

Pennsylvania Grade 3 Assessment

Mathematics and Reading

Technical Report

Spring 2005 Operational Test

Submitted to the Pennsylvania Department of Education

October 2005



Table of Contents

Overview	5
Test Design and Sample.....	6
Test Structure	6
Test Sample.....	7
Descriptive Statistics and Item Analysis	8
Item-Level Descriptive Statistics	8
Speededness.....	9
Rater Agreement	9
Differential Item Functioning (DIF)	11
Item Fit Assessment.....	12
IRT Calibration and Equating	14
Establishment of the Pennsylvania Grade 3 Score Scale	16
Test Form Statistics.....	16
Summary statistics for Content Standards.....	18
Appendix	73

Tables

Table 1 2005 Test Design	20
Table 2 Number of Items and Score Points by Item Type	20
Table 3 2005 Pennsylvania Grade 3 Sample Characteristics by Ethnicity	21
Table 4 2005 Pennsylvania Grade 3 Sample Characteristics by Gender	22
Table 5 Item Statistics for Reading Form A (N=25,743)*	23
Table 6 Item Statistics for Reading Form B (N=25,525)*	24
Table 7 Item Statistics for Reading Form C (N=25,226)*	25
Table 8 Item Statistics for Reading Form D (N=24,978)*	26
Table 9 Item Statistics for Reading Form E (N=24,689)*	27
Table 10 Item Statistics for Mathematics Common items (N=110,657)	28
Table 11 Item Statistics for Mathematics Matrix items	29
Table 12 Rater Agreement for Reading Constructed-Response Items	31
Table 13 Rater Agreement for Mathematics Constructed-Response Items	32
Table 14 Summary for Differential Item Functioning based on Criteria $\pm C$	34
Table 15 Scoring Table for Reading Form A	35
Table 16 Scoring Table for Reading Form B	36
Table 17 Scoring Table for Reading Form C	37
Table 18 Scoring Table for Reading Form D	38
Table 19 Scoring Table for Reading Form E	39
Table 20 Scoring Table for Mathematics	40
Table 21 Raw score Descriptive Statistics Based on All Samples	41
Table 22 Raw Score Descriptive Statistics by Ethnicity	42
Table 23 Raw Score Descriptive Statistics by Gender	43
Table 24 Descriptive Statistics for Reported Scale Scores Based on All Samples	44
Table 25 Descriptive Statistics for Reported Scale Scores by Ethnicity	45
Table 26 Descriptive Statistics for Reported Scale Scores by Gender	46
Table 27 Percentiles of Scale Score Ranges	47
Table 28 Percentiles of Reading Performance Cut Scores	50
Table 29 Percentiles of Mathematics Performance Cut Scores	50
Table 30 Number of Items per Each Reading Assessment Anchor *	51
Table 31 Number of Items per each Mathematics Assessment Anchor*	52
Table 32 Summary Statistics for Reading Reporting Categories/Assessment Anchors	53
Table 33 Summary Statistics for Mathematics Reporting Categories/Assessment Anchors (N = 126,577)*	54
Table 34 Factor Analysis Results for Reading Reporting Categories	55
Table 35 Factor Analysis Results for Reading Assessment Anchors	56
Table 36 Factor Analysis Results for Mathematics Reporting Categories	57
Table 37 Factor Analysis Results for Mathematics Assessment Anchors	58

Figures

Figure 1 Reading anchor items' B parameter correlations	60
Figure 2 Mathematics anchor items' B parameter correlations	60
Figure 3 Reading scale scores and SEMs across five forms.....	61
Figure 4 Mathematics scale score and SEM.....	62
Figure 5 Reading Form 1 Raw Score and Scale Score Distribution.....	63
Figure 6 Reading Form 2 Raw Score and Scale Score Distribution.....	64
Figure 7 Reading Form 3 Raw Score and Scale Score Distribution.....	65
Figure 8 Reading Form 4 Raw Score and Scale Score Distribution.....	66
Figure 9 Reading Form 5 Raw Score and Scale Score Distribution.....	67
Figure 10 Mathematics Raw Score and Scale Score Distribution	68
Figure 11 <i>Scree Plots for Reading Reporting Categories and Assessment Anchors</i>	69
Figure 12 <i>Scree Plots for Mathematics Reporting Categories and Assessment Anchors</i>	70
Figure 13 <i>State Mean Scale Scores</i>	71

Overview

The purpose of this report is to provide technical information about the 2005 spring operational test administration of the Pennsylvania Grade 3 Reading and Mathematics Assessments. Testing in Reading began in 2002, and testing in Mathematics was included in 2003. The score scales for Reading and Mathematics were set up during the 2003 operational testing. The 2004 score scales were transformed to the 2003 scale, and the 2005 score scales were transformed to the 2004 scale. New performance cuts were also established in 2005.

This report includes an overview of the operational test design, summary of the operational test items, and test form analyses. The report also provides a summary of raw score descriptive statistics at the item and test form levels and a discussion of the procedures used for form calibration and equating. Also included is a summary of the reported test scores, as individual test scores were reported based on the scale scores.

Test Design and Sample

Test Structure

The Pennsylvania Grade 3 Reading and Mathematics Assessments are part of the Pennsylvania School of System Assessment (PSSA) currently being administered in Pennsylvania. This year, in an effort to improve the way its standard based assessment fosters increases in student learning, Pennsylvania implemented the use of Assessment Anchors to replace the previously used content standards. Assessment Anchors are subcategories of broader Reporting Categories, and are designed to focus the PSSA test development on specific and meaningful knowledge and skills. The criterion-referenced assessment is intended to measure the two reporting categories for Reading, which include five assessment anchors, and the five reporting categories for Mathematics, which include eleven assessment anchors (see Appendix and Tables 30 and 31). Reading and Mathematics were administered together in one test book. Ten books were spiraled within classrooms. Reading has five unique forms repeated two times, whereas each of the 10 Mathematics forms is unique (See Table 1 for the test design).

Both Mathematics and Reading consist of common items, which were taken by every student and were common across all forms of the test, as well as matrix and embedded field-tested (FT) items. For Mathematics, the matrix and FT items are unique to each form of the test (with the exception of matrix open ended items, each of which are found in two forms). Mathematics matrix items do not contribute to students' reported scores, but do contribute to aggregated scores used for curriculum analysis. For Reading, matrix and FT items each appear twice, as there are five unique forms rather than ten. Common and matrix items both contribute to students' scores for Reading, because the

Reading assessment is shorter. Table 2 shows the number of items and score points by item type, multiple choice (MC) items and open ended (OE) items. Note that in Mathematics, OE items are scored using a four score point rubric, and for Reading, OE items are scored using a three score point rubric. For the reported individual scores, the total number-correct score was 61 for Mathematics (score of common items only) and 45 or 46 for Reading (sum of the scores of common and matrix items).

Test Sample

In April 2005, approximately 130,000 Pennsylvania Grade 3 students took one of the ten test books that contain both Reading and Mathematics. Table 3 shows the ethnic characteristics for the total number of students by form. The percentages are rounded to whole numbers. As the table indicates, most of the students were White (75%), approximately 15% of the students were African American, and approximately 6% of the students were Hispanic. As expected, these ratios were similar across all test forms because the test forms were spiraled within the classrooms. As shown in Table 4, slightly more male students (around 52%) than female students (48%) were in the tested population.

Descriptive Statistics and Item Analysis

Item-Level Descriptive Statistics

Tables 5–11 present item-level descriptive statistics for each of the operational test forms in test book order. Tables 5-9 contain Reading items for each of the five Reading forms, where matrix items as well as items that were common across forms were treated as unique items within each form. Therefore, there are five different statistics for each Reading common item. For Mathematics, however, only one set of statistics was provided for each common item (Table 10). Therefore, the statistics for Mathematics common items were computed based on the entire tested population. This was done because test forms were spiraled within classrooms, and statistics for common items were similar across alternative forms. The statistics for the unique matrix items in the ten Mathematics forms can be found in Table 11.

Tables 5-11 contain the following information: item type, item p -value, item correlation with the total test score (R-ITT), the percent of examinees that omitted an item, and fit information. The p -value for an MC item represents the proportion of students who answered the item correctly. The p -value for an OE item represents the proportion of the obtained mean raw score for the item to the number of points possible for the item. A point-biserial correlation between the item score and the total score on the test was also computed for the MC items. For the OE items, a Pearson correlation between the item score and the total score on the test was computed. For the item analysis, the studied item was excluded from the computation of the total score so as to not artificially inflate the correlation statistic. This effect would be most noticeable for OE items worth several points.

It should be noted that in 2005, Reading and Mathematics items were evaluated using the following three criteria: a p -value below 0.30 for MC items, a point-biserial below 0.15, and an omit-rate above 5%. Items falling within these criteria were flagged for further examination. Also note that item-level descriptive statistics were not given for suppressed items (Mathematics common item #21, Reading matrix item #30 from forms B and E).

Speededness

The degree to which a test is speeded can be evaluated by examining the percentage of students who fail to respond to the last items on the test. The omit rates shown in Tables 5–11 demonstrate that no forms are speeded. There were no differences between omit rates for items at the beginning of the test forms and items at the end of the test forms.

Rater Agreement

In order to monitor the reliability of the scoring of the OE items, approximately 5% of the common items and Reading matrix items and 10% of the Mathematics matrix items from student test booklets were submitted to a second rater for scoring. All other responses were read by a single rater. Indices of rater agreement and consistency were obtained using those students who had their OE items read by two raters. Tables 12–13 present the rater agreement statistics for the Reading and Mathematics OE items. These tables provide the percentages of pairs of raters' scores that did not differ (i.e., perfect agreement) and the percentages of pairs of raters' scores that differed by one point (i.e., adjacent agreement) for all OE items over all test forms. When rater agreement was

defined as two rater scores that differed by no more than one point (i.e., agreement), there was high rater agreement in terms of the percentage of agreement, which ranged from 98.3% to 99.9% for Reading and from 96.5% to 100% for Mathematics. Again, note that the maximum possible score points for the OE items was 3 or 4. In addition to the percentage of agreement, the tables present the mean item score and item standard deviation of the item scores assigned by each rater group. Examination of Tables 12-13 shows that the mean score points awarded by the two rater groups are very close. To further study rater agreement, intraclass correlations (Cronbach's alpha) and Kappa (Fleiss, et. al., 1969) coefficients were calculated and are reported as measures of rater agreement for each OE item. Ordinal rating scales (e.g., 0, 1, 2) used in scoring OE items contain a certain level of chance agreement that is expected. Although the intraclass correlation is reported in this report, it does not take into account chance agreement between the two raters, but Kappa does. Therefore, in general, Kappa will have values equal to or smaller than the intraclass correlation. If agreement is perfect, then Kappa is +1. If agreement is at chance levels, Kappa is 0. Landis and Koch (1977) suggest that values of Kappa greater than .75 indicate "excellent agreement", values between .40 and .74 represent "good agreement" beyond chance, and values below .40 denote "poor agreement." As Tables 12–13 show, Kappa coefficients ranged from 0.70 to 0.86 for Reading and from 0.82 to 0.97 for Mathematics. Note that the intraclass correlation and Kappa for Mathematics were higher than those for Reading. This trend has been often found in other large scale assessment programs. The values of Kappa for Reading and Mathematics meet the criteria of "good agreement" for Reading and the criteria of "excellent agreement" for Mathematics according to Landis and Koch.

