 Participation and Supported Independence ELA and Mathematics, and the fall 2011 grades 5 and 8 and spring 2012 grade 11 Science assessments were developed according to the test blueprints. The test blueprints for ELA Participation and Supported Independence are given in Table 1.1, which were unchanged from the previous year except for Participation grade 11. For Participation grade 11, there was one additional Comprehension item and one fewer Word Study item this year. The test blueprints for Mathematics Participation and Supported Independence for grades 3 to 5, and grades 6 to 8 are given in Table 1.2. Due to item availability, small changes were made to Supported Independence and Participation grades 6 to 8. For Supported Independence, Geometry and Data and Probability have each had two items, respectively, whereas last year they had three and one items. For Participation, Geometry went from three to four items and Measurement went from two to one items. The test blueprint for grade 11 Mathematics is given in Table 1.3, which is unchanged from last year. The test blueprints for Science Participation and Supported Independence are given in Table 1.4. The only change this year was to report the items separately for Constructing and Reflecting. For both Participation and Supported Independence and for all three content areas, three forms were developed for each grade group (elementary, grades 3 – 5; middle school, grades 6 – 8; and grade 11). The three forms contained the same operational, core items but differed by the field test items embedded in each form.

The Participation and Supported Independence assessments consist of a mixture of activity based and selected response items with picture cards that are scored by two raters. In addition to the numbers and the exact questions asked being different for Participation and Supported Independence, the two assessment levels differ in the number of answer choices for the selected response items (2 for Supported Independence and 3 for Participation) and the scoring rubrics. Students who take Participation are scored on a 3-point rubric by two raters and can receive credit when they get hand-over-hand assistance. Students who take Supported Independence are scored on a 2-point rubric by two raters and cannot get credit if they get hand-over-hand assistance.

Table 1.1
Operational ELA Test Blueprint: Grades 3 – 5,
Grades 6 – 8, and Grade 11

Strand	Participation		Supported Independence	
	Grades 3 – 5 and 6 –8	Grade 11	Grades 3 – 5 and 6 – 8	Grade 11
Word Study	4	3	4	4
Comprehension	2	3	5	5
Expressing Ideas	4	4	6	6
Total Core Items	10	10	15	15
Embedded Field-Test Items	5	5	5	5
Total Number of Items	15	15	20	20

Word Study and Comprehension are combined to compute the Accessing Information score in the student score reports.

Table 1.2
Operational Mathematics Test Blueprint: Grades 3 – 5 and 6 – 8

Strand	Participation		Supported Independence	
	Grades 3 – 5	Grades 6 – 8	Grades 3 – 5	Grades 6 – 8
Numbers and Operations	3	4	7	6
Algebra				2
Measurement	2	1	2	2
Geometry	4	4	4	3
Data and Probability	1	1	2	2
Total Core Items	10	10	15	15
Embedded Field-Test Items	5	5	5	5
Total Number of Items	15	15	20	20

Table 1.3
Operational Mathematics Test Blueprint: Grade 11

Strand	Participation	Supported Independence
Quantitative Literacy and Logic	6	10
Algebra and Functions	2	2
Geometry and Trigonometry	2	3
Total Core Items	10	15
Embedded Field-Test Items	5	5
Total Number of Items	15	20

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

Strand	Participation	Supported Independence
Constructing	1	1
Reflecting	1	1
Life Science	5	7
Physical Science	5	3
Earth Science	3	5
Total Core Items	15	17
Embedded Field-Test Items	5	5
Total Number of Items	20	22

2. Participation in Assessments and Subgroup Analysis

Participation in the assessments is monitored by gender and by racial/ethnic group. Participation counts by gender and grade are given in Tables 2.1 to 2.6, and participation counts by race/ethnicity and grade are given in Tables 2.7 to 2.12. The percent of male students across the grades and content areas for Participation ranged from 60.9% to 66.9%. For Supported Independence, the percent of males ranged from 63.6% to 70.3%. Across both levels and all three content areas, the largest racial/ethnic group was White students with 59.6% to 67.7% of the students, followed by Black students with 20.8% to 29.8% of the students, Hispanic students with 4.1% to 7.5% of the students, and Asian students with 0.8% to 3.7% of the students.

Table 2.1
2011–2012 N-Counts and Percents by Gender and Grade for ELA Participation

Grade	Female		Male		Total
	N	%	N	%	N
3	135	33.1	273	66.9	408
4	137	34.5	260	65.5	397
5	116	33.1	234	66.9	350
6	113	35.5	205	64.5	318
7	125	39.1	195	60.9	320
8	117	36.4	204	63.6	321
11	117	38.7	185	61.3	302

Table 2.2
2011–2012 N-Counts and Percents by Gender and Grade for ELA Supported Independence

Grade	Female		Male		Total
	N	%	N	%	N
3	149	31.6	322	68.4	471
4	144	29.8	340	70.2	484
5	154	30.3	354	69.7	508
6	163	33.3	326	66.7	489
7	177	36.4	309	63.6	486
8	179	35.4	326	64.6	505
11	207	35.8	371	64.2	578

Table 2.3
2011–2012 N-Counts and Percents by Gender and Grade for Mathematics Participation

Grade	Female		Male		Total
	N	%	N	%	N
3	134	33.2	270	66.8	404
4	136	34.3	260	65.7	396
5	115	33.2	231	66.8	346
6	113	35.8	203	64.2	316
7	125	39.1	195	60.9	320
8	117	36.4	204	63.6	321
11	117	38.7	185	61.3	302

Table 2.4
2011–2012 N-Counts and Percents by Gender and Grade for Mathematics Supported Independence

Grade	Female		Male		Total
	N	%	N	%	N
3	148	31.4	323	68.6	471
4	143	29.7	338	70.3	481
5	154	30.4	352	69.6	506
6	162	33.3	325	66.7	487
7	176	36.4	308	63.6	484
8	179	35.5	325	64.5	504
11	208	36.2	366	63.8	574

Table 2.5
2011–2012 N-Counts and Percents by Gender and Grade for Science Participation

Grade	Female		Male		Total
	N	%	N	%	N
5	114	33.4	227	66.6	341
8	117	37.7	193	62.3	310
11	116	38.7	184	61.3	300

Table 2.6
2011–2012 N-Counts and Percents by Gender and Grade for Science
Supported Independence

Grade	Female		Male		Total
	N	%	N	%	N
5	153	30.9	342	69.1	495
8	179	36.2	316	63.8	498
11	206	36.0	366	64.0	572

Table 2.7
2011–2012 N-Counts and Percents by Ethnicity and Grade for ELA Participation

Grade	Native Hawaiian or Other Pacific Islander		American Indian or Alaska Native		Black or African American		Hispanic of any Race		White		Two or more races		Asian		Total
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N
3	0	0.0	1	0.2	97	23.8	24	5.9	263	64.5	8	2.0	15	3.7	408
4	0	0.0	1	0.2	89	22.4	27	6.8	262	66.0	4	1.0	14	3.5	397
5	0	0.0	0	0.0	74	21.1	26	7.4	234	66.9	7	2.0	9	2.6	350
6	0	0.0	3	0.9	90	28.3	18	5.7	191	60.1	6	1.9	10	3.1	318
7	0	0.0	2	0.6	83	25.9	17	5.3	207	64.7	4	1.2	7	2.2	320
8	0	0.0	0	0.0	77	24.0	17	5.3	216	67.3	2	0.6	9	2.8	321
11	0	0.0	2	0.7	90	29.8	20	6.6	180	59.6	2	0.7	8	2.6	302

Table 2.8

2011–2012 N-Counts and Percents by Ethnicity and Grade for ELA Supported Independence

Grade	Native Hawaiian or Other Pacific Islander		American Indian or Alaska Native		Black or African American		Hispanic of any Race		White		Two or more races		Asian		Total
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N
3	0	0.0	4	0.8	119	25.3	28	5.9	304	64.5	8	1.7	8	1.7	471
4	0	0.0	4	0.8	127	26.2	28	5.8	317	65.5	3	0.6	5	1.0	484
5	0	0.0	3	0.6	134	26.4	30	5.9	322	63.4	6	1.2	13	2.6	508
6	2	0.4	3	0.6	119	24.3	33	6.7	320	65.4	8	1.6	4	0.8	489
7	2	0.4	2	0.4	115	23.7	20	4.1	323	66.5	8	1.6	16	3.3	486
8	2	0.4	5	1.0	115	22.8	22	4.4	340	67.3	10	2.0	11	2.2	505
11	0	0.0	3	0.5	149	25.8	33	5.7	379	65.6	8	1.4	6	1.0	578

Table 2.9

2011–2012 N-Counts and Percents by Ethnicity and Grade for Mathematics Participation

Grade	Native Hawaiian or Other Pacific Islander		American Indian or Alaska Native		Black or African American		Hispanic of any Race		White		Two or more races		Asian		Total
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N
3	0	0.0	1	0.2	95	23.5	24	5.9	261	64.6	8	2.0	15	3.7	404
4	0	0.0	1	0.3	88	22.2	27	6.8	262	66.2	4	1.0	14	3.5	396
5	0	0.0	0	0.0	84	20.8	19	7.5	209	67.1	7	2.0	9	2.6	346
6	0	0.0	3	0.9	89	28.2	18	5.7	190	60.1	6	1.9	10	3.2	316
7	0	0.0	2	0.6	83	25.9	17	5.3	207	64.7	4	1.2	7	2.2	320
8	0	0.0	0	0.0	77	24.0	17	5.3	216	67.3	2	0.6	9	2.8	321
11	0	0.0	2	0.7	90	29.8	20	6.6	180	59.6	2	0.7	8	2.6	302

Table 2.10
2011–2012 N-Counts and Percents by Ethnicity and Grade for Mathematics
Supported Independence

Grade	Native Hawaiian or Other Pacific Islander		American Indian or Alaska Native		Black or African American		Hispanic of any Race		White		Two or more races		Asian		Total
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N
3	0	0.0	4	0.8	119	25.3	27	5.7	305	64.8	8	1.7	8	1.7	471
4	0	0.0	4	0.8	127	26.4	27	5.6	315	65.6	3	0.6	5	1.0	481
5	0	0.0	3	0.6	133	26.3	30	5.9	321	63.4	6	1.2	13	2.6	506
6	2	0.4	3	0.6	116	23.8	33	6.8	321	65.9	8	1.6	4	0.8	487
7	2	0.4	2	0.4	115	23.8	20	4.1	321	66.3	8	1.7	16	3.3	484
8	2	0.4	5	1.0	114	22.6	21	4.2	341	67.7	10	2.0	11	2.2	504
11	0	0.0	3	0.5	147	25.6	33	5.7	377	65.7	8	1.4	6	1.0	574

Table 2.11
2011–2012 N-Counts and Percents by Ethnicity and Grade for Science Participation

Grade	Native Hawaiian or Other Pacific Islander		American Indian or Alaska Native		Black or African American		Hispanic of any Race		White		Two or more races		Asian		Total
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N
5	0	0.0	0	0.0	73	21.4	25	7.3	227	66.6	7	2.1	9	2.6	341
8	0	0.0	0	0.0	73	23.5	18	5.8	209	67.4	2	0.6	8	2.6	310
11	0	0.0	2	0.7	89	29.7	20	6.7	179	59.7	2	0.7	8	2.7	300

Table 2.12
2011–2012 N-Counts and Percents by Ethnicity and Grade for Science
Supported Independence

Grade	Native Hawaiian or Other Pacific Islander		American Indian or Alaska Native		Black or African American		Hispanic of any Race		White		Two or more races		Asian		Total
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N
5	0	0.0	2	0.4	132	26.7	31	6.3	311	62.8	6	1.2	13	2.6	495
8	2	0.4	5	1.0	114	23.0	25	5.1	329	66.5	10	2.0	10	2.0	495
11	0	0.0	3	0.5	150	26.2	31	5.4	374	65.4	8	1.4	6	1.0	572

Form Distribution

Recall from Section 1, three test forms were developed for each level and content area. 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 Assessment. Except for Detroit, each district received up to two forms at a grade. For Detroit, each school received the same form at a grade. Table 2.13 shows the materials ordered versus the number of forms that were used to test the students based on the sampling plan. The number of materials ordered usually exceeded the number tested by a somewhat considerable margin. In general, the number of students tested was fairly similar across forms.