Differential Item Functioning (DIF)

An item flagged for DIF is more difficult for a particular group of students than would be expected based on their total test scores, compared to the performance of the other group. The groups compared in the analysis were female and male students, and African–American, Hispanic, and White students. Other ethnic groups were not included in these analyses because their sample sizes were too small.

The statistical procedures used by CTB to identify items thought to exhibit substantial DIF are the same procedures used by ETS and NAEP. For multiple-choice items, the Mantel-Haenszel (χ^2_{MH}) statistic was used to evaluate potential DIF items. In this procedure, the “C”-level DIF items are flagged, where a “C” item indicates a large amount of DIF and has an absolute value of the Mantel-Haenszel (Δ_{MH}) significantly greater than zero (at the .05 level), and $|\Delta_{MH}|$ exceeds 1.5 (Zwick, Donoghue, and Grima, 1993).

For the constructed-response items, both the Mantel χ^2 and the standardized mean difference (SMD) statistics were used to evaluate DIF. Using these procedures, items can be flagged where the Mantel statistic is greater than zero with probability greater than .05, and the absolute value of the SMD is greater than .25. A detailed description of these procedures can be found in Zwick, et al., (1993).

Table 14 presents summary for Differential Item Functioning based on Criteria $\pm C$. Because the DIF statistics were computed based on test form, there were multiple statistics for common items. When a common item was flagged on only a few forms, this item was not flagged. Note that all items flagged based on DIF statistics were reviewed also by content editors to consider content perspective of those items.

Item Fit Assessment

A statistical procedure was used to identify items that did not fit the IRT model. Item model fit information was obtained for each item using a Z-statistic. The Z-statistic is a transformation of the chi-square (Q_I) statistic that takes into account differing numbers of score levels as well as sample size:

$$Z_j = \frac{(Q_{Ij} - DF_j)}{\sqrt{2DF_j}}$$

where Q_{Ij} is the item chi-square statistic,

j is an item, and

DF is the degrees of freedom for a given item j .

The Z-statistic is an index of the degree to which obtained proportions of students with each item score are close to the proportions that would be predicted by the estimated student ability and item parameters. These values, along with the associated chi-squares (Q_I), are computed for ten intervals corresponding to deciles of the ability distribution (Yen, 1984). Because the value of Z increases as the sample size increases, with other things being equal, the critical values for Z were established using the following equation (Yen, 1991a):

$$Z_{crit,j} = \frac{4N_j}{1500}$$

where $Z_{crit,j}$ is critical value of Z for item j , and

N_j is the number of students who responded to item j .

Tables 5–11 present items that were flagged statistically for poor fit for each test form. Many items displayed poor fit because the one-parameter (1PL)/one-parameter partial credit (1PPC) approach (See below IRT calibration and equating section) was used

to produce Z statistics. In the tables, the number “3” represents poor fit. Poor fit can easily happen for many items because the 1PL model does not consider the guessing factor.

IRT Calibration and Equating

Student item responses were calibrated using the combination of two IRT models. The 1PL was used to scale the SR items, and the 1PPC model was employed to scale the OE items. The 1PL defines an SR item in terms of the item difficulty (b_i). That is, the item discrimination (a_i) does not vary over items. In this model, the probability that a student with scale score θ responds correctly to item i is

$$P_i(\theta) = \frac{1}{1 + \exp[-1.7a_i(\theta - b)]}$$

The 2PPC model defines an OE item in terms of an item discrimination and a location parameter for each score point (Muraki, 1990, 1992):

$$P_{jk}(\theta) = P(x_j = k - 1 | \theta) = \frac{\exp Z_{jk}}{\sum_{i=1}^{m_j} \exp Z_{ji}}, k = 1, \dots, m_j,$$

where m_j is the number of score levels,

$$Z_{jk} = A_{jk} \theta + C_{jk},$$

$$C_{jk} = -\sum_{i=0}^{k-1} \gamma_{ji}, \text{ where } \gamma_{j0} = 0,$$

where γ_{ji} is a parameter freely estimated from the data.

The 1PPC model for the OE items can be considered a special case of the 2PPC model. In the 1PPC model, the discrimination does not vary over items. This is the same discrimination parameter that is applied to all test items. In the above equation for the 2PPC model, the following equation replaces A_{jk} with

$$A_k = \alpha (k - 1), \quad k = 1, 2, \dots, m_j,$$

where α represents a common discrimination parameter for all items.

The IRT calibrations were implemented using CTB's PARDUX software (Burket, 1991). PARDUX simultaneously estimates parameters for MC and OE items using marginal maximum likelihood procedures implemented via the EM algorithm (Bock and Aitkin, 1981; Thissen, 1982). Because the ten test forms were spiraled within classrooms, the groups of students who took the different forms can be considered randomly equivalent. Using the anchor items (i.e., items common to all forms), student item response data from alternate test forms were calibrated together. All items across all test forms converged during item calibration.

After 2005 items were calibrated using PARDUX, the scale of 2005 items are transformed to the 2004 scale using anchor items based on the Stocking and Lord (1983) procedure for multiple choice items, and the Stocking and Lord extension for open ended items. To check the stability of anchor items, item b parameters of 2004 anchor items and estimated item b parameters of 2005 anchor items were plotted in Figure 1. The horizontal axis represents 2004 item b parameters and the vertical axis represents 2005 item b parameters. As can be seen in the figure, most 2004 item parameters were aligned well with 2005 item parameters.

Establishment of the Pennsylvania Grade 3 Score Scale

The 2005 Pennsylvania score scale was transformed to the 2004 scale using anchor items by Year-to-Year equating. After transformation to the 2004 score scale, scoring tables for Reading and Mathematics were generated. Tables 15–20 show the scoring tables. Note that raw number-correct scores and scale scores have a one-to-one relationship because the 1PL/1PPC model was used for item calibration and scaling. These scale scores and standard errors of measurement (SEM) on the scoring tables were plotted in Figures 3 and 4. Also, Figures 5-10 show the distributions of raw scores and scale scores for the five Reading forms and for Mathematics. For Reading, all scale scores and SEMs across all five forms appeared to be similar. Because only common items are used for scoring across all alternative forms for Mathematics, only one curve line for each scale score and SEM appears in Figure 4.

Test Form Statistics

Table 21 presents raw score descriptive statistics for each test form. Listed by test form are the number of students, mean raw score, average p -value, standard deviation, minimum score, maximum score, reliability coefficient (Cronbach's alpha), and the SEM. The average p -value was calculated by dividing the mean raw score by the maximum possible raw score points for the test. The maximum possible raw score was 45 or 46 for Reading, depending on the form, 75 for common and matrix Mathematics items, and 61 for common-only Mathematics items. The p -values, which ranged from 0.71 to 0.73 for Reading and 0.77 to 0.79 for Mathematics, showed rather moderate difficulties. Examination of Table 21 reveals that for Reading, the maximum mean raw score form

difference between any two forms was 1.47 for Form B (32.23) and Form C (33.70). All Reading and Mathematics forms had high reliability, with the lowest reliability coefficient being 0.883. The reliability and the SEM for all alternative forms were similar.

Tables 22 and 23 present the mean, standard deviation, minimum and maximum raw score for each form by ethnicity and gender, respectively. Note that for Mathematics, two raw scores were used for computation. One is the raw score for common items only and the other is the raw score for both common items and matrix items. The results show the mean (average) performance of each subgroup on each test form. As the results indicate, for both Reading and Mathematics, White students generally performed better than African American and Hispanic students. Note that there were also substantially more White students than students in other ethnic groups. For Reading, female students generally performed better than male students, and for Mathematics, there was no difference in performance between female and male students.

Table 24 presents descriptive statistics for reported scale scores based on all samples. For 2005, the state mean was 1327 for Reading and 1365 for Mathematics. In 2004, the state mean was 1,296 for Reading and 1,341 for Mathematics. The state means for 2003, 2004 and 2005 are plotted in Figure 13. The mean and standard deviation were similar across alternative forms. Tables 25 and 26 show the scale score means and standard deviations of the ethnic and gender subgroups for each form. The results show how each subgroup performed compared to other students by test form. The pattern found in the raw score statistics was also found in the scale score statistics. That is, for both Reading and Mathematics, White students generally performed better than African

American and Hispanic students. Like the difference in raw score, there was almost one standard deviation (200) difference between the score of White students and the other two ethnic groups. For Reading, female students generally performed better than male students, and for Mathematics, there was no difference in performance between female and male students. The distributions of raw scores and scale scores can be found in Figures 5–10. The upper plot shows the raw score distribution and the lower plot shows the scale score distribution. Because Reading and Mathematics were relatively easy for Pennsylvania students, the distribution of raw score appeared to be positively skewed.

To help with test score interpretation, percentiles of scale scores are provided in Table 27. These percentiles are based on Pennsylvania students who took Grade 3 Reading and Mathematics. As mentioned previously, standard setting was conducted this year in order to establish cut scores for performance levels. Four performance levels were established: below basic, basic, proficient, and advanced. Tables 28 and 29 show the percent of students at each performance level by ethnicity, gender, and all students combined.

Summary statistics for Content Standards

Reading consists of two Reporting Categories (RCs) containing five Assessment Anchors (AAs), and Mathematics consists of five RCs containing 11 AAs. Tables 30 and 31 show the number of items in each AA for Reading and Mathematics. Tables 32 and 33 give raw mean, standard deviation, and mean p-value for each RC and AA. Note that raw scores are reported for both RCs and AAs. Raw scores for Mathematics were computed across all forms because only common items are used for individual student

reports. Note that for both Reading and Mathematics, the mean p-values vary considerably across RCs and AAs.

Factor analysis was done to examine the structure of both the Reporting Categories and Assessment Anchors for 2005 Pennsylvania Grade 3 Reading and Mathematics. Principal axis factor analysis was applied with one factor extraction. See Tables 34 and 35 as well as Figure 11 for Reading factor analysis results, and Tables 36 and 37 as well as Figure 12 for Mathematics factor analysis results. Commonalities, total variance explained, factor matrices with one factor, and scree plots were reported. As can be seen from the scree plots in Figures 11 and 12, a one factor model appeared to be appropriate for Reading and Mathematics.

Table 1
2005 Test Design

Test Book	Reading Form	Mathematics Form
1	A	A
2	B	B
3	C	C
4	D	D
5	E	E
6	A	F
7	B	G
8	C	H
9	D	I
10	E	J

Table 2
Number of Items and Score Points by Item Type

Content		Common Items			Matrix Items			Field-Tested Items		
		Total	MC	OE	Total	MC	OE	Total	MC	OE
Reading	Number of Items	25	24	1	17	16	1	11	10	1
	Score Points	27	24	3	19	16	3	13	10	3
Mathematics	Number of Items	56	54	2	11	10	1	8/9	7/8*	1
	Score Points	62	54	8	14	10	4	11/12	7/8*	4

* Odd-numbered Mathematics Forms have 8 MC field test items; even-numbered Mathematics Forms have 7 MC field test items.

Table 3
2005 Pennsylvania Grade 3 Sample Characteristics by Ethnicity

Content	Form	Number of Students *	Caucasian		African American		Percent Hispanic		Others	
			Count	Percent	Count	Percent	Count	Percent	Count	Percent
Reading	A	25,474	19,013	75	3,943	15	1,624	6	894	4
	B	25,263	18,804	74	3,976	16	1,543	6	940	4
	C	24,962	18,659	75	3,799	15	1,536	6	968	4
	D	24,701	18,484	75	3,777	15	1,523	6	917	4
	E	24,433	18,356	75	3,718	15	1,537	6	822	3
	Total	124,833	93,316	75	19,213	15	7,763	6	4,541	4
Mathematics	A	13,153	9,755	74	2,099	16	817	6	482	4
	B	12,979	9,670	75	2,036	16	800	6	473	4
	C	12,798	9,548	75	1,956	15	792	6	502	4
	D	12,688	9,509	75	1,908	15	798	6	473	4
	E	12,515	9,391	75	1,943	16	776	6	405	3
	F	12,382	9,275	75	1,853	15	833	7	421	3
	G	12,359	9,166	74	1,949	16	763	6	481	4
	H	12,269	9,146	75	1,880	15	773	6	470	4
	I	12,093	9,006	74	1,896	16	741	6	450	4
	J	11,991	8,990	75	1,794	15	786	7	421	4
	Total	125,227	93,456	75	19,314	15	7,879	6	4,578	4

*Students of unspecified ethnicity are not counted.

Table 4
2005 Pennsylvania Grade 3 Sample Characteristics by Gender

Content	Form	Number of Students *	Male		Female	
			Count	Percent	Count	Percent
Reading	A	25,525	13,164	52	12,361	48
	B	25,313	13,083	52	12,230	48
	C	25,018	12,881	51	12,137	49
	D	24,767	12,600	51	12,167	49
	E	24,492	12,717	52	11,775	48
	Total	125,115	64,445	52	60,670	48
Mathematics	A	13,183	6,839	52	6,344	48
	B	13,002	6,767	52	6,235	48
	C	12,825	6,602	51	6,223	49
	D	12,723	6,466	51	6,257	49
	E	12,540	6,508	52	6,032	48
	F	12,404	6,365	51	6,039	49
	G	12,387	6,354	51	6,033	49
	H	12,299	6,344	52	5,955	48
	I	12,124	6,194	51	5,930	49
	J	12,025	6,245	52	5,780	48
	Total	125,512	64,684	52	60,828	48

* Students of unspecified gender are not counted.

Table 5

*Item Statistics for Reading Form A (N=25,743)**

Item	Type	P-Val	R-ITT	Omit	FIT	Item	Type	P-Val	R-ITT	Omit	FIT
1	MC	0.60	0.27	0.00	3	22	MC	0.61	0.35	0.00	3
2	MC	0.89	0.44	0.00	1	23	MC	0.82	0.52	0.00	1
3	MC	0.86	0.37	0.00	1	24	MC	0.80	0.56	0.00	3
4	MC	0.74	0.42	0.01	1	25	MC	0.79	0.45	0.01	1
5	MC	0.53	0.29	0.00	3	26	MC	0.86	0.58	0.00	3
6	MC	0.73	0.52	0.01	1	27	MC	0.76	0.52	0.01	1
7	MC	0.47	0.36	0.01	3	28	MC	0.44	0.39	0.00	3
8	MC	0.76	0.50	0.01	1	29	MC	0.90	0.52	0.00	3
9	MC	0.77	0.58	0.00	3	30	MC	0.73	0.56	0.00	1
10	MC	0.84	0.61	0.00	3	31	MC	0.80	0.64	0.01	3
11	MC	0.86	0.53	0.02	1	32	MC	0.61	0.52	0.00	1
12	MC	0.77	0.50	0.00	1	33	MC	0.83	0.57	0.01	3
13	MC	0.87	0.55	0.00	3	34	CR	0.56	0.56	.	1
14	MC	0.75	0.61	0.00	3	35	MC	0.36	0.30	0.00	3
15	MC	0.29	0.30	0.00	3	36	MC	0.57	0.48	0.00	3
16	MC	0.78	0.48	0.00	1	37	MC	0.76	0.58	0.00	3
17	CR	0.66	0.67	.	1	38	MC	0.85	0.54	0.01	3
18	MC	0.78	0.55	0.00	1	39	MC	0.75	0.61	0.00	3
19	MC	0.81	0.39	0.00	1	40	MC	0.69	0.49	0.01	1
20	MC	0.78	0.54	0.00	1	41	MC	0.63	0.48	0.01	1
21	MC	0.80	0.54	0.00	1	42	MC	0.58	0.45	0.01	3

* OE omit rates are considered 'Blanks'

** FT OE omit rates were not calculated

Table 6
*Item Statistics for Reading Form B (N=25,525)**

Item	Type	P-Val	R-ITT	Omit	FIT	Item	Type	P-Val	R-ITT	Omit	FIT
1	MC	0.61	0.28	0.00	3	22	MC	0.61	0.36	0.00	3
2	MC	0.90	0.44	0.00	1	23	MC	0.83	0.52	0.00	1
3	MC	0.87	0.37	0.00	1	24	MC	0.81	0.54	0.00	3
4	MC	0.74	0.42	0.01	1	25	MC	0.79	0.45	0.01	1
5	MC	0.53	0.28	0.00	3	26	MC	0.78	0.53	0.00	1
6	MC	0.74	0.52	0.01	1	27	MC	0.79	0.55	0.00	1
7	MC	0.48	0.36	0.01	3	28	MC	0.59	0.38	0.00	3
8	MC	0.77	0.49	0.01	1	29	MC	0.71	0.48	0.00	1
9	MC	0.77	0.58	0.00	3	30					
10	MC	0.85	0.59	0.01	3	31	MC	0.80	0.56	0.00	3
11	MC	0.87	0.52	0.03	1	32	MC	0.61	0.50	0.00	3
12	MC	0.78	0.49	0.00	1	33	MC	0.75	0.57	0.00	3
13	MC	0.88	0.54	0.00	3	34	CR	0.61	0.56	.	3
14	MC	0.76	0.61	0.00	3	35	MC	0.75	0.50	0.00	1
15	MC	0.30	0.31	0.00	3	36	MC	0.85	0.46	0.00	1
16	MC	0.78	0.47	0.00	1	37	MC	0.84	0.49	0.00	1
17	CR	0.67	0.66	.	1	38	MC	0.76	0.49	0.01	1
18	MC	0.78	0.55	0.00	1	39	MC	0.39	0.29	0.00	3
19	MC	0.81	0.39	0.00	1	40	MC	0.62	0.46	0.01	1
20	MC	0.78	0.54	0.00	1	41	MC	0.79	0.41	0.00	1
21	MC	0.80	0.54	0.00	1	42	MC	0.52	0.37	0.02	3

* OE omit rates are considered 'Blanks'

** FT OE omit rates were not calculated

Table 7

*Item Statistics for Reading Form C (N=25,226)**

Item	Type	P-Val	R-ITT	Omit	FIT	Item	Type	P-Val	R-ITT	Omit	FIT
1	MC	0.60	0.28	0.00	3	22	MC	0.61	0.35	0.00	3
2	MC	0.90	0.42	0.00	1	23	MC	0.83	0.51	0.00	1
3	MC	0.87	0.37	0.00	1	24	MC	0.81	0.55	0.00	3
4	MC	0.74	0.43	0.01	1	25	MC	0.80	0.44	0.01	1
5	MC	0.53	0.28	0.00	3	26	MC	0.67	0.57	0.00	3
6	MC	0.74	0.51	0.01	1	27	MC	0.72	0.57	0.00	3
7	MC	0.48	0.36	0.01	3	28	MC	0.77	0.58	0.00	3
8	MC	0.77	0.49	0.01	1	29	MC	0.70	0.59	0.00	3
9	MC	0.77	0.59	0.00	3	30	MC	0.81	0.59	0.00	3
10	MC	0.85	0.60	0.01	3	31	MC	0.71	0.52	0.00	1
11	MC	0.87	0.52	0.03	1	32	MC	0.53	0.42	0.01	3
12	MC	0.78	0.48	0.00	1	33	MC	0.86	0.64	0.01	3
13	MC	0.88	0.55	0.00	3	34	CR	0.68	0.66	.	1
14	MC	0.76	0.61	0.00	3	35	MC	0.84	0.52	0.00	1
15	MC	0.30	0.30	0.00	3	36	MC	0.68	0.52	0.00	1
16	MC	0.79	0.49	0.00	1	37	MC	0.78	0.51	0.03	1
17	CR	0.67	0.65	.	1	38	MC	0.83	0.56	0.04	1
18	MC	0.78	0.55	0.00	1	39	MC	0.79	0.59	0.01	3
19	MC	0.81	0.37	0.00	1	40	MC	0.71	0.53	0.00	1
20	MC	0.78	0.53	0.00	1	41	MC	0.77	0.53	0.00	1
21	MC	0.80	0.54	0.00	1	42	MC	0.62	0.47	0.00	1

* OE omit rates are considered 'Blanks'