The percent of students by various subgroups and form for the 2011 – 2012 school year are given in the tables on the following pages. 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. For Participation, the percents for the three content areas are given in Tables 2.14 – 2.16. For Supported Independence, the percents for the three content areas are given in Tables 2.17 – 2.19. The subgroups consist of gender, three racial/ethnic groups (Black, Hispanic, and White), and Economically Disadvantaged. As seen from the tables for both levels and all three content areas, each form is usually well represented by the various subgroups. Moreover, for each form at a grade, the percent of students across the subgroups was generally consistent with the percents for the grade population.

Table 2.13

2011–2012 Material Orders versus Test Forms Used

Assessment	Ordered	Tested (Participation)	Tested (Supported Independence)
Participation/Supported Independence ELA /Mathematics Grades 3-5 Form 1	1451	409	508
Participation/Supported Independence ELA /Mathematics Grades 3-5 Form 2	1454	401	512
Participation/Supported Independence ELA /Mathematics Grades 3-5 Form 3	1470	406	515
Participation/Supported Independence ELA/Mathematics Grades 6-8 Form 1	1322	310	495
Participation/Supported Independence ELA/Mathematics Grades 6-8 Form 1	1312	342	523
Participation/Supported Independence ELA/Mathematics Grades 6-8 Form 1	1360	354	523
Participation ELA/Mathematics Grade 11 Form 1	170	109	
Participation ELA/Mathematics Grade 11 Form 2	177	101	
Participation ELA/Mathematics Grade 11 Form 3	169	114	
Supported Independence ELA/Mathematics Grade 11 Form 1	276		227
Supported Independence ELA/Mathematics Grade 11 Form 2	276		189
Supported Independence ELA/Mathematics Grade 11 Form 3	275		192
Participation/Supported Independence Science Grade 5 Form 1	539	112	
Participation/Supported Independence Science Grade 5 Form 2	520	146	
Participation/Supported Independence Science Grade 5 Form 3	477	95	
Participation/Supported Independence Science Grade 8 Form 1	504		114
Participation/Supported Independence Science Grade 8 Form 2	455		107
Participation/Supported Independence Science Grade 8 Form 3	453		103
Participation Science Grade 11 Form 1	167	115	
Participation Science Grade 11 Form 2	167	104	
Participation Science 11 Form 3	168	103	
Supported Independence Science Grade 11 Form 1	279		213
Supported Independence Science Grade 11 Form 2	279		186
Supported Independence Science Grade 11 Form 3	279		203

* Participation and Supported Independence material orders were combined in Fall 2011. There were separate orders in Spring 2012. Participation and Supported Independence ELA and Mathematics items appear in the same test booklet. Participation and Supported Independence Science items appear in the same test booklet.

Table 2.14

2011–2012 Percent of Students by Subgroup and Form ELA Participation

	N	Female	Male	Black	Hispanic	White	Economic Disadv
Grade 3							
All Forms	408	33.1	66.9	23.8	5.9	64.5	51.0
Form 1	129	31.8	68.2	22.5	4.7	67.4	51.2
Form 2	150	30.0	70.0	30.7	6.7	56.7	57.3
Form 3	129	38.0	62.0	17.1	6.2	70.5	43.4
Grade 4							
All Forms	397	34.5	65.5	22.4	6.8	66.0	55.4
Form 1	136	37.5	62.5	21.3	9.6	63.2	55.9
Form 2	124	32.3	67.7	24.2	4.8	65.3	53.2
Form 3	137	33.6	66.4	21.9	5.8	69.3	56.9
Grade 5							
All Forms	350	33.1	66.9	21.1	7.4	66.9	53.4
Form 1	120	40.0	60.0	30.0	5.8	61.7	56.7
Form 2	104	21.2	78.8	18.3	7.7	67.3	52.9
Form 3	126	36.5	63.5	15.1	8.7	71.4	50.8
Grade 6							
All Forms	318	35.5	64.5	28.3	5.7	60.1	50.4
Form 1	99	40.4	59.6	31.3	5.1	58.6	43.4
Form 2	114	35.1	64.9	28.9	4.4	58.8	50.9
Form 3	105	31.4	68.6	24.8	7.6	62.9	56.2
Grade 7							
All Forms	320	39.1	60.9	25.9	5.3	64.7	49.7
Form 1	105	37.1	62.9	30.5	2.9	63.8	52.4
Form 2	103	48.5	51.5	29.1	8.7	58.8	53.4
Form 3	112	32.1	67.9	18.8	4.5	71.4	43.8
Grade 8							
All Forms	321	36.4	63.6	24.0	5.3	67.3	44.2
Form 1	88	36.4	63.6	26.1	4.5	68.2	43.2
Form 2	112	37.5	62.5	23.2	8.9	61.6	46.4
Form 3	121	35.5	64.5	23.1	2.5	71.9	43.0
Grade 11							
All Forms	302	38.7	61.3	29.8	6.6	59.6	57.9
Form 1	103	42.7	57.3	26.2	7.8	62.1	64.1
Form 2	89	37.1	62.9	28.1	7.9	59.6	57.3
Form 3	110	36.4	63.6	34.5	4.5	57.3	52.7

Table 2.15

2011–2012 Percent of Students by Subgroup and Form

Mathematics Participation

	N	Female	Male	Black	Hispanic	White	Economic Disadv
Grade 3							
All Forms	404	33.2	66.8	23.5	5.9	64.6	51.2
Form 1	129	31.8	68.2	22.5	4.7	67.4	51.2
Form 2	148	29.7	70.3	30.4	6.8	56.8	58.1
Form 3	127	38.6	61.4	16.5	6.3	70.9	43.3
Grade 4							
All Forms	396	34.3	65.7	22.2	6.8	66.2	55.6
Form 1	135	37.0	63.0	20.7	9.6	63.7	56.3
Form 2	124	32.3	67.7	24.2	4.8	65.3	53.2
Form 3	137	33.6	66.4	21.9	5.8	69.3	56.9
Grade 5							
All Forms	346	33.2	66.8	20.8	7.5	67.1	53.8
Form 1	118	39.8	60.2	28.8	5.9	62.7	55.9
Form 2	103	21.4	78.6	18.4	7.8	67.0	53.4
Form 3	125	36.8	63.2	15.2	8.8	71.2	52.0
Grade 6							
All Forms	316	35.8	64.2	28.2	5.7	60.1	50.4
Form 1	98	40.8	59.2	31.6	5.1	58.2	43.9
Form 2	114	35.1	64.9	28.9	4.4	58.8	50.9
Form 3	104	31.7	68.3	24.0	7.7	63.5	55.8
Grade 7							
All Forms	320	39.1	60.9	25.9	5.3	64.7	49.7
Form 1	105	37.1	62.9	30.5	2.9	63.8	52.4
Form 2	103	48.5	51.5	29.1	8.7	58.3	53.4
Form 3	112	32.1	67.9	18.8	4.5	71.4	43.8
Grade 8							
All Forms	321	36.4	63.6	24.0	5.3	67.3	44.2
Form 1	87	36.8	63.2	25.3	4.6	69.0	42.5
Form 2	113	37.2	62.8	23.9	8.8	61.1	46.9
Form 3	121	35.5	64.5	23.1	2.5	71.9	43.0
Grade 11							
All Forms	302	38.7	61.3	29.8	6.6	59.6	57.9
Form 1	103	42.7	57.3	26.2	7.8	62.1	64.1
Form 2	89	37.1	62.9	28.1	7.9	59.6	57.3
Form 3	110	36.4	63.6	34.5	4.5	57.3	52.7

Table 2.16
2011–2012 Percent of Students by Subgroup and Form
Science Participation

	N	Female	Male	Black	Hispanic	White	Economic Disadv
Grade 5							
All Forms	341	33.4	66.6	21.4	7.3	66.6	53.4
Form 1	106	30.2	69.8	13.2	6.6	76.4	57.5
Form 2	143	37.8	62.2	30.8	4.9	60.1	54.5
Form 3	92	30.4	69.6	16.3	12.0	65.2	46.7
Grade 8							
All Forms	310	37.7	62.3	23.5	5.8	67.4	44.8
Form 1	112	38.4	61.6	22.3	6.2	67.9	46.4
Form 2	100	40.0	60.0	25.0	7.0	65.0	47.0
Form 3	98	34.7	65.3	23.5	4.1	69.4	40.8
Grade 11							
All Forms	300	38.7	61.3	29.7	6.7	59.7	57.7
Form 1	111	33.3	66.7	22.5	5.4	66.7	55.0
Form 2	89	41.6	58.4	28.1	6.7	61.8	58.4
Form 3	100	42.0	58.0	39.0	8.0	50.0	60.0

Table 2.17
2011–2012 Percent of Students by Subgroup and Form
ELA Supported Independence

	N	Female	Male	Black	Hispanic	White	Economic Disadv
Grade 3							
All Forms	471	31.6	68.4	25.3	5.9	64.5	57.7
Form 1	160	28.1	71.9	27.5	5.0	61.9	55.0
Form 2	144	29.9	70.1	23.6	6.2	64.6	54.9
Form 3	167	36.5	63.5	24.6	6.6	67.1	62.9
Grade 4							
All Forms	484	29.8	70.2	26.2	5.8	65.5	58.3
Form 1	165	35.2	64.8	29.1	3.0	64.8	57.9
Form 2	167	26.9	73.1	26.9	6.0	65.3	56.9
Form 3	152	27.0	73.0	22.4	8.6	66.4	61.2
Grade 5							
All Forms	508	30.3	69.7	26.4	5.9	63.4	60.2
Form 1	168	33.9	66.1	33.3	4.8	58.3	60.7
Form 2	175	26.9	73.1	24.0	6.9	67.4	57.7
Form 3	165	30.3	69.7	21.8	6.1	64.2	62.4
Grade 6							
All Forms	489	33.3	66.7	24.3	6.7	65.4	56.9
Form 1	161	33.5	66.5	23.0	3.1	71.4	52.2
Form 2	157	29.3	70.7	26.1	9.6	61.8	65.6
Form 3	171	36.8	63.2	24.0	7.6	63.2	53.2
Grade 7							
All Forms	486	36.4	63.6	23.7	4.1	66.5	60.3
Form 1	164	34.8	65.2	24.4	3.7	64.6	56.7
Form 2	170	35.9	64.1	24.1	2.4	67.6	52.9
Form 3	152	38.8	61.2	22.4	6.6	67.1	61.2
Grade 8							
All Forms	505	35.4	64.6	22.8	4.4	67.3	58.4
Form 1	149	34.9	65.1	20.1	5.4	69.8	61.1
Form 2	174	36.2	63.8	23.0	3.4	64.9	59.2
Form 3	182	35.2	64.8	24.7	4.4	67.6	55.5
Grade 11							
All Forms	578	35.8	64.2	25.8	5.7	65.6	55.9
Form 1	217	35.0	65.0	28.1	1.4	66.4	52.5
Form 2	182	40.1	59.9	26.4	6.0	64.3	55.5
Form 3	179	32.4	67.6	22.3	10.6	65.9	60.3

Table 2.18
2011–2012 Percent of Students by Subgroup and Form
Mathematics Supported Independence