** FT OE omit rates were not calculated

Table 8

*Item Statistics for Reading Form D (N=24,978)**

Item	Type	P-Val	R-ITT	Omit	FIT	Item	Type	P-Val	R-ITT	Omit	FIT
1	MC	0.60	0.28	0.00	3	22	MC	0.61	0.34	0.00	3
2	MC	0.90	0.44	0.00	1	23	MC	0.83	0.51	0.00	1
3	MC	0.87	0.37	0.00	1	24	MC	0.81	0.55	0.00	3
4	MC	0.74	0.41	0.01	1	25	MC	0.79	0.45	0.01	1
5	MC	0.55	0.29	0.00	3	26	MC	0.86	0.57	0.00	3
6	MC	0.75	0.52	0.01	1	27	MC	0.77	0.51	0.01	1
7	MC	0.48	0.35	0.01	3	28	MC	0.44	0.39	0.00	3
8	MC	0.77	0.49	0.01	1	29	MC	0.90	0.53	0.00	3
9	MC	0.78	0.58	0.00	3	30	MC	0.73	0.56	0.00	1
10	MC	0.84	0.60	0.01	3	31	MC	0.81	0.63	0.01	3
11	MC	0.87	0.52	0.03	1	32	MC	0.61	0.52	0.00	1
12	MC	0.78	0.49	0.00	1	33	MC	0.84	0.56	0.01	3
13	MC	0.88	0.56	0.01	3	34	CR	0.56	0.55	.	1
14	MC	0.76	0.60	0.00	3	35	MC	0.88	0.42	0.00	1
15	MC	0.30	0.30	0.00	3	36	MC	0.85	0.46	0.00	1
16	MC	0.78	0.48	0.00	1	37	MC	0.29	0.26	0.00	3
17	CR	0.67	0.67	.	1	38	MC	0.86	0.57	0.00	3
18	MC	0.78	0.55	0.00	1	39	MC	0.74	0.54	0.00	1
19	MC	0.81	0.37	0.00	1	40	MC	0.74	0.49	0.00	1
20	MC	0.78	0.53	0.00	1	41	MC	0.70	0.50	0.00	1
21	MC	0.80	0.55	0.00	1	42	MC	0.78	0.48	0.01	1

* OE omit rates are considered 'Blanks'

** FT OE omit rates were not calculated

Table 9

*Item Statistics for Reading Form E (N=24,689)**

Item	Type	P-Val	R-ITT	Omit	FIT	Item	Type	P-Val	R-ITT	Omit	FIT
1	MC	0.60	0.28	0.00	3	22	MC	0.61	0.36	0.00	3
2	MC	0.90	0.42	0.00	1	23	MC	0.83	0.53	0.00	1
3	MC	0.87	0.37	0.00	1	24	MC	0.81	0.55	0.00	3
4	MC	0.75	0.42	0.01	1	25	MC	0.80	0.44	0.01	1
5	MC	0.53	0.28	0.00	3	26	MC	0.79	0.53	0.00	1
6	MC	0.75	0.52	0.00	1	27	MC	0.79	0.55	0.00	1
7	MC	0.48	0.36	0.00	3	28	MC	0.58	0.39	0.00	3
8	MC	0.77	0.50	0.01	1	29	MC	0.71	0.48	0.00	1
9	MC	0.78	0.58	0.00	3	30					
10	MC	0.85	0.60	0.01	3	31	MC	0.80	0.55	0.00	3
11	MC	0.87	0.51	0.02	1	32	MC	0.61	0.52	0.00	3
12	MC	0.78	0.49	0.00	1	33	MC	0.76	0.56	0.00	3
13	MC	0.88	0.55	0.00	3	34	CR	0.61	0.56	.	3
14	MC	0.76	0.61	0.00	3	35	MC	0.76	0.41	0.00	1
15	MC	0.30	0.30	0.00	3	36	MC	0.68	0.44	0.00	1
16	MC	0.79	0.48	0.00	1	37	MC	0.65	0.55	0.00	1
17	CR	0.67	0.67	.	1	38	MC	0.67	0.43	0.01	1
18	MC	0.79	0.54	0.00	1	39	MC	0.64	0.41	0.00	3
19	MC	0.81	0.39	0.00	1	40	MC	0.56	0.46	0.00	1
20	MC	0.79	0.52	0.00	1	41	MC	0.93	0.44	0.00	1
21	MC	0.80	0.53	0.00	1	42	MC	0.89	0.51	0.01	3

* OE omit rates are considered 'Blanks'

** FT OE omit rates were not calculated

Table 10

Item Statistics for Mathematics Common items (N=110,657)

Item	Type	P-Val	R-ITT	Omit	Fit	Item	Type	P-Val	R-ITT	Omit	Fit
1	MC	0.80	0.39	0.00	1	29	MC	0.86	0.48	0.00	1
2	MC	0.94	0.33	0.00	1	30	MC	0.93	0.39	0.01	1
3	MC	0.97	0.29	0.01	1	31	MC	0.84	0.42	0.00	3
4	MC	0.85	0.53	0.00	3	32	MC	0.80	0.53	0.01	1
5	MC	0.79	0.56	0.00	3	33	MC	0.86	0.42	0.01	1
6	MC	0.85	0.38	0.00	1	34	OE	0.74	0.44	0.01	1
7	MC	0.85	0.40	0.00	1	35	MC	0.85	0.45	0.00	1
8	MC	0.97	0.28	0.00	1	36	MC	0.84	0.41	0.00	1
9	MC	0.83	0.38	0.00	1	37	MC	0.76	0.50	0.00	1
10	MC	0.87	0.42	0.00	1	38	MC	0.91	0.40	0.01	1
11	MC	0.79	0.40	0.00	1	39	MC	0.70	0.48	0.00	1
12	MC	0.89	0.44	0.00	1	40	MC	0.89	0.46	0.00	1
13	MC	0.89	0.38	0.00	1	41	MC	0.79	0.37	0.00	3
14	MC	0.82	0.44	0.00	1	42	MC	0.73	0.59	0.00	1
15	MC	0.81	0.47	0.00	1	43	MC	0.98	0.29	0.00	1
16	OE	0.91	0.45	0.00	1	44	MC	0.90	0.46	0.01	1
17	MC	0.79	0.46	0.00	1	45	MC	0.70	0.43	0.01	1
18	MC	0.83	0.34	0.00	3	46	MC	0.63	0.43	0.00	1
19	MC	0.89	0.31	0.00	1	47	MC	0.92	0.41	0.00	1
20	MC	0.69	0.27	0.01	3	48	MC	0.78	0.48	0.00	3
21	MC	0.00	.	1.00	1	49	MC	0.77	0.31	0.00	1
22	MC	0.76	0.54	0.01	3	50	MC	0.62	0.44	0.02	1
23	MC	0.97	0.39	0.00	1	51	MC	0.74	0.52	0.00	3
24	MC	0.67	0.48	0.00	1	52	MC	0.71	0.61	0.01	3
25	MC	0.83	0.46	0.00	3	53	MC	0.81	0.55	0.01	1
26	MC	0.74	0.45	0.00	1	54	MC	0.66	0.46	0.00	1
27	MC	0.84	0.46	0.01	1	55	MC	0.90	0.39	0.00	3
28	MC	0.55	0.61	.	1	56	MC	0.51	0.50	.	3

Table 11

Item Statistics for Mathematics Matrix items

Form	Item	Type	P-Val	R-ITT	Omit	FIT	Form	Item	Type	P-Val	R-ITT	Omit	FIT
1 (N = 13,307)	57	MC	0.94	0.44	0.00	1	4 (N = 12,827)	57	MC	0.85	0.43	0.00	1
	58	MC	0.89	0.52	0.00	3		58	MC	0.95	0.36	0.00	1
	59	MC	0.90	0.51	0.00	3		59	MC	0.89	0.39	0.00	1
	60	MC	0.84	0.54	0.00	3		60	MC	0.77	0.54	0.01	3
	61	MC	0.47	0.32	0.00	3		61	MC	0.92	0.40	0.00	1
	62	MC	0.79	0.27	0.00	3		62	MC	0.89	0.29	0.00	1
	63	MC	0.79	0.48	0.00	1		63	MC	0.91	0.38	0.00	1
	64	MC	0.93	0.37	0.00	1		64	MC	0.91	0.37	0.00	1
	65	MC	0.92	0.46	0.00	1		65	MC	0.77	0.33	0.00	3
	66	MC	0.91	0.43	0.00	1		66	MC	0.88	0.41	0.00	1
2 (N = 13,116)	67	CR	0.54	0.54	.	3	5 (N = 12,633)	67	CR	0.55	0.56	.	3
	57	MC	0.84	0.34	0.00	1		57	MC	0.73	0.49	0.00	1
	58	MC	0.93	0.46	0.00	1		58	MC	0.89	0.45	0.00	1
	59	MC	0.84	0.51	0.00	1		59	MC	0.86	0.42	0.00	1
	60	MC	0.85	0.44	0.00	1		60	MC	0.85	0.33	0.00	1
	61	MC	0.86	0.42	0.00	1		61	MC	0.86	0.61	0.00	3
	62	MC	0.85	0.45	0.00	1		62	MC	0.89	0.43	0.00	1
	63	MC	0.78	0.46	0.00	1		63	MC	0.82	0.37	0.00	1
	64	MC	0.63	0.38	0.00	3		64	MC	0.72	0.43	0.01	1
	65	MC	0.96	0.34	0.00	1		65	MC	0.96	0.35	0.00	1
3 (N = 12,930)	66	MC	0.43	0.36	0.00	3	6 (N = 12,501)	66	MC	0.82	0.45	0.00	1
	67	CR	0.68	0.48	.	3		67	CR	0.64	0.48	.	3
	57	MC	0.88	0.37	0.00	1		57	MC	0.44	0.48	0.00	1
	58	MC	0.81	0.44	0.00	1		58	MC	0.85	0.42	0.00	1
	59	MC	0.91	0.42	0.00	1		59	MC	0.86	0.38	0.00	1
	60	MC	0.68	0.52	0.00	3		60	MC	0.86	0.54	0.00	3
	61	MC	0.68	0.33	0.00	3		61	MC	0.62	0.42	0.00	3
	62	MC	0.83	0.55	0.00	3		62	MC	0.76	0.35	0.00	3
	63	MC	0.94	0.35	0.00	1		63	MC	0.65	0.46	0.00	1
	64	MC	0.87	0.39	0.00	1		64	MC	0.89	0.42	0.00	1
	65	MC	0.91	0.41	0.00	1		65	MC	0.88	0.44	0.00	1
	66	MC	0.95	0.25	0.00	1		66	MC	0.83	0.53	0.00	1
	67	CR	0.80	0.63	.	1		67	CR	0.80	0.63	.	1

Table 11 (Cont.)

Item Statistics for Mathematics Matrix items

Form	Item	Type	P-Val	R-ITT	Omit	FIT	Form	Item	Type	P-Val	R-ITT	Omit	FIT
7 (N = 12,487)	57	MC	0.90	0.51	0.00	3	9 (N = 12,239)	57	MC	0.92	0.45	0.00	1
	58	MC	0.76	0.49	0.00	1		58	MC	0.86	0.44	0.00	1
	59	MC	0.89	0.42	0.00	1		59	MC	0.77	0.52	0.00	1
	60	MC	0.63	0.47	0.00	1		60	MC	0.64	0.51	0.00	1
	61	MC	0.65	0.43	0.00	1		61	MC	0.92	0.44	0.00	1
	62	MC	0.77	0.54	0.00	3		62	MC	0.84	0.30	0.00	3
	63	MC	0.57	0.43	0.00	1		63	MC	0.66	0.39	0.00	1
	64	MC	0.80	0.56	0.00	3		64	MC	0.66	0.38	0.00	3
	65	MC	0.88	0.57	0.00	3		65	MC	0.73	0.31	0.00	3
	66	MC	0.99	0.20	0.00	1		66	MC	0.79	0.48	0.00	1
8 (N = 12,405)	67	CR	0.66	0.45	.	3		67	CR	0.86	0.48	.	3
	57	MC	0.73	0.37	0.00	3	10 (N = 12,132)	57	MC	0.56	0.44	0.00	3
	58	MC	0.86	0.45	0.00	1		58	MC	0.91	0.39	0.00	1
	59	MC	0.88	0.43	0.00	1		59	MC	0.83	0.35	0.00	1
	60	MC	0.65	0.50	0.00	3		60	MC	0.88	0.40	0.02	1
	61	MC	0.64	0.39	0.00	1		61	MC	0.84	0.42	0.00	1
	62	MC	0.73	0.44	0.00	1		62	MC	0.95	0.39	0.00	1
	63	MC	0.95	0.37	0.00	1		63	MC	0.95	0.37	0.00	1
	64	MC	0.78	0.52	0.02	1		64	MC	0.87	0.42	0.00	1
	65	MC	0.98	0.10	0.00	1		65	MC	0.76	0.47	0.00	1
	66	MC	0.84	0.58	0.00	3		66	MC	0.76	0.40	0.00	1
8 (N = 12,405)	67	CR	0.86	0.48	.	3		67	CR	0.66	0.44	.	3

Table 12

Rater Agreement for Reading Constructed-Response Items

Form	Item	Mean of G1	Mean of G2	SD of G1	SD of G2	Percent of Perfect Agreement	Percent of Adjacent Agreement	Percent of Agreement	Intraclass Correlation	Kappa
A	17	2.04	2.05	0.76	0.73	83.61	15.77	99.38	0.92	0.84
A	34	1.70	1.71	0.70	0.71	74.70	24.24	98.94	0.86	0.71
A	53	1.86	1.88	0.63	0.65	78.67	21.27	99.94	0.87	0.74
B	17	2.06	2.04	0.77	0.77	83.20	15.91	99.11	0.92	0.83
B	34	1.86	1.87	0.84	0.83	81.72	17.72	99.44	0.93	0.86
B	53	1.62	1.60	1.01	1.01	74.23	24.59	98.82	0.93	0.86
C	17	2.08	2.09	0.72	0.71	85.98	13.57	99.55	0.92	0.85
C	34	2.13	2.12	0.73	0.73	86.23	13.41	99.64	0.93	0.86
C	53	2.00	2.00	0.71	0.68	74.05	25.60	99.65	0.86	0.72
D	17	2.05	2.06	0.73	0.72	85.58	13.88	99.46	0.92	0.85
D	34	1.71	1.71	0.74	0.73	75.04	23.28	98.32	0.86	0.72
D	53	1.29	1.33	0.74	0.75	77.80	21.85	99.65	0.90	0.79
E	17	2.01	2.01	0.73	0.71	84.37	14.34	98.71	0.91	0.81
E	34	1.81	1.81	0.80	0.80	81.59	18.03	99.62	0.92	0.85
E	53	1.73	1.74	0.71	0.73	71.00	28.30	99.30	0.85	0.70

G1: Rater group 1 G2: Rater group 2

Percent of Agreement is the sum of percents of perfect and adjacent agreements.

Table 13

Rater Agreement for Mathematics Constructed-Response Items

Form	Item	N	Mean of G1	Mean of G2	SD of G1	SD of G2	Percent of Perfect Agreement	Percent of Perfect Agreement	Percent of Agreement	Intraclass Correlation	Kappa
A	28	679	2.21	2.20	0.85	0.84	84.24	15.17	99.41	0.94	0.88
A	56	674	2.04	2.04	0.92	0.92	89.76	9.35	99.11	0.96	0.92
A	67	1197	2.16	2.14	1.15	1.14	81.79	16.21	98.00	0.95	0.91
A	76	1168	2.68	2.66	1.13	1.13	82.96	15.67	98.63	0.95	0.91
B	28	651	2.26	2.27	0.90	0.88	83.26	16.44	99.70	0.94	0.89
B	56	648	2.09	2.09	0.95	0.94	90.43	8.95	99.38	0.97	0.94
B	67	1326	2.70	2.70	1.12	1.13	88.24	8.60	96.84	0.96	0.91
B	75	1213	2.14	2.12	1.22	1.21	92.58	6.84	99.42	0.98	0.97
C	28	572	2.25	2.27	0.86	0.87	83.22	15.03	98.25	0.92	0.85
C	56	572	2.10	2.10	0.92	0.91	87.59	11.71	99.30	0.96	0.91
C	67	1305	3.19	3.19	0.81	0.81	94.79	5.13	99.92	0.98	0.96
C	76	1222	2.55	2.56	0.93	0.93	94.11	5.65	99.76	0.98	0.96
D	28	637	2.23	2.23	0.86	0.84	82.89	16.48	99.37	0.93	0.87
D	56	633	2.04	2.05	0.89	0.88	91.15	8.21	99.36	0.97	0.93
D	67	1348	2.27	2.27	1.13	1.14	78.41	19.81	98.22	0.95	0.89
D	75	1275	2.82	2.82	1.19	1.20	91.14	8.63	99.77	0.98	0.96
E	28	628	2.24	2.26	0.84	0.84	81.85	17.83	99.68	0.93	0.87
E	56	624	2.09	2.08	0.95	0.96	86.06	12.82	98.88	0.95	0.90
E	67	1432	2.52	2.54	1.17	1.17	84.57	11.94	96.51	0.95	0.90
E	76	1225	3.05	3.05	1.10	1.08	91.27	7.43	98.70	0.97	0.95
F	28	654	2.22	2.26	0.87	0.87	84.71	14.83	99.54	0.94	0.89
F	56	650	2.05	2.04	0.91	0.90	89.85	9.08	98.93	0.96	0.91
F	67	1215	3.17	3.17	0.79	0.80	95.39	4.36	99.75	0.98	0.95
F	75	1252	1.89	1.91	1.14	1.13	82.35	16.13	98.48	0.96	0.91
G	28	601	2.26	2.24	0.89	0.91	81.70	17.80	99.50	0.94	0.88
G	56	600	2.02	2.01	0.95	0.93	89.33	9.83	99.16	0.96	0.93
G	67	1275	2.66	2.65	0.77	0.77	91.84	7.84	99.68	0.96	0.92
G	76	1277	2.67	2.67	0.93	0.92	89.74	10.26	100.00	0.97	0.94

G1: Rater group 1

G2: Rater group 2

Percent of Agreement is the sum of percents of perfect and adjacent agreements.

Table 13 (cont.)

Rater Agreement for Mathematics Constructed-Response Items

Form	Item	N	Mean of G1	Mean of G2	SD of G1	SD of G2	Percent of Perfect Agreement	Percent of Perfect Agreement	Percent of Agreement	Intraclass Correlation	Kappa
H	28	555	2.18	2.17	0.86	0.85	86.31	13.33	99.64	0.95	0.90
H	56	553	2.03	2.01	0.92	0.93	90.05	8.86	98.91	0.96	0.92
H	67	1206	3.48	3.46	0.91	0.93	95.85	3.48	99.33	0.98	0.96
H	75	1178	3.08	3.08	1.10	1.10	85.57	12.22	97.79	0.96	0.91
I	28	682	2.25	2.26	0.85	0.84	81.96	17.16	99.12	0.93	0.86
I	56	680	2.06	2.05	0.93	0.95	87.50	11.62	99.12	0.96	0.91
I	67	1207	3.45	3.46	0.91	0.90	95.19	3.56	98.75	0.97	0.93
I	76	1245	1.62	1.64	0.90	0.91	83.53	15.50	99.03	0.94	0.88
J	28	723	2.17	2.17	0.84	0.86	85.06	14.80	99.86	0.95	0.89
J	56	720	1.99	1.99	0.91	0.93	90.14	8.61	98.75	0.96	0.92
J	67	1196	2.61	2.61	0.72	0.73	92.14	7.69	99.83	0.96	0.92
J	75	1208	2.48	2.48	0.88	0.88	77.15	21.69	98.84	0.91	0.82

G1: Rater group 1

G2: Rater group 2

Percent of Agreement is the sum of percents of perfect and adjacent agreements.