	N	Female	Male	Black	Hispanic	White	Economic Disadv
Grade 3							
All Forms	471	31.4	68.6	25.3	5.7	64.8	57.5
Form 1	160	28.1	71.9	27.5	5.0	61.9	55.0
Form 2	144	29.9	70.1	23.6	6.2	64.6	54.9
Form 3	167	35.9	64.1	24.6	6.0	67.7	62.3
Grade 4							
All Forms	481	29.7	70.3	26.4	5.6	65.5	58.2
Form 1	164	35.4	64.6	29.3	3.0	64.6	57.3
Form 2	166	26.5	73.5	27.1	5.4	65.7	56.6
Form 3	151	27.2	72.8	22.5	8.6	66.2	60.9
Grade 5							
All Forms	506	30.4	69.6	26.3	5.9	63.4	60.5
Form 1	168	33.9	66.1	33.3	4.8	58.3	60.7
Form 2	174	27.0	73.0	24.1	6.9	67.2	58.0
Form 3	164	30.5	69.5	21.3	6.1	64.6	62.8
Grade 6							
All Forms	487	33.3	66.7	23.8	6.8	65.9	56.5
Form 1	159	33.3	66.7	21.4	3.1	73.0	50.9
Form 2	157	29.3	70.7	26.1	9.6	61.8	65.6
Form 3	171	36.8	63.2	24.0	7.6	63.2	53.2
Grade 7							
All Forms	484	36.4	63.6	23.8	4.1	66.3	60.3
Form 1	162	34.6	65.4	24.7	3.7	64.2	56.8
Form 2	170	35.9	64.1	24.1	2.4	67.6	62.9
Form 3	152	38.8	61.2	22.4	6.6	67.1	61.2
Grade 8							
All Forms	504	35.5	64.5	22.6	4.2	67.7	57.9
Form 1	149	34.9	65.1	20.1	4.7	70.5	59.7
Form 2	173	36.4	63.6	22.5	3.5	65.3	59.0
Form 3	182	35.2	64.8	24.7	4.4	67.6	55.5
Grade 11							
All Forms	574	36.2	63.8	25.6	5.7	65.7	55.6
Form 1	214	36.0	64.0	27.6	1.4	66.8	51.9
Form 2	182	40.1	59.9	26.4	6.0	64.3	55.5
Form 3	178	32.6	67.4	22.5	10.7	65.7	60.1

Table 2.19
2011–2012 Percent of Students by Subgroup and Form
Science Supported Independence

	N	Female	Male	Black	Hispanic	White	Economic Disadv
Grade 5							
All Forms	495	30.9	69.1	26.7	6.3	62.8	60.2
Form 1	171	30.4	69.6	25.1	3.5	68.4	55.0
Form 2	186	26.9	73.1	29.0	5.9	60.2	65.6
Form 3	138	37.0	63.0	25.4	10.1	59.4	59.4
Grade 8							
All Forms	495	36.2	63.8	23.0	5.1	66.5	59.0
Form 1	180	33.3	66.7	24.4	5.0	62.8	58.9
Form 2	158	38.0	62.0	25.3	4.4	67.7	65.2
Form 3	157	37.6	62.4	19.1	5.7	69.4	52.9
Grade 11							
All Forms	572	36.0	64.0	26.2	5.4	65.4	56.1
Form 1	200	36.0	64.0	28.5	2.5	64.5	50.5
Form 2	178	38.2	61.8	17.4	4.5	77.0	53.9
Form 3	194	34.0	66.0	32.0	9.3	55.7	63.9

Subgroup Analysis

Summary statistics by gender and grade for all three content areas for Participation and Supported Independence are given in Tables 2.20 to 2.25, and the summary statistics by race/ethnicity and grade for the three content areas for Participation and Supported Independence are given in Tables 2.26 to 2.31. Summary statistics by race/ethnicity are given for White students, Black students, and Other students. Excluding White and Black students, the Other students category combines all other ethnic subgroups but consists primarily of Hispanic students.

For ELA Participation, males' average scores were higher than females in grades 3 – 5, 7, and 8, and lower at grades 6 and 11. For ELA Supported Independence, females' average scores were higher than males at all grades except grade 6. For Mathematics Participation, males' average scores were higher than females at all grades except grade 6. For Mathematics Supported Independence, males' average scores were higher than females at grades 3 and 6 – 8, and lower at grades 4, 5 and 11. For Science Participation, males' average scores were higher than females at grades 5 and 8 and lower at grade 11. For Science Supported Independence, females' average scores were higher than males at all three grades.

For Participation and Supported Independence ELA, the black students' average scores were higher than the White students except at grades 4 and 5 for Participation and grades 4, 5, and 7 for Supported Independence. For Participation and Supported Independence Mathematics, Black students' average scores were higher than White students except at grade 5 Participation and grade 7 Supported Independence. For Participation and Supported Independence Science, Black students' average scores were higher at grades 8 and 11 and lower at grade 5. Students from the Other ethnic subgroups sometimes had the average scores that were higher, in between, or lower than White and Black students depending on the grade, test level, and content area.

Table 2.20
2011–2012 Operational Form Score Summaries by Gender and Grade – ELA Participation

	Female			Male		
Grade	Mean	SD	N	Mean	SD	N
3	25.52	17.44	135	26.64	16.94	273
4	24.68	17.51	137	26.00	16.40	260
5	27.28	17.84	116	28.03	16.68	234
6	26.04	17.64	113	25.43	18.45	205
7	28.15	19.18	125	28.49	16.86	195
8	27.42	19.97	117	28.72	19.00	204
11	31.75	18.01	117	30.06	17.33	185

Table 2.21
2011–2012 Operational Form Score Summaries by Gender and Grade – ELA Supported Independence

	Female			Male		
Grade	Mean	SD	N	Mean	SD	N
3	39.20	12.85	149	38.77	12.67	322
4	41.99	12.81	144	39.09	14.08	340
5	43.64	13.36	154	42.14	12.61	354
6	39.28	14.16	163	40.67	13.87	326
7	42.48	12.85	177	39.23	14.29	309
8	43.72	11.89	179	41.38	14.30	326
11	42.90	14.25	207	40.92	14.75	371

Table 2.22
2011–2012 Operational Form Score Summaries by Gender and Grade – Mathematics Participation

	Female			Male		
Grade	Mean	SD	N	Mean	SD	N
3	25.39	19.45	134	28.35	18.43	270
4	25.42	18.23	136	28.66	17.29	260
5	26.53	19.06	115	30.45	17.25	231
6	25.26	17.78	113	24.48	18.66	203
7	27.70	18.88	125	27.95	17.30	195
8	26.64	20.60	117	28.03	18.11	204
11	31.62	18.85	117	31.70	17.72	185

Table 2.23
2011–2012 Operational Form Score Summaries by Gender and Grade –
Mathematics Supported Independence

	Female			Male		
Grade	Mean	SD	N	Mean	SD	N
3	37.05	15.31	148	37.80	13.16	323
4	40.28	13.10	143	37.83	14.67	338
5	41.51	13.55	154	40.58	13.48	352
6	33.91	12.57	162	36.24	13.27	325
7	36.02	14.11	176	36.59	13.79	308
8	38.05	13.08	179	38.42	13.58	325
11	40.40	12.97	208	39.20	13.72	366

Table 2.24
2011–2012 Operational Form Score Summaries by Gender and Grade –
Science Participation

	Female			Male		
Grade	Mean	SD	N	Mean	SD	N
5	40.46	28.14	114	43.59	26.30	227
8	41.38	30.74	117	44.68	26.74	193
11	46.17	28.16	116	44.21	27.78	184

Table 2.25
2011–2012 Operational Form Score Summaries by Gender and Grade –
Science Supported Independence

	Female			Male		
Grade	Mean	SD	N	Mean	SD	N
5	50.07	13.34	153	49.36	13.79	342
8	47.61	12.76	179	45.93	14.42	316
11	49.04	14.22	206	47.72	15.70	366

Table 2.26
2011–2012 Operational Form Score Summaries by Ethnicity and Grade – ELA Participation

	White			Black			Other		
Grade	Mean	SD	N	Mean	SD	N	Mean	SD	N
3	25.29	17.01	263	27.94	17.94	97	27.27	15.70	48
4	26.06	16.16	223	24.01	17.17	77	25.59	19.51	46
5	28.53	17.06	234	28.23	16.41	74	22.83	17.69	42
6	25.33	17.96	191	25.58	18.57	90	27.49	18.38	37
7	28.17	17.46	207	29.65	18.64	83	26.03	17.79	30
8	27.96	19.71	216	28.00	18.61	77	31.07	18.84	28
11	30.09	17.82	180	33.97	16.68	90	25.12	17.49	32

Table 2.27
2011–2012 Operational Form Score Summaries by Ethnicity and Grade – ELA Supported Independence

	White			Black			Other		
Grade	Mean	SD	N	Mean	SD	N	Mean	SD	N
3	38.30	12.72	304	40.08	13.13	119	39.83	11.61	48
4	40.33	13.74	317	39.77	12.99	127	37.52	16.24	40
5	42.87	12.12	322	41.62	14.00	134	43.38	14.21	52
6	39.25	14.15	320	41.78	13.98	119	42.52	12.36	50
7	40.68	13.97	323	38.35	14.16	115	43.54	11.76	48
8	42.16	13.42	340	43.13	14.19	115	40.38	12.80	50
11	41.34	15.01	379	42.81	13.19	149	40.30	15.42	50

Table 2.28
2011–2012 Operational Form Score Summaries by Ethnicity and Grade – Mathematics Participation

	White			Black			Other		
Grade	Mean	SD	N	Mean	SD	N	Mean	SD	N
3	26.38	18.74	261	29.66	19.41	95	28.19	17.79	48
4	27.53	17.20	262	28.61	18.34	88	25.59	19.11	46
5	30.09	17.83	232	28.46	17.83	72	25.14	18.58	42
6	23.87	18.09	190	24.47	19.00	89	30.03	17.45	37
7	27.52	17.73	207	30.01	18.22	83	24.13	18.05	30
8	27.26	19.25	216	27.49	19.15	77	29.64	17.49	28
11	30.74	18.17	180	35.80	17.51	90	25.28	17.63	32

Table 2.29
2011–2012 Operational Form Score Summaries by Ethnicity and Grade – Mathematics Supported Independence

	White			Black			Other		
Grade	Mean	SD	N	Mean	SD	N	Mean	SD	N
3	36.75	13.94	305	39.75	14.37	119	37.34	11.49	47
4	38.61	13.98	315	39.66	13.76	127	34.54	17.37	39
5	40.67	12.90	321	40.95	14.88	133	41.85	13.40	52
6	35.21	12.85	321	35.81	13.98	116	36.36	12.54	50
7	36.78	13.87	321	33.14	14.14	115	41.56	11.67	48
8	38.25	13.59	341	39.13	13.32	114	36.61	12.20	49
11	39.12	13.89	377	41.19	12.30	147	38.94	13.22	50

Table 2.30
2011–2012 Operational Form Score Summaries by Ethnicity and Grade – Science Participation

	White			Black			Other		
Grade	Mean	SD	N	Mean	SD	N	Mean	SD	N
5	44.89	26.67	227	38.74	27.53	73	36.27	26.13	41
8	42.71	28.71	209	44.49	27.91	73	50.11	26.23	28
11	43.08	27.78	179	51.88	27.39	89	36.28	26.61	32

Table 2.31
2011–2012 Operational Form Score Summaries by Ethnicity and Grade – Science Supported Independence

	White			Black			Other		
Grade	Mean	SD	N	Mean	SD	N	Mean	SD	N
5	49.59	13.18	311	49.21	14.45	132	50.48	14.46	52
8	46.63	13.66	329	47.22	13.99	114	44.44	14.85	52
11	47.66	15.96	374	50.42	12.58	150	45.40	15.89	48

3. Score Reliability and Summary Statistics

Internal consistency reliability estimates were computed as Cronbach's Coefficient Alpha (Cronbach, 1951) for the operational forms in ELA, Mathematics, and Science. These reliability estimates are reported in Tables 3.1 to 3.6. Across all grades for all three content areas of Participation, the estimated reliabilities ranged from the mid .80s to the low .90s. Across all grades for all three content areas of Supported Independence, the estimated reliabilities ranged from the low .80s to the high .80s. These estimated reliabilities indicated a high degree of internal consistency across all 34 assessments.

The standard error of measurement (SEM) for each assessment is also reported in these tables as well as the N, mean, and standard deviation. The standard deviations for all the assessments were larger than one might expect, especially for the Participation assessments. This is because a few to several percent of the students obtained a valid, earned zero score, and likewise, a few to several percent of the students received a perfect score. Due to the large number of students obtaining each extreme score, the standard deviations were much larger than is often observed on typical assessments. Due to the large standard deviations, the corresponding SEMs were also quite large despite the high reliabilities.