Table 14

Summary for Differential Item Functioning based on Criteria $\pm C$

Content	Form	Focal Group*	Item	Type	N Ref Group	N Focal Group	Delta	Criteria
MA	1	F	26	1	6839	6344	-1.852	-C
MA	1	F	31	1	6839	6344	-1.592	-C
MA	2	F	26	1	6767	6235	-1.880	-C
MA	2	F	31	1	6767	6235	-1.767	-C
MA	3	F	26	1	6602	6223	-1.817	-C
MA	4	F	26	1	6466	6257	-2.235	-C
MA	4	F	31	1	6466	6257	-1.656	-C
MA	5	F	26	1	6508	6032	-1.931	-C
MA	6	F	3	1	6365	6039	1.583	+C
MA	6	F	26	1	6365	6039	-1.746	-C
MA	6	F	39	1	6365	6039	-1.501	-C
MA	7	F	26	1	6354	6033	-1.786	-C
MA	7	F	31	1	6354	6033	-1.596	-C
MA	8	F	26	1	6344	5955	-1.809	-C
MA	8	F	31	1	6344	5955	-1.544	-C
MA	9	F	26	1	6194	5930	-1.818	-C
MA	10	F	26	1	6245	5780	-1.663	-C
MA	10	F	31	1	6245	5780	-1.503	-C
MA	3	H	16	1	9548	792	-1.574	-C
MA	6	H	16	1	9275	833	-1.780	-C
MA	7	H	16	1	9166	763	-1.818	-C
MA	7	H	64	1	9166	763	1.618	+C
MA	8	H	16	1	9146	773	-1.575	-C
MA	8	H	63	1	9146	773	1.608	+C
MA	9	H	16	1	9006	741	-1.528	-C
MA	2	AA	16	1	9670	2036	-1.76512	-C
MA	3	AA	16	1	9548	1956	-1.80402	-C
MA	5	AA	16	1	9391	1943	-1.9509	-C
MA	6	AA	16	1	9275	1853	-1.7334	-C
MA	7	AA	16	1	9166	1949	-1.51846	-C
MA	8	AA	16	1	9146	1880	-1.9696	-C
MA	9	AA	16	1	9006	1896	-1.57843	-C

*F = Female, H = Hispanic, AA = African American

Table 15
Scoring Table for Reading Form A

NC	SS	SEM	NC	SS	SEM
0	300	217			
1	377	176	26	1143	57
2	503	127	27	1162	58
3	580	105	28	1182	59
4	636	93	29	1202	60
5	681	84	30	1223	61
6	719	78	31	1244	62
7	753	74	32	1267	63
8	782	70	33	1290	64
9	809	67	34	1315	66
10	834	65	35	1341	68
11	858	63	36	1368	71
12	880	61	37	1398	73
13	901	60	38	1431	77
14	921	59	39	1466	80
15	941	58	40	1506	85
16	960	57	41	1551	92
17	978	57	42	1604	100
18	997	56	43	1669	113
19	1015	56	44	1756	134
20	1033	56	45	1893	182
21	1051	56	46	1999	237
22	1069	56			
23	1087	56			
24	1105	56			
25	1124	57			

Table 16
Scoring Table for Reading Form B

NC	SS	SEM	NC	SS	SEM
0	300	229			
1	395	176	26	1151	57
2	521	126	27	1170	58
3	597	105	28	1190	59
4	652	92	29	1210	60
5	696	84	30	1231	61
6	734	78	31	1253	62
7	766	73	32	1276	64
8	796	69	33	1300	65
9	822	66	34	1325	67
10	847	64	35	1352	70
11	869	62	36	1381	72
12	891	60	37	1412	75
13	912	59	38	1447	79
14	932	58	39	1485	84
15	951	57	40	1529	90
16	970	57	41	1580	99
17	988	56	42	1643	111
18	1006	56	43	1727	132
19	1024	56	44	1861	180
20	1042	56	45	1999	257
21	1060	56			
22	1077	56			
23	1095	56			
24	1114	56			
25	1132	57			

Table 17
Scoring Table for Reading Form C

NC	SS	SEM	NC	SS	SEM
0	300	224			
1	387	176	26	1123	56
2	512	126	27	1141	56
3	587	104	28	1159	57
4	642	91	29	1178	58
5	686	83	30	1198	59
6	722	77	31	1218	60
7	754	72	32	1239	61
8	783	68	33	1261	63
9	808	65	34	1284	64
10	832	63	35	1309	66
11	854	61	36	1335	69
12	875	59	37	1364	71
13	895	58	38	1395	75
14	915	57	39	1428	79
15	933	56	40	1467	84
16	951	55	41	1510	90
17	969	55	42	1561	99
18	986	55	43	1624	111
19	1003	54	44	1708	132
20	1020	54	45	1843	181
21	1037	54	46	1999	270
22	1054	54			
23	1071	54			
24	1088	55			
25	1105	55			

Table 18
Scoring Table for Reading Form D

NC	SS	SEM	NC	SS	SEM
0	300	209			
1	362	176	26	1124	57
2	489	127	27	1143	58
3	565	105	28	1162	59
4	621	93	29	1182	60
5	666	84	30	1203	61
6	704	78	31	1225	62
7	737	73	32	1248	63
8	767	70	33	1271	65
9	793	67	34	1296	67
10	818	64	35	1323	69
11	842	63	36	1351	71
12	863	61	37	1382	74
13	884	60	38	1415	78
14	904	58	39	1452	82
15	924	58	40	1493	87
16	943	57	41	1540	93
17	961	56	42	1595	102
18	979	56	43	1662	114
19	997	56	44	1751	135
20	1015	55	45	1890	183
21	1033	55	46	1999	239
22	1051	56			
23	1068	56			
24	1087	56			
25	1105	57			

Table 19
Scoring Table for Reading Form E

NC	SS	SEM	NC	SS	SEM
0	300	220			
1	381	177	26	1146	57
2	508	127	27	1165	58
3	584	106	28	1185	59
4	641	93	29	1205	59
5	686	84	30	1225	60
6	724	78	31	1247	62
7	757	74	32	1269	63
8	787	70	33	1293	65
9	814	67	34	1318	67
10	839	65	35	1344	69
11	863	63	36	1373	72
12	885	61	37	1404	75
13	906	60	38	1438	79
14	926	59	39	1476	84
15	945	58	40	1519	90
16	964	57	41	1570	98
17	983	56	42	1633	111
18	1001	56	43	1717	132
19	1019	56	44	1851	180
20	1037	56	45	1999	264
21	1055	56			
22	1073	56			
23	1091	56			
24	1109	56			
25	1127	57			

Table 20
Scoring Table for Mathematics

NC	SS	SEM	NC	SS	SEM
0	200	163			
1	200	163	36	1110	51
2	290	134	37	1125	51
3	375	112	38	1140	52
4	440	100	39	1156	52
5	492	91	40	1172	53
6	535	84	41	1188	53
7	574	79	42	1204	54
8	608	75	43	1221	55
9	638	71	44	1239	56
10	666	68	45	1257	57
11	692	66	46	1276	58
12	716	64	47	1297	60
13	739	62	48	1318	61
14	760	60	49	1340	63
15	781	59	50	1364	66
16	800	57	51	1389	68
17	819	56	52	1417	71
18	837	55	53	1448	75
19	854	54	54	1483	80
20	871	54	55	1522	86
21	887	53	56	1568	94
22	903	52	57	1624	104
23	919	52	58	1696	119
24	934	51	59	1792	142
25	949	51	60	1945	191
26	964	51	61	1999	213
27	979	51			
28	993	50			
29	1008	50			
30	1023	50			
31	1037	50			
32	1051	50			
33	1066	50			
34	1080	50			
35	1095	51			

Table 21
Raw score Descriptive Statistics Based on All Samples

Content	Form	N	Mean						
		Count	Mean	P-Value	SD	Min	Max	Alpha	SEM
Reading	A	25,743	32.45	0.71	9.07	3	46	0.927	2.45
	B	25,525	32.23	0.72	8.55	3	45	0.917	2.46
	C	25,226	33.70	0.73	9.24	4	46	0.933	2.39
	D	24,978	33.36	0.73	8.66	3	46	0.924	2.38
	E	24,689	32.53	0.72	8.57	2	45	0.919	2.44
Mathematics (Both Common items and Matrix items)	A	13,307	57.67	0.77	11.76	9	75	0.936	2.97
	B	13,116	58.36	0.78	11.47	12	75	0.933	2.97
	C	12,930	59.46	0.79	11.22	13	75	0.939	2.77
	D	12,827	58.72	0.78	11.32	12	75	0.933	2.93
	E	12,633	58.74	0.78	11.48	10	75	0.935	2.93
	F	12,501	58.47	0.78	11.74	11	75	0.941	2.86
	G	12,487	58.29	0.78	11.39	13	75	0.939	2.81
	H	12,405	59.08	0.79	11.56	8	75	0.939	2.85
	I	12,239	59.03	0.79	11.45	12	75	0.938	2.85
	J	12,132	58.64	0.78	11.18	11	75	0.937	2.81
Mathematics (Common Items only)	Total	126,577	47.67	0.78	9.51	4	61	0.927	2.56

Table 22
Raw Score Descriptive Statistics by Ethnicity

Content	Form	White				African American				Hispanic			
		N Count	Raw Score Mean	Raw Score SD	Mean P-Value	N Count	Raw Score Mean	Raw Score SD	Mean P-Value	N Count	Raw Score Mean	Raw Score SD	Mean P-Value
Reading	A	19,013	34.26	8.12	0.74	3,943	26.52	9.45	0.59	1,624	25.90	9.67	0.56
	B	18,804	33.83	7.65	0.75	3,976	27.32	9.06	0.62	1,543	25.84	9.58	0.60
	C	18,659	35.56	8.15	0.77	3,799	27.52	9.83	0.60	1,536	26.96	10.01	0.59
	D	18,484	35.12	7.60	0.76	3,777	27.47	9.18	0.61	1,523	26.74	9.60	0.59
	E	18,356	34.22	7.55	0.76	3,718	26.80	9.14	0.60	1,537	26.57	9.61	0.59
Mathematics (Both Common items and Matrix items)	A	9,755	60.10	10.13	0.80	2,099	49.35	13.06	0.67	817	49.81	13.23	0.67
	B	9,670	60.74	9.84	0.81	2,036	50.31	12.79	0.68	800	50.54	12.97	0.69
	C	9,548	61.83	9.46	0.82	1,956	50.96	12.74	0.69	792	51.76	12.88	0.71
	D	9,509	61.17	9.46	0.82	1,908	49.95	12.85	0.68	798	50.33	13.36	0.69
	E	9,391	61.11	9.80	0.81	1,943	50.18	12.85	0.68	776	51.03	13.70	0.68
	F	9,275	61.08	9.81	0.81	1,853	49.20	13.19	0.66	833	49.71	13.51	0.67
	G	9,166	60.75	9.57	0.81	1,949	49.92	12.83	0.67	763	49.86	13.30	0.67
	H	9,146	61.54	9.80	0.82	1,880	50.28	13.12	0.69	773	51.77	13.15	0.71
	I	9,006	61.40	9.76	0.82	1,896	50.43	13.17	0.68	741	51.26	12.44	0.69
	J	8,990	61.10	9.23	0.81	1,794	49.59	13.08	0.68	786	51.29	12.74	0.69
Mathematics (Common items only)	Total	93,456	49.69	8.03	0.81	19,314	40.55	10.76	0.66	7,879	41.16	10.88	0.67

Table 23
Raw Score Descriptive Statistics by Gender

Content	Form	Male				Female			
		N Count	Raw Score Mean	Raw Score SD	Mean P-Value	N Count	Raw Score Mean	Raw Score SD	Mean P-Value
Reading	A	13,164	31.67	9.39	0.69	12,361	33.38	8.59	0.73
	B	13,083	31.34	8.83	0.70	12,230	33.24	8.10	0.74
	C	12,881	33.02	9.54	0.72	12,137	34.50	8.82	0.75
	C	12,600	32.63	8.99	0.71	12,167	34.18	8.20	0.74
	D	12,717	31.70	8.87	0.70	11,775	33.49	8.09	0.74
Mathematics (Both Common items and Matrix items)	A	6,839	57.99	11.71	0.77	6,344	57.46	11.73	0.77
	B	6,767	58.60	11.46	0.78	6,235	58.20	11.42	0.78
	C	6,602	59.97	11.32	0.80	6,223	59.02	11.09	0.79
	D	6,466	59.02	11.31	0.79	6,257	58.46	11.33	0.78
	E	6,508	58.86	11.59	0.78	6,032	58.64	11.44	0.78
	F	6,365	58.79	11.70	0.78	6,039	58.24	11.76	0.78
	G	6,354	58.70	11.29	0.78	6,033	57.95	11.47	0.77
	H	6,344	59.49	11.52	0.79	5,955	58.73	11.57	0.78
	I	6,194	59.35	11.45	0.79	5,930	58.79	11.43	0.78
	J	6,245	58.87	11.21	0.78	5,780	58.48	11.08	0.78
Mathematics (Common Items only)	Total	64,684	47.97	9.46	0.79	60,828	47.43	9.50	0.78

Table 24

Descriptive Statistics for Reported Scale Scores Based on All Samples

Content	Form	N Count	Scale Score Mean	Scale Score SD	Skewness	Kurtosis
Reading	A	25,743	1,322	229	-0.21	-0.25
	B	25,525	1,324	218	-0.17	-0.06
	C	25,226	1,334	238	-0.12	-0.20
	D	24,978	1,328	226	-0.25	-0.13
	E	24,689	1,327	221	-0.15	0.00
	Total	126,161	1,327	227	-0.18	-0.12
Mathematics	A	13,307	1,354	222	-0.07	0.00
	B	13,116	1,366	221	-0.05	0.04
	C	12,930	1,367	216	-0.11	0.05
	D	12,827	1,368	221	-0.07	0.05
	E	12,633	1,368	220	-0.08	0.02
	F	12,501	1,366	223	-0.03	0.05
	G	12,487	1,368	217	-0.05	0.12
	H	12,405	1,363	221	-0.09	0.00
	I	12,239	1,368	220	-0.05	0.06
	J	12,132	1,365	220	-0.06	0.10
	Total	126,577	1,365	220	-0.07	0.05

Table 25
Descriptive Statistics for Reported Scale Scores by Ethnicity

Content	Test Form	White				African American				Hispanic			
		Mean	SD	MIN	MAX	Mean	SD	MIN	MAX	Mean	SD	MIN	MAX
Reading	A	1,366	215	580	1,999	1,174	210	636	1,893	1,161	213	681	1,999
	B	1,364	206	597	1,999	1,201	203	597	1,861	1,169	210	652	1,727
	C	1,381	224	642	1,999	1,178	214	686	1,999	1,166	217	686	1,999
	D	1,373	210	565	1,999	1,176	205	666	1,890	1,161	213	666	1,890
	E	1,369	207	641	1,999	1,182	202	508	1,999	1,181	216	641	1,999
	Total	1,371	213	565	1,999	1,182	207	508	1,999	1,168	214	641	1,999
Mathematics	A	1,398	206	608	1,999	1,202	206	638	1,945	1,211	208	666	1,945
	B	1,410	207	574	1,999	1,218	203	608	1,945	1,225	208	666	1,945
	C	1,411	199	692	1,999	1,210	199	638	1,945	1,226	204	692	1,792
	D	1,413	203	608	1,999	1,204	202	440	1,945	1,217	211	716	1,945
	E	1,412	204	666	1,999	1,208	195	574	1,945	1,230	215	692	1,999
	F	1,411	205	535	1,999	1,202	203	638	1,945	1,213	210	666	1,945
	G	1,411	200	692	1,999	1,219	203	608	1,999	1,220	209	638	1,945
	H	1,407	205	638	1,999	1,209	203	574	1,945	1,229	212	535	1,792
	I	1,411	205	666	1,999	1,216	205	638	1,945	1,221	195	716	1,945
	J	1,410	201	608	1,999	1,200	208	638	1,945	1,229	211	692	1,999
	Total	1,409	204	535	1,999	1,209	203	440	1,999	1,222	208	535	1,999

Table 26
Descriptive Statistics for Reported Scale Scores by Gender

Content	Test Form	Male				Female			
		Mean	SD	MIN	MAX	Mean	SD	MIN	MAX
Reading	A	1,345	223	636	1,999	1,302	232	580	1,999
	B	1,352	215	597	1,999	1,300	218	597	1,999
	C	1,355	235	686	1,999	1,316	239	642	1,999
	D	1,350	220	666	1,999	1,309	229	565	1,999
	E	1,352	217	508	1,999	1,305	222	584	1,999
	Total	1,351	222	508	1,999	1,306	228	565	1,999
Mathematics	A	1,349	222	638	1,999	1,360	220	608	1,999
	B	1,362	223	716	1,999	1,372	219	574	1,999
	C	1,359	216	716	1,999	1,377	216	638	1,999
	D	1,362	222	608	1,999	1,375	218	440	1,999
	E	1,365	220	574	1,999	1,371	219	638	1,999
	F	1,364	226	608	1,999	1,370	220	535	1,999
	G	1,364	221	638	1,999	1,373	212	608	1,999
	H	1,357	223	574	1,999	1,371	218	535	1,999
	I	1,363	222	666	1,999	1,375	218	638	1,999
	J	1,361	220	638	1,999	1,370	219	608	1,999
	Total	1,361	222	574	1,999	1,371	218	440	1,999

Table 27
Percentiles of Scale Score Ranges

Percentile	Reading Score Range	Mathematics Score Range
1	300-815	200-850
2	816-847	851-894
3	848-872	895-926
4	873-893	927-953
5	894-912	954-975
6	913-929	976-994
7	930-945	995-1013
8	946-961	1014-1029
9	962-978	1030-1043
10	979-991	1044-1056
11	992-1006	1057-1069
12	1007-1020	1070-1080
13	1021-1034	1081-1091
14	1035-1046	1092-1102
15	1047-1057	1103-1112
16	1058-1070	1113-1122
17	1071-1080	1123-1131
18	1081-1090	1132-1140
19	1091-1100	1141-1148
20	1101-1110	1149-1156
21	1111-1123	1157-1164
22	1124-1128	1165-1171
23	1129-1141	1172-1178
24	1142-1145	1179-1185
25	1146-1158	1186-1192
26	1159-1162	1193-1198
27	1163-1169	1199-1205
28	1170-1179	1206-1211
29	1180-1183	1212-1217
30	1184-1191	1218-1223
31	1192-1201	1224-1229
32	1202-1203	1230-1235
33	1204-1209	1236-1240
34	1210-1220	1241-1245
35	1221-1223	1246-1251
36	1224-1227	1252-1256
37	1228-1236	1257-1261
38	1237-1244	1262-1266
39	1245-1247	1267-1271
40	1248-1250	1272-1276

Table 27 (cont.)

Percentiles of Scale Score Ranges

Percentile	Reading Score Range	Mathematics Score Range
41	1251-1258	1277-1281
42	1259-1266	1282-1286
43	1267-1268	1287-1291
44	1269-1271	1292-1296
45	1272-1276	1297-1300
46	1277-1284	1301-1304
47	1285-1290	1305-1309
48	1291-1293	1310-1313
49	1294-1296	1314-1318
50	1297-1299	1319-1322
51	1300-1307	1323-1326
52	1308-1313	1327-1330
53	1314-1317	1331-1335
54	1318-1320	1336-1339
55	1321-1323	1340-1343
56	1324-1327	1344-1347
57	1328-1335	1348-1351
58	1336-1340	1352-1356
59	1341-1343	1357-1360
60	1344-1347	1361-1364
61	1348-1351	1365-1368
62	1352-1353	1369-1372
63	1354-1363	1373-1376
64	1364-1367	1377-1380
65	1368-1370	1381-1384
66	1371-1376	1385-1388
67	1377-1381	1389-1393
68	..	1394-1398
69	1382-1391	1399-1402
70	1392-1396	1403-1407
71	1397-1400	1408-1411
72	1401-1406	1412-1416
73	1407-1412	1417-1422
74	1413-1415	1423-1427
75	1416-1426	1428-1433
76	1427-1430	1434-1439
77	1431-1435	1440-1445
78	1436-1442	1446-1452
79	1443-1448	1453-1461
80	1449-1455	1462-1469

Table 27 (cont.)

Percentiles of Scale Score Ranges

Percentile	Reading Score Range	Mathematics Score Range
81	1456-1466	1470-1477
82	1467-1469	1478-1486
83	1470-1479	1487-1495
84	1480-1488	1496-1504
85	1489-1497	1505-1513
86	1498-1507	1514-1522
87	1508-1515	1523-1533
88	1516-1528	1534-1544
89	1529-1540	1545-1555
90	1541-1553	1556-1566
91	1554-1564	1567-1580
92	1565-1577	1581-1595
93	1578-1595	1596-1610
94	1596-1614	1611-1627
95	1615-1635	1628-1652
96	1636-1661	1653-1677
97	1662-1702	1678-1709
98	1703-1740	1710-1763
99	1741-1999	1764-1999

Table 28
Percentiles of Reading Performance Cut Scores

		Below Basic	Basic	Proficient	Advanced
All students		17	15	36	32
Ethnicity	Caucasian	11	13	38	38
	African American	36	23	30	11
	Hispanic	39	22	29	10
	Other ethnicities	15	16	36	34
Gender	Male	19	16	36	29
	Female	14	14	37	35

Table 29
Percentiles of Mathematics Performance Cut Scores

		Below Basic	Basic	Proficient	Advanced
All students		8	12	31	50
Ethnicity	Caucasian	4	9	30	58
	African American	22	24	34	21
	Hispanic	21	22	34	23
	Other ethnicities	6	10	30	55
Gender	Male	8	11	30	52
	Female	8	13	31	48

* The total sum may not be 100 due to rounding percent.

Table 30
*Number of Items per Each Reading Assessment Anchor **

Objective	Form A		Form B		Form C		Form D		Form E	
	MC	OE	MC	OE	MC	OE	MC	OE	MC	OE
1.A.1	14	2	12	1	11	1	14	2	12	1
1.A.2	14	0	14	1	13	1	15	0	14	1
2.B.1	9	0	11	0	12	0	9	0	10	0
2.B.2	1	0	1	0	1	0	1	0	2	0
2.B.3	2	0	1	0	3	0	1	0	1	0
Total	40	2	39	2	40	2	40	2	39	2

* Table does not include FT items.