Table 3.1
2011–2012 Operational Form Summaries, including Sample Size, Score Statistics, Coefficient Alpha, and SEM by Grade – ELA Participation

Grade	N	Mean	SD	Cronbach's Alpha	SEM
3	408	26.15	17.09	0.86	6.36
4	397	25.55	16.78	0.85	6.53
5	350	27.79	17.05	0.85	6.57
6	318	25.65	18.14	0.88	6.37
7	320	28.36	17.77	0.87	6.52
8	321	28.24	19.34	0.90	6.23
11	302	30.72	17.59	0.88	6.16

Table 3.2

2011–2012 Operational Form Summaries, including Sample Size, Score Statistics, Coefficient Alpha, and SEM by Grade – ELA Supported Independence

Grade	N	Mean	SD	Cronbach's Alpha	SEM
3	471	38.91	12.72	0.81	5.53
4	484	39.95	13.76	0.85	5.30
5	508	42.59	12.85	0.84	5.13
6	489	40.20	13.97	0.86	5.17
7	486	40.41	13.86	0.86	5.21
8	505	42.21	13.53	0.86	5.03
11	578	41.63	14.59	0.87	5.18

Table 3.3

2011–2012 Operational Form Summaries, including Sample Size, Score Statistics, Coefficient Alpha, and SEM by Grade – Mathematics Participation

Grade	N	Mean	SD	Cronbach's Alpha	SEM
3	404	27.37	18.80	0.90	5.85
4	396	27.55	17.66	0.88	6.13
5	346	29.15	17.94	0.88	6.12
6	316	24.76	18.32	0.88	6.23
7	320	27.85	17.90	0.87	6.43
8	321	27.52	19.04	0.89	6.21
11	302	31.67	18.14	0.88	6.28

Table 3.4

2011–2012 Operational Form Summaries, including Sample Size, Score Statistics, Coefficient Alpha, and SEM by Grade – Mathematics Supported Independence

Grade	N	Mean	SD	Cronbach's Alpha	SEM
3	471	37.56	13.86	0.86	5.20
4	481	38.56	14.25	0.87	5.08
5	506	40.86	13.48	0.86	4.98
6	487	35.47	13.08	0.81	5.73
7	484	36.39	13.90	0.84	5.60
8	504	38.29	13.39	0.82	5.62
11	574	39.63	13.45	0.83	5.59

Table 3.5

2011–2012 Operational Form Summaries, including Sample Size, Score Statistics, Coefficient Alpha, and SEM by Grade – Science Participation

Grade	N	Mean	SD	Cronbach's Alpha	SEM
5	341	42.54	26.93	0.91	7.87
8	310	43.44	28.31	0.93	7.52
11	300	44.97	27.90	0.93	7.61

Table 3.6

2011–2012 Operational Form Summaries, including Sample Size, Score Statistics, Coefficient Alpha, and SEM by Grade – Science Supported Independence

Grade	N	Mean	SD	Cronbach's Alpha	SEM
5	495	49.58	13.64	0.82	5.81
8	495	46.54	13.86	0.81	6.01
11	572	48.20	15.19	0.86	5.74

4. Rater Consistency

Each item of the ELA, Mathematics, and Science assessments is scored based on a rubric. There is a 3–point rubric with 3 condition codes at the Participation level and a 2–point rubric with the same 3 condition codes at the Supported Independence level. Each activity has two raters — the Primary Assessment Administrator (PAA) and the Shadow Assessment Administrator (SAA). Items with a missing score for a rater are coded using a special code in the student data file, but for the purposes of scoring, a missing score for a rater is treated as a 0. There are very few item scores that are missing scores for a rater. A student's score is the sum of the item scores across both raters.

The consistency of each double–scored item is summarized in Tables 4.1 to 4.6. Across the 70 ELA Participation items, the percent perfect agreement rates ranged from 94% to 98% with a median of 96.5% across all of the perfect agreement rates in the table. Across the 70 Mathematics Participation items, the percent perfect agreement rates ranged from 94% to 99% with a median of 97% across all of the perfect agreement rates. Across the 45 Science Participation items, the percent perfect agreement rates ranged from 95% to 99% with a median of 97% across all of the perfect agreement rates. Across the 105 ELA Supported Independence items, the percent perfect agreement rates ranged from 95% to 99% with a median of 97% across all of the perfect agreement rates. Across the 105 Mathematics Supported Independence items, the percent perfect agreement rates ranged from 95% to 99% with a median of 97% across all of the perfect agreement rates. Across the 51 Science Supported Independence items, the percent perfect agreement rates ranged from 94% to 99% with a median of 97% across all of the perfect agreement rates.

Table 4.1
2011–2012 Operational Form Percent Perfect Interrater Agreement Rates – Participation ELA

Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 11	
Item #	N = 417	Item #	N = 481	Item #	N = 506	Item #	N = 487	Item #	N = 484	Item #	N = 504	Item #	N = 574
73064	97	73064	97	73064	97	72143	96	72143	97	72143	97	18417	97
73605	94	73605	98	73605	96	72027	94	72027	98	72027	96	18421	97
73070	97	73070	96	73070	97	73028	95	73028	96	73028	98	18422	97
73009	95	73009	97	73009	96	73030	94	73030	95	73030	94	18456	97
72004	96	72004	96	72004	97	73083	94	73083	98	73083	97	18461	95
72060	97	72060	97	72060	97	73081	94	73081	97	73081	96	18463	97
72020	95	72020	95	72020	97	72036	95	72036	96	72036	97	18464	96
73134	97	73134	98	73134	97	73412	94	73412	97	73412	97	18465	96
73406	97	73406	97	73406	96	73416	95	73416	97	73416	96	18468	97
73405	96	73405	96	73405	97	73139	94	73139	96	73139	96	28711	96

Table 4.2

2011–2012 Operational Form Percent Perfect Interrater Agreement Rates – Participation Mathematics

Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 11	
Item #	N = 404	Item #	N = 396	Item #	N = 346	Item #	N = 316	Item #	N = 320	Item #	N = 321	Item #	N = 302
72083	97	72083	96	72083	99	72221	96	72221	96	72221	98	18396	96
72122	97	72122	97	72122	97	72204	96	72204	97	72204	96	18403	96
72118	96	72118	97	72118	97	72138	96	72138	96	72138	96	18407	98
72113	97	72113	97	72113	99	73087	95	73087	96	73087	97	18425	97
72093	96	72093	96	72093	98	73088	95	73088	97	73088	95	18734	97
72090	97	72090	98	72090	99	73211	95	73211	96	73211	97	18735	97
73026	98	73026	97	73026	97	73246	97	73246	98	73246	97	18781	98
73071	96	73071	96	73071	97	73243	96	73243	95	73243	96	18792	97
73072	98	73072	97	73072	97	73244	97	73244	98	73244	95	18831	97
73209	97	73209	98	73209	97	73253	94	73253	96	73253	95	18845	95

Table 4.3

2011–2012 Operational Form Percent Perfect Interrater Agreement Rates – Participation Science

Grade 5		Grade 8		Grade 11	
Item #	N = 341	Item #	N = 310	Item #	N = 300
11132	98	11359	97	14963	97
11138	96	11378	96	14980	97
11141	99	11384	96	14985	96
11096	98	11228	96	14988	96
11092	97	11231	97	14991	97
11022	96	11318	97	15002	96
11076	97	11278	95	15024	98
11075	98	11279	97	15027	96
11063	98	11275	97	15039	97
11060	97	11274	97	15044	95
12035	97	12119	97	15051	98
12109	98	12111	95	15052	96
12104	99	12112	96	15063	97
12101	99	12054	95	15067	97
12100	98	12050	96	15265	98

Table 4.4
2011–2012 Operational Form Percent Perfect Interrater Agreement Rates –
Supported Independence ELA

Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 11	
Item #	N = 471	Item #	N = 484	Item #	N = 508	Item #	N = 489	Item #	N = 486	Item #	N = 505	Item #	N = 578
52002	97	52002	97	52002	97	52070	96	52070	96	52070	96	17649	96
52071	96	52071	98	52071	97	52075	96	52075	97	52075	96	17650	97
52012	96	52012	97	52012	99	52026	98	52026	97	52026	98	17660	97
52017	97	52017	98	52017	99	52196	97	52196	95	52196	96	17667	96
52024	96	52024	98	52024	97	52165	97	52165	97	52165	98	17970	96
52191	98	52191	98	52191	96	54148	97	54148	98	54148	98	17981	95
52182	98	52182	98	52182	98	54114	97	54114	98	54114	97	17990	97
54106	99	54106	98	54106	97	53041	96	53041	98	53041	98	17994	97
54140	98	54140	98	54140	96	53034	97	53034	99	53034	97	18046	95
54138	98	54138	96	54138	99	52247	97	52247	99	52247	98	18050	95
54137	97	54137	97	54137	98	53014	96	53014	97	53014	97	18051	97
54136	98	54136	98	54136	98	53013	97	53013	98	53013	96	18052	95
54133	97	54133	96	54133	97	53079	97	53079	98	53079	97	18056	95
53021	96	53021	96	53021	96	53076	97	53076	98	53076	98	18064	96
53074	97	53074	98	53074	98	53073	96	53073	97	53073	96	28808	96

Table 4.5
2011–2012 Operational Form Percent Perfect Interrater Agreement Rates –
Supported Independence Mathematics

Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 11	
Item #	N = 471	Item #	N = 481	Item #	N = 506	Item #	N = 487	Item #	N = 484	Item #	N = 504	Item #	N = 574
52193	97	52193	98	52193	96	52100	95	52100	95	52100	97	17694	95
52194	96	52194	99	52194	99	52215	96	52215	97	52215	97	17695	97
52164	97	52164	99	52164	98	52151	96	52151	97	52151	98	17697	97
52119	97	52119	98	52119	97	52150	96	52150	96	52150	97	17973	97
52117	97	52117	98	52117	96	54210	97	54210	97	54210	98	17986	97
52104	96	52104	97	52104	96	54208	97	54208	97	54208	98	18003	99
54200	97	54200	98	54200	98	54207	97	54207	96	54207	97	18011	97
54205	99	54205	98	54205	98	53038	97	53038	97	53038	96	18025	97
54201	97	54201	99	54201	98	52239	96	52239	98	52239	98	18027	96
54228	97	54228	99	54228	97	53087	99	53087	99	53087	99	18035	97
54222	96	54222	96	54222	97	53090	96	53090	98	53090	98	18040	98
53024	97	53024	98	53024	98	53089	97	53089	96	53089	96	24653	96
52241	98	52241	98	52241	97	54229	95	54229	96	54229	97	24654	97
53083	99	53083	99	53083	98	53084	97	53084	98	53084	97	28809	96
53082	97	53082	96	53082	96	54253	97	54253	98	54253	98	28948	97

Table 4.6

2011–2012 Operational Form Percent Perfect Interrater Agreement Rates – Supported Independence Science

Grade 5		Grade 8		Grade 11	
Item #	N = 495	Item #	N = 495	Item #	N = 511
22017	98	22228	95	15568	96
22005	99	22227	97	15570	98
22024	99	22203	97	15575	97
22172	97	22266	98	15626	97
22189	99	22269	96	15649	97
22161	98	22283	97	15659	97
22100	98	22280	96	15782	95
22055	99	22378	97	15797	98
22070	99	22488	97	15872	96
22145	98	22390	98	15830	98
22150	99	22310	97	15834	94
22122	98	22297	97	15836	98
23100	98	22326	96	15839	94
23106	97	23111	98	15848	97
23139	98	23112	98	15855	96
23143	98	23152	98	15857	97
23142	98	23118	98	15862	97

5. Classification Accuracy and Consistency

The number and percent of students at each of the three performance levels by content area and grade are given in Table 5.1 for Participation and in Table 5.2 for Supported Independence. Across grades and content areas of Participation, the largest percentage of students were classified at the *Attained* performance level except for grades 6 and 8 ELA, grades 3 and 5 – 8 Mathematics, and grade 11 Science. In these cases, more students were in the *Emerging* performance level. For Supported Independence ELA and Mathematics, the largest percentage of students was at the *Surpassed* performance level except for grades 5 ELA and grades 5, 7, and 11 for Mathematics. In those grades, *Attained* had the greatest percentage of students. For Supported Independence Science, the *Attained* performance level had the highest percentage. Across both Participation and Supported Independence, more of the students were classified as being proficient than not proficient as more examinees were in the *Attained* and *Surpassed* performance levels. These two performance levels are counted as proficient for the purposes of adequate yearly progress (AYP). The one exception is Science Participation grade 11 where 53% of students were classified as being *Emerging.*

Table 5.1
Percent of Students by Proficiency Level - Participation

Grade	Emerging		Attained		Surpassed	
	N	%	N	%	N	%
ELA						
3	150	36.8	188	46.1	70	17.2
4	159	40.1	176	44.3	62	15.6
5	127	36.3	166	47.4	57	16.3
6	125	39.3	123	38.7	70	22.0
7	107	33.4	127	39.7	86	26.9
8	125	38.9	117	36.4	79	24.6
11	77	25.5	151	50.0	74	24.5
Math						
3	142	35.1	111	27.5	151	37.4
4	140	35.4	153	38.6	103	26.0
5	134	38.7	126	36.4	86	24.9
6	149	47.2	103	32.6	64	20.3
7	129	40.3	124	38.8	67	20.9
8	159	49.5	111	34.6	51	15.9
11	113	37.4	134	44.4	55	18.2
Science						
5	97	28.4	196	57.5	48	14.1
8	95	30.6	174	56.1	41	13.2
11	159	53.0	102	34.0	39	13.0

Table 5.2

Percent of Students by Proficiency Level - Supported Independence

Grade	Emerging		Attained		Surpassed	
	N	%	N	%	N	%
ELA						
3	60	12.7	197	41.8	214	45.4
4	108	22.3	186	38.4	190	39.3
5	109	21.5	200	39.4	199	39.2
6	65	13.3	154	31.5	270	55.2
7	97	20.0	163	33.5	226	46.5
8	92	18.2	189	37.4	224	44.4
11	114	19.7	207	35.8	257	44.5
Math						
3	54	11.5	196	41.6	211	46.9
4	77	16.0	187	38.9	217	45.1
5	77	15.2	229	45.3	200	39.5
6	40	8.2	203	41.7	244	50.1
7	56	11.6	233	48.1	195	40.3
8	44	8.7	228	45.2	232	46.0
11	58	10.1	291	50.7	225	39.2
Science						
5	83	16.8	305	61.6	107	21.6
8	133	26.9	240	48.5	122	24.6
11	96	16.8	381	66.6	95	16.6

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 Participation and Supported Independence classification accuracy and consistency indices were calculated by applying the classification accuracy and consistency indices from Livingston and Lewis (1995) using the BB–CLASS computer program (Brennan, 2004). The BB-CLASS program takes in the reliability estimates for the assessments, the distribution of examinee scores, and the cut scores and computes the Livingston and Lewis indices from these data assuming a beta binomial model. These indices are presented in Tables 5.3 and 5.4. 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 5.3

Estimated Classification Accuracy and Consistency by Content Area and Grade – Participation

	2 Categories Emerging vs Attained plus Surpassed		3 Categories Emerging vs Attained vs Surpassed	
Grade	Accuracy	Consistency	Accuracy	Consistency
ELA				
3	0.89	0.85	0.81	0.74
4	0.88	0.83	0.80	0.72
5	0.89	0.84	0.78	0.71
6	0.90	0.86	0.78	0.71
7	0.90	0.85	0.78	0.70
8	0.91	0.87	0.81	0.74
11	0.91	0.87	0.82	0.74
Mathematics				
3	0.91	0.88	0.83	0.77
4	0.90	0.86	0.81	0.74
5	0.90	0.86	0.80	0.72
6	0.90	0.86	0.80	0.73
7	0.89	0.85	0.80	0.72
8	0.90	0.87	0.80	0.73
11	0.90	0.86	0.78	0.70
Science				
5	0.92	0.89	0.84	0.78
8	0.93	0.90	0.85	0.79
11	0.92	0.89	0.85	0.79

Table 5.4

Estimated Classification Accuracy and Consistency by Subject and Grade – Supported Independence

	2 Categories Emerging vs Attained plus Surpassed		3 Categories Emerging vs Attained vs Surpassed	
Grade	Accuracy	Consistency	Accuracy	Consistency
ELA				
3	0.93	0.89	0.79	0.71
4	0.91	0.87	0.79	0.71
5	0.91	0.87	0.77	0.68
6	0.94	0.91	0.83	0.76
7	0.92	0.89	0.81	0.74
8	0.92	0.89	0.81	0.73
11	0.92	0.89	0.82	0.74
Mathematics				
3	0.94	0.92	0.83	0.76
4	0.93	0.91	0.82	0.75
5	0.94	0.91	0.82	0.75
6	0.95	0.92	0.81	0.74
7	0.93	0.91	0.80	0.72
8	0.94	0.92	0.81	0.73
11	0.94	0.92	0.81	0.73
Science				
5	0.91	0.87	0.78	0.70
8	0.89	0.84	0.76	0.67
11	0.92	0.89	0.76	0.69

Across all grades for all three content areas of Participation, when categorizing students into the NCLB categories of proficient (*Attained* + *Surpassed*) and not proficient (*Emerging*), classification accuracy ranged from 88% to 93% while classification consistency ranged from 83% to 90%. When categorizing students into three categories (*Emerging*, *Attained*, and *Surpassed*) the classification accuracy ranged from 78% to 85% and the classification consistency ranged from 70% to 79%. Across all grades for all three content areas of Supported Independence, the two-category classification accuracy ranged from 89% to 95% and the two-

category classification consistency ranged from 84% to 92%. The three-category classification accuracy ranged from 76% to 83% and the three-category classification consistency ranged 67% to 76%. Across all 34 assessments, there was at least 88% classification accuracy and at least 83% classification consistency when students were classified as proficient or not proficient. 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 later 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 appropriate for students with disabilities such as those that take the MI–Access Participation and Supported Independence assessments.

6. Interrelations Among Strands

One important source of reliability evidence is the consistency of the relations of test subcomponents – interrelations among strands within the test. These interrelations provide a picture of the internal structure of a test, indicating the extent to which item types and items within subsections of the content area "hang together." The Pearson product-moment intercorrelations between section scores for ELA Participation and Supported Independence are reported in Tables 6.1 and 6.2. The Pearson product-moment intercorrelations between section scores for Mathematics Participation and Supported Independence at grades 3 to 8 are reported in Tables 6.3 and 6.4, and for grade 11 the intercorrelations are reported in Tables 6.5 and 6.6. The Pearson product-moment intercorrelations between section scores for Science Participation and Supported Independence are reported in Tables 6.7 and 6.8. For all three content areas for both Participation and Supported Independence, the intercorrelations between section scores were generally high indicating a high degree of association between the subcontent areas. The few exceptions to this pattern were when a subsection consisted of a single item. In these cases, some lower correlations were observed. For each content area within each level, the correlational structure was generally consistent across the grades.

The N, mean, standard deviation, and Cronbach's Coefficient Alpha along with the minimum and maximum score of the section scores are given in Tables 6.9 and 6.10 for ELA Participation and Supported Independence, in Tables 6.11 and 6.12 for Mathematics Participation and Supported Independence at grades 3 to 8, in Tables 6.13 and 6.14 for grade 11 Mathematics, and in Tables 6.15 and 6.16 for Science Participation and Supported Independence. For the most part the results were what one would expect. The Cronbach's Coefficient Alpha estimates were moderately high for the subsections, except for the sections that consisted of a single item in which a Cronbach's Coefficient Alpha estimate did not exist.

Table 6.1

2011–2012 Operational Form Intercorrelations Between Section Scores by Grade – ELA Participation

Grade		Expressing Ideas	Text Comprehension
3	Text Comprehension	0.73	
	Word Recognition	0.62	0.64
4	Text Comprehension	0.66	
	Word Recognition	0.57	0.66
5	Text Comprehension	0.69	
	Word Recognition	0.63	0.59
6	Text Comprehension	0.66	
	Word Recognition	0.70	0.65
7	Text Comprehension	0.66	
	Word Recognition	0.66	0.69
8	Text Comprehension	0.77	
	Word Recognition	0.76	0.77
11	Text Comprehension	0.71	
	Word Recognition	0.67	0.68

Table 6.2

2011–2012 Operational Form Intercorrelations Between Section Scores by Grade – ELA Supported Independence

Grade		Expressing Ideas	Text Comprehension
3	Text Comprehension	0.64	
	Word Recognition	0.47	0.56
4	Text Comprehension	0.76	
	Word Recognition	0.54	0.57
5	Text Comprehension	0.70	
	Word Recognition	0.52	0.56
6	Text Comprehension	0.73	
	Word Recognition	0.54	0.55
7	Text Comprehension	0.74	
	Word Recognition	0.52	0.57
8	Text Comprehension	0.72	
	Word Recognition	0.57	0.57
11	Text Comprehension	0.68	
	Word Recognition	0.68	0.62

Table 6.3

Fall 2011 Operational Form Intercorrelations Between Section Scores by Grade – Mathematics Participation

Grade		Data Probability	Geometry	Measurement
3	Geometry	0.64		
	Measurement	0.56	0.70	
	Numbers and Operations	0.55	0.70	0.68
4	Geometry	0.57		
	Measurement	0.52	0.62	
	Numbers and Operations	0.48	0.63	0.64
5	Geometry	0.57		
	Measurement	0.50	0.64	
	Numbers and Operations	0.54	0.66	0.65
6	Geometry	0.62		
	Measurement	0.50	0.53	
	Numbers and Operations	0.66	0.75	0.60
7	Geometry	0.51		
	Measurement	0.48	0.54	
	Numbers and Operations	0.57	0.70	0.59
8	Geometry	0.65		
	Measurement	0.60	0.63	
	Numbers and Operations	0.68	0.72	0.65

Table 6.4
Fall 2011 Operational Form Intercorrelations Between Section Scores by Grade – Mathematics Supported Independence

Grade		Data Probability	Geometry	Measurement	Algebra
3	Geometry	0.47			
	Measurement	0.57	0.44		
	Numbers and Operations	0.63	0.57	0.59	
4	Geometry	0.51			
	Measurement	0.61	0.45		
	Numbers and Operations	0.69	0.57	0.63	
5	Geometry	0.46			
	Measurement	0.57	0.44		
	Numbers and Operations	0.66	0.56	0.64	
6	Geometry	0.33			
	Measurement	0.45	0.32		
	Numbers and Operations	0.48	0.39	0.51	
	Algebra	0.31	0.40	0.28	0.45
7	Geometry	0.40			
	Measurement	0.46	0.31		
	Numbers and Operations	0.55	0.49	0.53	
	Algebra	0.37	0.41	0.33	0.55
8	Geometry	0.34			
	Measurement	0.50	0.31		
	Numbers and Operations	0.48	0.48	0.51	
	Algebra	0.31	0.45	0.33	0.51

Table 6.5
Spring 2012 Operational Form Intercorrelations Between Section Scores – Mathematics Participation

Grade		Geometry	Logic
11	Logic	0.71	
	Algebra	0.57	0.67

Table 6.6
Spring 2012 Operational Form Intercorrelations Between Section Scores – Mathematics Supported Independence

Grade		Geometry	Logic
11	Logic	0.50	
	Algebra	0.34	0.56

Table 6.7
2011–2012 Operational Form Intercorrelations Between Section Scores by Grade – Science Participation

Grade		Constructing	Reflecting	Life Science	Physical Science
5	Reflecting	0.47			
	Life Science	0.63	0.62		
	Physical Science	0.59	0.64	0.76	
	Earth Science	0.59	0.59	0.74	0.75
8	Reflecting	0.45			
	Life Science	0.56	0.63		
	Physical Science	0.62	0.67	0.78	
	Earth Science	0.53	0.72	0.71	0.79
11	Reflecting	0.46			
	Life Science	0.63	0.69		
	Physical Science	0.61	0.67	0.82	
	Earth Science	0.62	0.62	0.79	0.79

Table 6.8

2011–2012 Operational Form Intercorrelations Between Section Scores by Grade – Science Supported Independence

Grade		Constructing	Reflecting	Life Science	Physical Science
5	Reflecting	0.36			
	Life Science	0.34	0.46		
	Physical Science	0.40	0.36	0.44	
	Earth Science	0.36	0.45	0.60	0.53
8	Reflecting	0.22			
	Life Science	0.33	0.34		
	Physical Science	0.25	0.32	0.56	
	Earth Science	0.31	0.33	0.59	0.59
11	Reflecting	0.13			
	Life Science	0.29	0.22		
	Physical Science	0.24	0.19	0.60	
	Earth Science	0.35	0.26	0.71	0.53

Table 6.9

2011–2012 Operational Form Summary Statistics for Section Scores by Grade – ELA Participation

Grade		N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
3	Expressing Ideas	408	0	24	11.12	7.19	0.70
	Text Comprehension	408	0	18	7.32	6.04	0.69
	Word Recognition	408	0	18	7.72	6.16	0.66
4	Expressing Ideas	397	0	24	10.73	7.12	0.69
	Text Comprehension	397	0	18	7.24	5.97	0.66
	Word Recognition	397	0	18	7.58	6.27	0.66
5	Expressing Ideas	350	0	24	11.44	7.44	0.71
	Text Comprehension	350	0	18	7.82	5.82	0.61
	Word Recognition	350	0	18	8.52	6.28	0.67
6	Expressing Ideas	318	0	24	10.42	7.91	0.71
	Text Comprehension	318	0	18	7.00	6.21	0.69
	Word Recognition	318	0	18	8.23	6.39	0.73
7	Expressing Ideas	320	0	24	11.45	7.59	0.70
	Text Comprehension	320	0	18	8.63	6.27	0.68
	Word Recognition	320	0	18	8.28	6.27	0.70
8	Expressing Ideas	321	0	24	11.69	8.31	0.78
	Text Comprehension	321	0	18	8.41	6.38	0.69
	Word Recognition	321	0	18	8.14	6.33	0.70
11	Expressing Ideas	302	0	24	12.95	7.53	0.75
	Text Comprehension	302	0	18	8.31	6.18	0.73
	Word Recognition	302	0	18	9.46	6.06	0.66

Table 6.10

2011–2012 Operational Form Summary Statistics for Section Scores by Grade – ELA Supported Independence

Grade		N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
3	Expressing Ideas	471	0	24	16.10	5.56	0.64
	Text Comprehension	471	0	20	12.86	5.25	0.62
	Word Recognition	471	0	16	9.95	4.29	0.53
4	Expressing Ideas	484	0	24	16.83	5.87	0.71
	Text Comprehension	484	0	20	12.85	5.59	0.71
	Word Recognition	484	0	16	10.27	4.33	0.53
5	Expressing Ideas	508	0	24	17.46	5.66	0.71
	Text Comprehension	508	0	20	14.04	5.34	0.70
	Word Recognition	508	0	16	11.09	3.93	0.46
6	Expressing Ideas	489	0	24	15.19	6.43	0.75
	Text Comprehension	489	0	20	13.03	5.45	0.72
	Word Recognition	489	0	16	11.99	4.22	0.61
7	Expressing Ideas	486	0	24	15.50	6.31	0.74
	Text Comprehension	486	0	20	12.62	5.50	0.70
	Word Recognition	486	0	16	12.29	4.16	0.61
8	Expressing Ideas	505	0	24	16.33	6.24	0.76
	Text Comprehension	505	0	20	13.20	5.28	0.68
	Word Recognition	505	0	16	12.68	4.04	0.62
11	Expressing Ideas	578	0	24	15.38	6.75	0.76
	Text Comprehension	578	0	20	14.62	5.40	0.69
	Word Recognition	578	0	16	11.99	4.37	0.68

Table 6.11

Fall 2011 Operational Form Summary Statistics for Section Scores by Grade – Mathematics Participation

Grade		N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
3	Data Probability	404	0	6	3.42	2.76	
	Geometry	404	0	24	12.83	8.57	0.82
	Measurement	404	0	12	5.22	4.24	0.52
	Numbers and Recognition	404	0	18	5.90	5.98	0.79
4	Data Probability	396	0	6	3.66	2.67	
	Geometry	396	0	24	12.88	8.37	0.80
	Measurement	396	0	12	5.30	4.29	0.54
	Numbers and Recognition	396	0	18	5.71	5.57	0.70
5	Data Probability	346	0	6	3.72	2.67	
	Geometry	346	0	24	13.58	8.47	0.81
	Measurement	346	0	12	5.72	4.36	0.56
	Numbers and Recognition	346	0	18	6.12	5.51	0.66
6	Data Probability	316	0	6	3.03	2.68	
	Geometry	316	0	24	9.28	7.71	0.73
	Measurement	316	0	6	2.03	2.48	
	Numbers and Recognition	316	0	24	10.42	8.07	0.75
7	Data Probability	320	0	6	3.21	2.70	
	Geometry	320	0	24	10.93	7.73	0.71
	Measurement	320	0	6	2.16	2.47	
	Numbers and Recognition	320	0	24	11.56	8.03	0.75
8	Data Probability	321	0	6	3.32	2.71	
	Geometry	321	0	24	10.71	8.32	0.79
	Measurement	321	0	6	2.29	2.54	
	Numbers and Recognition	321	0	24	11.20	7.98	0.73

Table 6.12
Fall 2011 Operational Form Summary Statistics for Section Scores by Grade – Mathematics Supported Independence

Grade		N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
3	Data Probability	471	0	8	4.37	2.78	0.65
	Geometry	471	0	16	13.28	3.51	0.57
	Measurement	471	0	8	4.16	2.85	0.56
	Numbers and Recognition	471	0	28	15.75	7.47	0.75
4	Data Probability	481	0	8	4.55	2.86	0.71
	Geometry	481	0	16	13.36	3.61	0.62
	Measurement	481	0	8	4.36	2.89	0.60
	Numbers and Recognition	481	0	28	16.29	7.50	0.76
5	Data Probability	506	0	8	4.74	2.76	0.67
	Geometry	506	0	16	13.69	3.27	0.59
	Measurement	506	0	8	4.76	2.77	0.58
	Numbers and Recognition	506	0	28	17.67	7.23	0.75
6	Data Probability	487	0	8	6.22	2.27	0.38
	Geometry	487	0	12	7.85	3.15	0.37
	Measurement	487	0	8	4.76	2.95	0.57
	Numbers and Recognition	487	0	24	12.61	6.55	0.68
	Algebra	487	0	8	4.03	2.65	0.37
7	Data Probability	484	0	8	6.29	2.32	0.46
	Geometry	484	0	12	8.49	3.12	0.40
	Measurement	484	0	8	4.66	2.95	0.54
	Numbers and Recognition	484	0	24	12.6	6.86	0.72
	Algebra	484	0	8	4.09	2.69	0.38
8	Data Probability	504	0	8	6.57	2.14	0.43
	Geometry	504	0	12	8.66	3.13	0.39
	Measurement	504	0	8	4.90	2.90	0.52
	Numbers and Recognition	504	0	24	13.78	6.68	0.71
	Algebra	504	0	8	4.38	2.66	0.29

Table 6.13
Spring 2012 Operational Form Summary Statistics for Section Scores – Mathematics Participation

Grade		N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
11	Algebra	302	0	12	5.60	4.37	0.54
	Geometry	302	0	12	7.09	4.38	0.59
	Logic	302	0	36	18.98	11.44	0.82

Table 6.14
Spring 2012 Operational Form Summary Statistics for Section Scores – Mathematics Supported Independence

Grade		N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
11	Algebra	574	0	8	4.10	2.73	0.32
	Geometry	574	0	12	9.89	2.42	0.28
	Logic	574	0	40	25.64	10.33	0.80

Table 6.15

2011–2012 Operational Form Summary Statistics for Section Scores by Grade – Science Participation

Grade		N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
5	Constructing	341	0	6	3.26	2.72	
	Reflecting	341	0	6	2.29	2.56	
	Life Science	341	0	30	14.85	9.91	0.79
	Physical Science	341	0	30	13.89	9.43	0.75
	Earth Science	341	0	18	8.24	6.03	0.65
8	Constructing	310	0	6	2.70	2.69	
	Reflecting	310	0	6	2.40	2.59	
	Life Science	310	0	30	16.07	10.57	0.84
	Physical Science	310	0	30	14.30	10.13	0.82
	Earth Science	310	0	18	7.96	6.05	0.68
11	Constructing	300	0	6	3.13	2.63	
	Reflecting	300	0	6	3.43	2.68	
	Life Science	300	0	30	15.19	10.10	0.81
	Physical Science	300	0	30	14.37	9.55	0.77
	Earth Science	300	0	18	8.85	6.17	0.67

Table 6.16

2011–2012 Operational Form Summary Statistics for Section Scores by Grade – Science Supported Independence

Grade		N	Minimum Score	Maximum Score	Mean	Standard Deviation	Cronbach's Alpha
5	Constructing	495	0	4	2.88	1.63	
	Reflecting	495	0	4	3.18	1.45	
	Life Science	495	0	28	21.63	5.90	0.65
	Physical Science	495	0	12	7.79	3.54	0.47
	Earth Science	495	0	20	14.10	5.00	0.56
8	Constructing	495	0	4	3.10	1.54	
	Reflecting	495	0	4	3.22	1.42	
	Life Science	495	0	28	18.25	6.38	0.62
	Physical Science	495	0	12	7.73	3.41	0.40
	Earth Science	495	0	20	14.24	4.96	0.58
11	Constructing	572	0	4	2.15	1.82	
	Reflecting	572	0	4	2.53	1.80	
	Life Science	572	0	28	21.94	7.20	0.80
	Physical Science	572	0	12	8.63	3.22	0.41
	Earth Science	572	0	20	12.94	5.42	0.64

7. Summary of Items Flagged for Difficulty and Discrimination

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

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.

The adjusted p-value is equal to the item mean divided by the maximum item score. This PL flag criterion was modified to have an adjusted p-value of less than 0.33 for 2011-2012 for Participation and Supported Independence assessments to be consistent with the Functional Independence assessments. Previously, the criterion was set at having an adjusted p-value less than 0.10.

Table 7.1 summarizes the number of flagged items by level, content area, and grade. The table also shows the total number of items for each assessment. For Participation ELA, from one to four items were flagged PL across grades. No items were flagged for PH or CL. For Supported Independence ELA, no items were flagged PL, two items were flagged PH in grades 6 and 7 and three items were flagged as PH in grade 8. There was one item flagged CL in grades 3 and 4. For Participation Mathematics, one to three items were flagged PL in grades 3 through 8 depending on the grade. No items were flagged PL at grade 11. No items were flagged PH or CL in any grade for Participation Mathematics. For Supported Independence, two items were flagged PL in grade 8 and three items were flagged PL in grades 6 and 7. In each grade, one or two items were flagged PH. There was one item flagged CL in grade 3 and 7, two items in grade 8, and three items flagged in grade 6. For Participation Science, three items were flagged PL in grade 5 and five items were flagged in grade 8. No items were flagged PH or CL. For Supported Independence Science, no items were flagged based on any of the three criteria.

Table 7.1

Number of Items Flagged for Difficulty and Discrimination

Grade	Participation				Supported Independence			
	Total Number of Items	Difficulty Flag		Item-Total Correlation Flag[3]	Total Number of Items	Difficulty Flag		Item-Total Correlation Flag[3]
		PL[1]	PH[2]			PL[1]	PH[2]	
ELA								
3	25	2	0	0	30	0	0	1
4	25	4	0	0	30	0	0	1
5	25	2	0	0	30	0	0	0
6	25	4	0	0	30	0	2	0
7	25	1	0	0	30	0	2	0
8	25	1	0	0	30	0	3	0
11	25	1	0	0	30	0	0	0
Mathematics								
3	25	2	0	0	30	0	2	1
4	25	3	0	0	30	0	1	0
5	25	3	0	0	30	0	2	0
6	25	3	0	0	30	3	1	3
7	25	2	0	0	30	3	1	1
8	25	1	0	0	30	2	1	2
11	25	0	0	0	30	0	2	0
Science								
5	30	3	0	0	32	0	0	0
8	30	5	0	0	32	0	0	0
11	30	0	0	0	32	0	0	0

[1]PL = p-value < 0.33

[2]PH = p-value > 0.90

[3]Item-total correlation < 0.25

8. Spring Online Pilot Analyses

The Michigan Department of Education Bureau of Assessment and Accountability explored the use of an online answer document in the spring 2012 for grade 11. The online answer document was similar in design and layout to the traditional paper answer document that has been used previously for MI-Access. For Participation and Supported Independence, the answer document was completed by the primary and shadow raters who observed the assessment. The procedure for filling out the online document was slightly different than has been used on the paper answer document in the past. For the online answer document, both the primary and shadow ratings were input into the computer by the primary rater and the shadow rater was asked to review and certify that the input of their ratings by the primary were accurate. For the paper answer document, both the primary and shadow rater fill in their ratings on the answer document. The shadow rater was not asked to formally certify that their ratings were accurate. Scores for the online answer document were used operationally to assign scores to students since students did not interact with the answer documents in either the paper or online versions.

Tables 8.1 and 8.2 shows the demographic breakdowns of all students that took each assessment, the students that had their scores input through the online system, and the students whose scores were recorded on the paper answer document for Participation and Supported Independence. The tables indicate that the demographic characteristics were notably different from each other as it pertained to the ethnicities of the students. For Participation, more students that used the online answer document were white, whereas for Supported Independence more students that used the online answer document were black. Most of the other demographic characteristics for the two populations were similar. The tables also show that more students used the paper answer document compared to the online answer document.

Table 8.1
2011–2012 Percent of Students by Subgroup and Content Area for Online and Paper Forms for Participation

	N	Female	Male	Black	Hispanic	White	Economic Disadv
ELA							
All Forms	302	38.7	61.3	29.8	6.6	59.6	57.9
Online	97	42.3	57.7	18.6	8.2	68.0	55.7
Paper	205	37.1	62.9	35.1	5.9	55.6	59.0
Mathematics							
All Forms	302	38.7	61.3	29.8	6.6	59.6	57.9
Online	97	42.3	57.7	18.6	8.2	68.0	55.7
Paper	205	37.1	62.9	35.1	5.9	55.6	59.0
Science							
All Forms	300	38.7	61.3	29.7	6.7	59.7	57.7
Online	95	42.1	57.9	18.9	8.4	67.4	54.7
Paper	205	37.1	62.9	35.1	5.9	55.6	59.0

Table 8.2
2011–2012 Percent of Students by Subgroup and Content Area for Online and Paper Forms for Supported Independence

	N	Female	Male	Black	Hispanic	White	Economic Disadv
ELA							
All Forms	578	35.8	64.2	25.8	5.7	65.6	55.9
Online	118	39.8	60.2	35.6	4.2	55.9	55.9
Paper	460	34.8	65.2	23.3	6.1	68.0	55.9
Mathematics							
All Forms	574	36.2	63.8	25.6	5.7	65.7	55.6
Online	118	39.8	60.2	35.6	4.2	55.9	55.9
Paper	456	35.3	64.7	23.0	6.1	68.2	55.5
Science							
All Forms	572	36.0	64.0	26.2	5.4	65.4	56.1
Online	119	40.3	59.7	36.1	4.2	56.3	58.0
Paper	453	34.9	65.1	23.6	5.7	67.8	55.6

Tables 8.3 and 8.4 provide score summaries for the students that had their scores input on online and on paper. For Participation ELA, online students had higher average scores. For Participation Mathematics and Science, students who took paper had higher average scores. The standard deviations were less for online than for paper. For Supported Independence the score differences between online and paper were less than for Participation. In each case, the difference in average scores was less than one point. Online scores again were less variable than the paper scores. The similarity in scores from the two methods of recording scores is desirable. However, it is important to note that having one's scores recorded on paper or online was not random. Differences or similarities in scores for the two populations in the tables could be a function of some of these similarities and differences between the populations or the way that the student scores were recorded.

Table 8.3
2011–2012 Score Summaries by Online and Paper – Participation

	Online			Paper		
Content Area	Mean	SD	N	Mean	SD	N
ELA	32.06	16.38	97	30.08	18.13	205
Mathematics	30.89	16.33	97	32.04	18.96	205
Science	43.32	25.82	95	45.73	28.84	205

Table 8.4
2011–2012 Score Summaries by Online and Paper –
Supported Independence

	Online			Paper		
Content Area	Mean	SD	N	Mean	SD	N
ELA	41.91	13.62	118	41.56	14.84	460
Mathematics	39.56	12.56	118	39.66	13.69	456
Science	47.51	13.90	119	48.38	13.95	453

Tables 8.5 and 8.6 provide some comparisons of the percent exact agreement rates for the primary and shadow raters for the online and paper answer documents. The pattern in the two tables is similar; the ratings submitted online have higher average percent exact agreement than the ratings that were submitted on paper. The differences were larger for Participation compared to Supported Independence. Although the differences appeared to be small, they were larger than one might expect given that the percent exact agreement rates tend be very high and have had a restricted range historically for both Participation and Supported Independence. The exact reason for the differences in the percent exact agreement rates between the two mediums was hard to identify and would be a good area for future investigation. It is possible that the ratings were being input slightly differently online and via paper such that the small changes affected the extent to which the raters agree with each other. It is also possible that the certification and review process worked extremely well and afforded the shadow raters with an opportunity to review their ratings that they would not otherwise had. This review process may have removed some discrepancies that were not likewise removed with the paper answer document. Other explanations for the differences are also possible.

Table 8.5
2011–2012 Percent Exact Rater Agreement for Primary and Shadow Raters for Online and Paper –
Participation Operational Item Ratings

	Online			Paper		
Content Area	Mean	Min	Max	Mean	Min	Max
ELA	98.45	95.87	100.00	95.71	94.15	96.59
Mathematics	98.35	96.91	100.00	96.00	94.63	97.07
Science	99.03	97.89	100.00	95.58	93.17	97.56

Table 8.6

2011–2012 Average Percent Exact Rater Agreement for Primary and Secondary Raters for Online and Paper–Supported Independence Operational Item Ratings

	Online			Paper		
Content Area	Mean	Min	Max	Mean	Min	Max
ELA	97.12	94.07	100.00	95.72	94.13	97.17
Mathematics	97.34	94.07	100.00	96.61	95.18	98.86
Science	96.98	93.28	100.00	96.39	94.04	97.57

9. Verification of Psychometric Procedures

Assessment and Evaluation Services (AES) served as the subcontractor for the Independent Psychometric Quality Assurance Review. AES reviewed and replicated all psychometric procedures connected to the item analysis, scaling and equating of the assessments. AES has experience in performing quality control services in testing programs in Ohio, New Jersey, New York, Virginia, and Washington. In those states, AES verifies similar psychometric analyses as described in the Michigan program. AES has verified analyses involving Item Response Theory (IRT) equating, scaling, and item analysis. AES staff has expertise in IRT models including, but not limited to, the Rasch model, Partial Credit Model, and three-parameter IRT model. AES also has experience in coordinating this work with various contractors.

In the past, AES has found that working with the psychometric staff of another company and/or state department requires extensive planning and coordinated scheduling. It is essential that the quality control work be extensive and accurate, but it is equally important that it be completed in a timely fashion so that overall project schedules can be met. This requires that both AES and the psychometric staff of the prime contractor and/or state department work closely in planning for the transfer of data and analysis results to AES. Likewise, it is important for AES to complete the checking and transfer our results to the state office for verification.

AES met with MI-Access psychometric staff and Questar staff to discuss the plans and schedules regarding the implementation of this contract. The intent was to coordinate the activities between the contractor, state, and AES to ensure that the verification procedures were implemented in a smooth and accurate manner.

AES provided the verification of item analysis activities for the MI-Access 2011-2012 Participation and Supported Independence assessments. Verifications were done for census form item analysis and field test item analysis.

This section provides a description of the steps AES undertook to provide replication of the MI-Access 2011-2012 analyses for the Participation and Supported Independence assessments. The workflow was organized so that Michigan psychometric staff and AES staff worked independently on each step. Once major portions of the analysis were completed, AES compiled the two sets of results into a comparison spreadsheet. These spreadsheets were then examined by Michigan psychometric staff and AES staff to determine if the replication was successful. When discrepancies between department and AES results occurred during the steps they were often resolved before the comparison spreadsheets were completed.

The MI-Access grades 3-8 and grade 11 Participation and Supported Independence assessments are given in three subject areas: ELA, Mathematics, and Science. ELA and Mathematics were assessed in fall 2011 at grades 3 through 8. Science was assessed at grades 5 and 8. ELA, Mathematics, and Science were assessed in spring 2012 at Grade 11. In most cases, the test forms were structured so that forms contained the same census test items and different field tests

The major analyses for the project are detailed below in steps from the AES perspective. The project has been partitioned into two analysis sets for description; a post assessment file and a field test file. The analyses occurred sequentially. Grades 3-8 analyses were done in the fall and winter of 2011-2012, while the grade 11 analyses were done in the spring and summer of 2012.

Analysis Set 1- Classical Item Analyses for English Language Arts, Mathematics, and Science

The classical item analyses took place after student work had been scored and was based on a data set which included almost all students.

Step 1-Check Data File for Unreasonable Values

In large data files, there are often implausible item response values that are found. This is particularly true in scored data files. AES examined the item score fields for values which were not plausible given the form designation and the item key. When values were found, AES notified Michigan psychometric staff so these instances could be investigated.

Step 2- Analysis of Operational Items

The data was analyzed and classical item statistics were generated. This included analyses of both the primary and shadow scores as well as combined scores for both raters. Adjusted p-values, n-counts, point-biserial correlations, and frequencies and percentages of various ratings given by both raters were computed.

AES Verification Files for Analysis Set 1

The steps detailed above yield one AES verification file. A verification file consists of three spreadsheets and is used to evaluate the verification analyses.

Primary Spreadsheet - This is provided by the primary technical analysis. In the case of MI-Access, the Michigan psychometric staff performs the primary technical analyses. This spreadsheet provides values for variables for each item or each score point depending on the comparison.

AES Spreadsheet-This is provided by AES. It takes the identical form to the Primary Spreadsheet and contains the AES values for each variable or each score point depending on the comparison.

Verification Spreadsheet-This spreadsheet is simply a comparison of the Primary and AES spreadsheets to see if any differences exist. With numerical variables it is usually a simple subtraction of the Primary value minus the AES value. When alpha numeric codes or flag values are compared differences are noted by the display of the values.

The AES verification file for the item analyses of MI-Access focused on checking for data equivalency and verification of item statistics.

Verification File One- Total Classical Statistic Verification
This file is based on the initial item analysis and provides a row in the spreadsheets for each item. The key variables compared were n-counts, adjusted p-values, percent receiving each score point and condition code, ratings for primary and shadow raters, and the point-biserial correlations.

Analysis Set 2-Field Test Item Analysis

Field test items were analyzed to provide item data for committee review and values for the development of future test forms. All subject forms contained field test items. The analyses consisted of information about the performance of the item for the total population and item performance for race/ethnicitry, gender, and economically disadvantaged groups.

Step 1-Check Data File for Unreasonable Values

In large data files, there are often implausible item response values that are found. This is particularly true in scored data files. AES examined the item score fields for values which were not plausible given the form designation and the item key. When values were found, AES notified Michigan psychometric staff so these instances could be investigated.

Step 2-Item Statistics for the Total Group

Analyses were run by grade/subject on all forms to develop adjusted p-values, n-counts, and point-biserial correlations for total group of students that took each assessment. The ratings for the primary and shadow raters were also analyzed for the total group of students.

Step 3-Item Statistics for Race/Ethnicity, Gender, and Economically Disadvantaged Groups

The data was analyzed by race/ethnicity, gender, accommodated, and economically disadvantaged groupings. The specific groups investigated were Black/White, male/female, and economically disadvantaged/non-economically disadvantaged. N-counts, p-values, and point-biserials correlations were calculated for each group.

AES Verification Files for Analysis Set 2

The steps detailed above yield two AES verification files. A verification file consists of three spreadsheets and is used to evaluate the verification analyses.

Primary Spreadsheet - This is provided by the primary technical analysis. In the case of MI-Access, the Michigan psychometric staff performs the primary technical analyses. This spreadsheet provides values for variables for each item or each score point depending on the comparison.

AES Spreadsheet-This is provided by AES. It takes the identical form to the Primary Spreadsheet and contains the AES values for each variable or each score point depending on the comparison.

Verification Spreadsheet-This spreadsheet is simply a comparison of the Primary and AES spreadsheets to see if any differences exist. With numerical variables it is usually a simple subtraction of the Primary value minus the AES value. When alpha numeric codes or flag values are compared differences are noted by the display of the values.

The two AES verification files for the field test Analyses of MI-Access are focused on checking for data equivalency for the total group of students and the various subgroups for MI-Access Participation and

Supported Independence.

Verification File One- Total Classical Statistic Verification
This file is based on the item analysis for total group of students and provides a row in the spreadsheets for each item. The key variables compared were n-counts, adjusted p-values, point-biserial correlations, and percentage of scores at score point and condition code. Analyses of the primary and shadow ratings were also checked.

Verification File Two- Sub-group Classical Statistic Verification
This file is based on the item analysis of all items and provides a row in the spreadsheets for each item. The key variables compared were n-counts, adjusted p-values, point-biserial correlations, and percentage of scores at score point and condition code. Analyses of the primary and shadow ratings were also checked. Rather than one total verification file for each test, a verification file was produced for each sub-group under investigation. Subgroup variables include race/ethnicity, gender, or economically disadvantaged grouping classifications.

10. Validity Argument

An important part of ensuring the technical quality of assessments is examining the validity of the score interpretations and uses of the assessments. The approach taken in this chapter regarding validity is similar to the approach presented in Kane (2006). Kane's (2006) view of validity is to view validity as an argument where positive and negative evidence is collected and an argument is made based on the evidence for or against the intended uses and interpretations. For the MI-Access Supported Independence and Participation assessments, the assessments are used to make decisions about whether or not students' possess adequate mastery of extended grade level content standards by comparing test performance on the assessments to previously established cut-scores. Students that obtain scores that exceeded the cut-scores for the *Attained* performance level are considered to be proficient for that specific grade and assessment. These proficiency data are then used in combination with data on other Michigan assessments to make adequate yearly progress determinations for schools. Score results are also presented to districts, schools, teachers, and parents to provide information about how the student performed on the assessments. Data presented in the other chapters of this technical report and from previous years and other special studies for the MI-Access Supported Independence and Participations assessments constitute the currently compiled validity evidence for or against the intended uses and score interpretations. Ideally, the goal is that the positive evidence would outweigh the negative evidence and that the intended uses and interpretations of the scores are more or less supported by the evidence that has been collected.

One important piece of validity evidence relates to the content of the assessment and ensuring that the tests that are given measure an appropriate sampling of the content. In Chapter 1 of this technical report addendum, the test blueprints for the MI-Access Supported Independence and Participation were provided and outlined. These test blueprints were very similar to previous test administrations and covered a range of content standards represented in the extended grade level content standards underlying the alternate assessments. In the past, more detailed alignment studies of the content standards for each assessment have been conducted using the Webb alignment procedure (Webb, 1997; 2007) and the Linking for Academic for Learning procedure (Flowers, Wakeman, & Browder, 2009). The results from these studies by and large provided support to the fact that assessments had appropriate alignment, although a few areas for improvement were suggested. These included looking at the specific content of some of the items and ensuring that the items are better written to specific extended grade level content expectations. Additional training has been included in item writer training and item review to address these concerns. This has included additional information on Webb's depth of knowledge and more focused reviews of items, the content standards associated with them, and greater attention being placed on looking at the depth of knowledge of the items. These data provide some positive evidence in support of the uses of the assessments.

Another important piece of validity evidence is related to ensuring that the different forms of the assessment are taken by appropriate samples of students and that the scores from the assessments are comparable for different subgroups of the population. To this end, the Michigan Department of Education has developed sampling plans to make sure that the forms for MI-Access Supported Independence and Participation are distributed in an appropriate fashion to schools and that the samples of students that take in each form have similar characteristics. The results in Chapter 2 provide detailed subgroup analyses and breakdowns of the students that took each form. These results showed that for the most part the groups of students that took each form had fairly similar characteristics. This is what would be expected if the sampling plan was working and being implemented effectively.

Analyses of subgroup performance did show some small differences in average scores for several assessments, test levels, and subgroup comparisons. The differences in average scores differed depending on the assessment, test level, and grade in question. For the most part, the differences in average scores were less than 3 raw score points and many of the differences were less than one raw score point. These minor

differences in subgroup performance are somewhat expected if the different groups of students have similar opportunities to learn the content and demonstrate their knowledge on the assessments. This provides some positive evidence in support of the uses of the assessments.

Other analyses of subgroup performance for the spring online pilot from 2012 were presented in Chapter 8. In these cases, the investigations focused on whether not students that had their scores recorded on paper answer documents had similar scores as students that had their scores recorded via online answer documents. The spring online pilot did not involve the students directly interacting with the online system or taking test items via computer. For Participation, the scores differences were less than two points for the paper and online answer document entry and for Supported Independence the scores differences were less than one point. The group with higher scores differed depending on the assessment and test level. In some cases, students that had scores submitted on a paper answer document had higher scores and other cases online students had higher scores. These results are about what one would expect in terms of differences in scores between the groups and tend to provide some support for using the scores interchangeably.

Another common way that states and large-scale assessment programs ensure that assessments are comparable is through the use of equating methods. The Michigan Department of Education has explored the use of equating methods for the MI-Access Participation and Supported Independence assessments in the past (Wyse, Dean, Viger, & Vansickle, in press). These investigations provided evidence to suggest that in many cases employing equating methods to adjust scores on the MI-Access assessments would introduce more error into the scores and create several additional complications than not using formal equating methods. These investigations also demonstrated that the use of equating methods may create situations in which the equating transformations would not be invariant across various subgroups of the population. Hence, the MI-Access assessments do not use formal equating methods to maintain score comparability. Instead, score comparability is maintained through a social moderation process and by building tests to the same test blueprints. This process includes providing statistical targets in terms of the difficulty range for the assessment and the item total test correlations to content specialists, careful review of the items by content specialists and educators, and the removal of items that do not fall within acceptable ranges. Similar to previous years, this process was used again this year and the assessments generated were close to the desired targets.

Chapter 7 provides some analyses of the item total test correlations and average adjusted p-values for the each of the assessments. The results indicated that there were few items that fell outside the item total test correlation or average adjusted p-value ranges that the state had established as being acceptable for these assessments. This provides some positive evidence that most of the items selected had item statistics that would be viewed as being acceptable based on these criteria.

Another piece of validity evidence is related to data on reliability and on the classification accuracy and consistency of the assessments. These data are found in Chapters 3, 4, 5, and part of chapter 8. The results from these chapters in most cases lend support to uses of the MI-Access Participation and Supported Independence assessments. The tests have estimated internal consistency reliability estimates in the 80s and 90s. These internal consistency estimates exceeded the commonly suggest 0.80 level. The standard errors of measurement for the assessment were slightly larger than one would expect given the higher internal consistency estimates, but this can be explained by the odd score distributions and the fact that several of the students received somewhat extreme scores on the assessments.

In Chapter 4, data on the rater consistency of the primary and shadow raters that scored each of the examinees responses to each of the test questions was presented. The levels of rater consistency were well into the mid to high 90s on most of the test items indicating that a high level of agreement existed in viewing and scoring the responses for the students. This again provides positive evidence for scores from the assessments. Part of chapter 8 also looked at rater consistency for the online and paper answer document versions to see if there were any differences between the two ways of recording answers. Results did indicate

that there were some small, but noticeable differences depending on the way that the answers were recorded. In particular, the online answer documents often had higher levels of rater agreement than the paper answer document versions. This is an area that requires some additional investigation. While having higher agreement would probably be viewed as desirable, it is important to ensure that the scores from two mediums have similar levels of consistency. The online version did require an additional affirmation step that was not required for the paper answer documents. There are several other possible explanations for the differences that were observed for the paper and online answer document submissions. Additional explorations of the online and paper answer documents are planned for 2012-2013.

Some additional analyses of the rater agreement data that were not reported in the technical report did show that there were some odd interactions that were possible between the rater agreement rates and Cronbach's Alpha. In particular, some additional analyses indicated that it was possible to have a higher Cronbach's Alpha for raters with lower percent agreement rates when looking at some of the data from the spring 2012 pilot. This is intuitively not what would one expect. One would expect that the estimate of reliability would always be higher when the agreement rate is increased. The reason for these odd results appears to be a function of the fact that Cronbach's Alpha does not capture the impact that raters have as a source of potential error in scoring. Some additional analyses related to this phenomenon are planned for the fall 2012 administrations.

Chapter 5 on classification accuracy and consistency provides some evidence about the reliability of the classification decisions that were made on these assessments. The chapter shows that there were different performance level classifications based on the content area and grade level under consideration and that more students were generally classified as being proficient than not. These performance levels were established several years ago though the application of the Bookmark standard setting procedure using representative stakeholders from across Michigan (Lewis, Mitzel, & Green, 1996; Mitzel, Lewis, Patz, & Green, 2001). The results from the standard setting meeting were reviewed and approved by the State Board of Education and were used to establish the cut-scores for classifying students into different performance categories. The classification decisions had high classification accuracy and consistency, often in the 80s and 90s, when using the scores to make classification decisions used in adequate yearly progress designations. The classification accuracy and consistency was less with three performance levels as compared to using only two performance levels. This suggests that classification decisions were more reliable when making the decisions for NCLB adequate yearly progress than for other more finely grained decisions. Again, these data provide some positive evidence in support of the cut scores and consistency with which decisions can be made based on these cut-scores.

One last piece of evidence is related to the interrelationships between strands and the reliability of these content strand subscores that was presented in Chapter 6. This chapter suggested that for the most part there were somewhat similar patterns between the content strands across grade levels when the content strands were repeated across grades. The correlations in the Chapter 9 also suggest that the correlations between strands were usually fairly moderate. This suggests that there were differences in performance across strands. This is to be expected if the content strands are measuring somewhat unique aspects of achievement associated with the content area and are not perfectly reliable indicators. There is no evidence to suggest that the assessments are functioning in an undesirable way based on the interrelationships between strands.

Chapter 6 also displays the results on the averages for the strand scores and their reliability. The reliabilities of the strand scores were notably lower than reliabilities for the assessments as whole. This suggests that some caution is needed when interpreting the total scores for the strands in various score reports. These scores are not as reliable as the scores for the whole test and they contain less precision in comparison to the total test score. This means that the strand scores may be more variable than the total scores. The lowest strand score reliabilities tended to be found for the strands with the fewest items. However, it is important to point out that these findings are not that different from what is observed in many other large-scale assessments. In many

cases, the strand scores are not as reliable as the overall score for the whole assessment. Again, there is no evidence to suggest that the assessments are functioning in an inappropriate manner.

In viewing all of the evidence as a whole, one can see that by and large the positive evidence for the assessments appears to outweigh the negative evidence for the assessments. This suggests that based on the evidence collected that the intended uses and interpretations of the scores seems to be more supported than rejected.

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing.* Washington, DC: American Research Association.

Brennan, R. L. (2004). BB–CLASS: Beta–Binomial Classification Consistency and Accuracy, version 1.1. CASMA, University of Iowa, Iowa City, IA: Author.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.

Flowers, C., Wakeman, S., & Browder, D. (2009). Links for academic learning (LAL): A conceptual model for investigating alignment of alternate assessments based on alternate achievement standards. *Educational Measurement: Issues and Practice, 28*, 25-37.

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement (4th edition*, pp. 17-64). Washington, DC: American Council on Education.

Lewis, D. M., Mitzel, H. C., & Green, D. R. (1996, June). Standard setting: A bookmark approach. In D. R. Green (Chair), *IRT-based standard setting procedures utilizing behavioral anchoring.* Paper presented at the symposium of the Council of Chief State School Offices National Council of Large-Scale Assessment. Phoenix, AZ.

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

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 249-281). Mahwah, NJ: Lawrence Erlbaum.

Webb, N. L. (1997). *Criteria for the alignment of expectations and assessments in mathematics and science* (Council of Chief State School Officers and National Institute of Science Education Research Monograph No. 6). Madison: University of Wisconsin-Madison, Wisconsin Center for Educational Research.

Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education*, *20*, 7-25.

Wyse, A. E., Dean, V. J., Viger, S. G., & Vansickle, T. R. (in press). Considerations for equating alternate assessments: Two case studies of alternate assessments based on alternate achievement standards. *Applied Measurement in Education*.

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