Table 31

*Number of Items per each Mathematics Assessment Anchor**

Standards	A		B		C		D		E		F		G		H		I		J	
	MC	OE	MC	OE	MC	OE	MC	OE	MC	OE	MC	OE	MC	OE	MC	OE	MC	OE	MC	OE
1.A.1	14		13		13		13		13		11		12		12		12		12	
1.A.2	4	1	3	1	4	2	4	1	4	1	5	2	5	1	3	1	3	1	4	1
1.A.3	9		10		9		10		10		11		10		12		12		10	
2.B.1	7		6		7		7		7		7		7		6	1	6	1	7	
2.B.2	3		4		4		3		3		4		3		5		4		4	
3.C.1	6		8		8		7		7		7		8	1	7		7		8	1
3.C.2	4		3		3		3		3		3		3		3		3		3	
4.D.1	6	1	4	2	5	1	6	1	5	2	5	1	4	1	5	1	5	1	5	1
4.D.2	1		2		2		1		1		1		2		1		1		1	
5.E.1	6	1	7		5		6	1	7		6		6		6		7		6	
5.E.2	4		4		4		4		4		4		4		4		4		4	
Total	64	3	64	3	64	3	64	3	64	3	64	3	64	3	64	3	64	3	64	3

Table 32
Summary Statistics for Reading Reporting Categories/Assessment Anchors

Form	N	Reporting Category	Assessment Anchor	Total Number of Items	Total Score Points	Mean	Mean P-Value	SD
A	25743	R3A		30	34	24.28	0.71	6.65
			R3A1	16	20	14.21	0.71	3.86
			R3A2	14	14	10.08	0.72	3.18
		R3B		12	12	8.17	0.68	2.77
			R3B1	9	9	6.61	0.73	2.19
			R3B2	1	1	0.29	0.29	0.45
			R3B3	2	2	1.27	0.63	0.75
B	25525	R3A		28	32	22.82	0.71	6.13
			R3A1	13	15	10.55	0.70	2.91
			R3A2	15	17	12.27	0.72	3.64
		R3B		13	13	9.42	0.72	2.79
			R3B1	11	11	8.33	0.76	2.44
			R3B2	1	1	0.30	0.30	0.46
			R3B3	1	1	0.79	0.79	0.41
C	25226	R3A		26	30	22.28	0.74	6.00
			R3A1	12	14	10.20	0.73	2.86
			R3A2	14	16	12.08	0.76	3.49
		R3B		16	16	11.42	0.71	3.58
			R3B1	12	12	9.17	0.76	2.73
			R3B2	1	1	0.30	0.30	0.46
			R3B3	3	3	1.95	0.65	1.01
D	24978	R3A		31	35	25.68	0.73	6.50
			R3A1	16	20	14.32	0.72	3.79
			R3A2	15	15	11.36	0.76	3.10
		R3B		11	11	7.68	0.70	2.51
			R3B1	9	9	6.68	0.74	2.14
			R3B2	1	1	0.30	0.30	0.46
			R3B3	1	1	0.70	0.70	0.46
E	24689	R3A		28	32	23.37	0.73	6.20
			R3A1	13	15	11.07	0.74	2.98
			R3A2	15	17	12.29	0.72	3.62
		R3B		13	13	9.17	0.71	2.74
			R3B1	10	10	7.14	0.71	2.30
			R3B2	2	2	1.23	0.62	0.54
			R3B3	1	1	0.79	0.79	0.40

Table 33

Summary Statistics for Mathematics Reporting Categories/Assessment Anchors
*(N = 126,577)**

Reporting Category	Assessment Anchor	Total Number of Items	Total Score Points	Mean	Mean P-Value	SD
M3A		23	26	20.09	0.77	4.23
	M3A1	10	10	8.22	0.82	1.88
	M3A2	3	3	2.57	0.86	0.74
	M3A3	9	9	7.10	0.79	1.74
M3B		9	9	7.52	0.84	1.67
	M3B1	6	6	4.92	0.82	1.35
	M3B2	3	3	2.60	0.87	0.63
M3C		8	8	6.74	0.84	1.44
	M3C1	5	5	4.02	0.80	1.21
	M3C2	3	3	2.72	0.91	0.52
M3D		6	9	5.90	0.66	1.85
	M3D1	4	4	3.10	0.78	1.09
	M3D2	1	1	0.76	0.76	0.43
M3E		9	9	7.41	0.82	1.89
	M3E1	5	5	3.90	0.78	1.31
	M3E3	4	4	3.51	0.88	0.87

* Table includes common items only

Table 34
Factor Analysis Results for Reading Reporting Categories

Communalities

	Initial	Extraction
R3A	.440	.662
R3B	.440	.662

Extraction Method: Principal Axis Factoring.

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.663	83.152	83.152	1.324	66.217	66.217
2	.337	16.848	100.000			

Extraction Method: Principal Axis Factoring.

Factor Matrix(a)

	Factor
	1
R3A	.814
R3B	.814

Extraction Method: Principal Axis Factoring.
a. 1 factors extracted. 8 iterations required.

Table 35

Factor Analysis Results for Reading Assessment Anchors

Communalities

	Initial
R3A1	.304
R3A2	.621
R3B1	.588
R3B2	.095
R3B3	.256

Extraction Method: Principal Axis Factoring.

Total Variance Explained

Factor	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.498	49.958	49.958
2	1.027	20.532	70.491
3	.749	14.979	85.470
4	.488	9.752	95.222
5	.239	4.778	100.000

Extraction Method: Principal Axis Factoring.

Factor Matrix(a)

- a Attempted to extract 2 factors. In iteration 25, the communality of a variable exceeded 1.0.
Extraction was terminated.

Table 36
Factor Analysis Results for Mathematics Reporting Categories

Communalities

	Initial	Extraction
M3A	.709	.828
M3B	.543	.603
M3C	.406	.445
M3D	.541	.595
M3E	.611	.685

Extraction Method: Principal Axis Factoring.

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.505	70.098	70.098	3.156	63.117	63.117
2	.527	10.531	80.629			
3	.400	7.998	88.627			
4	.345	6.906	95.533			
5	.223	4.467	100.000			

Extraction Method: Principal Axis Factoring.

Factor Matrix(a)

	Factor 1
M3A	.910
M3B	.776
M3C	.667
M3D	.771
M3E	.828

Extraction Method: Principal Axis Factoring.
a. 1 factors extracted. 6 iterations required.

Table 37
Factor Analysis Results for Mathematics Assessment Anchors

Communalities

	Initial	Extraction
M3A1	.624	.702
M3A2	.301	.331
M3A3	.542	.600
M3B1	.502	.544
M3B2	.214	.233
M3C1	.357	.394
M3C2	.159	.166
M3D1	.479	.512
M3D2	.259	.285
M3E1	.562	.619
M3E3	.333	.365

Extraction Method: Principal Axis Factoring.

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.260	47.816	47.816	4.750	43.184	43.184
2	.862	7.839	55.654			
3	.746	6.781	62.435			
4	.702	6.378	68.813			
5	.660	5.999	74.812			
6	.623	5.662	80.474			
7	.591	5.370	85.844			
8	.436	3.967	89.812			
9	.425	3.862	93.674			
10	.372	3.385	97.059			
11	.324	2.941	100.000			

Extraction Method: Principal Axis Factoring.

Table 35 (cont.)

Factor Analysis Results for Mathematics Assessment Anchors

Factor Matrix(a)

	Factor
	1
M3A1	.838
M3A2	.575
M3A3	.774
M3B1	.738
M3B2	.482
M3C1	.628
M3C2	.407
M3D1	.716
M3D2	.534
M3E1	.787
M3E3	.604

Extraction Method: Principal Axis Factoring.
a. 1 factors extracted. 4 iterations required.

Figure 1
Reading anchor items' B parameter correlations

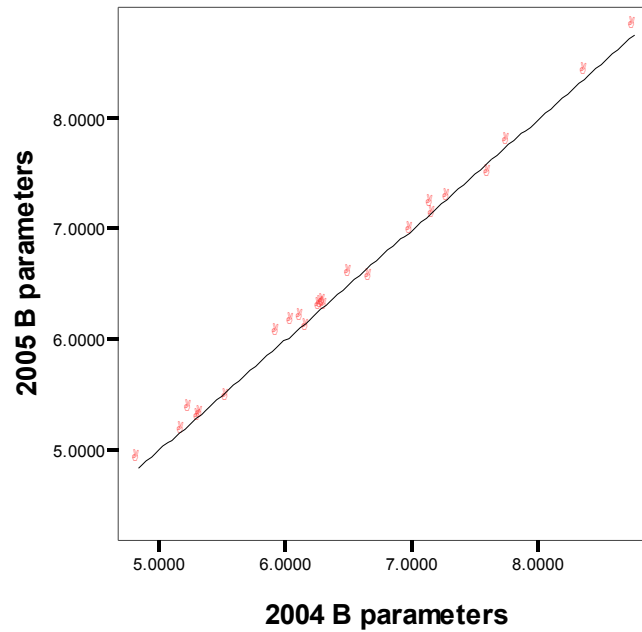


Figure 2
Mathematics anchor items' B parameter correlations

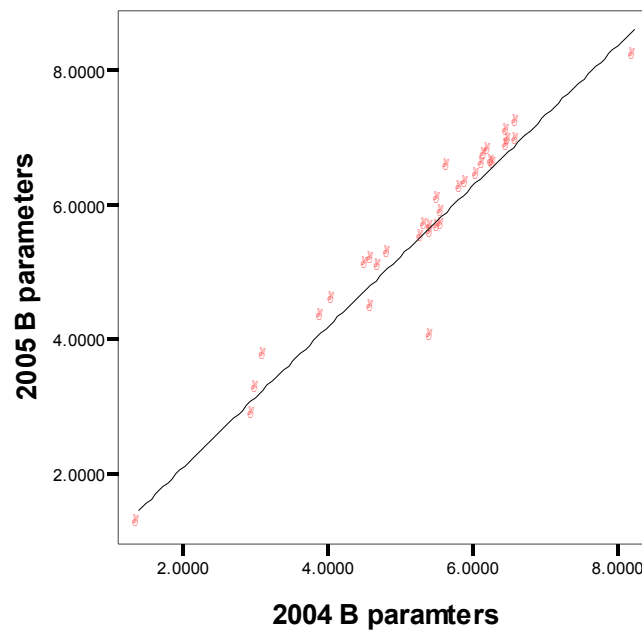


Figure 3
Reading scale scores and SEMs across five forms

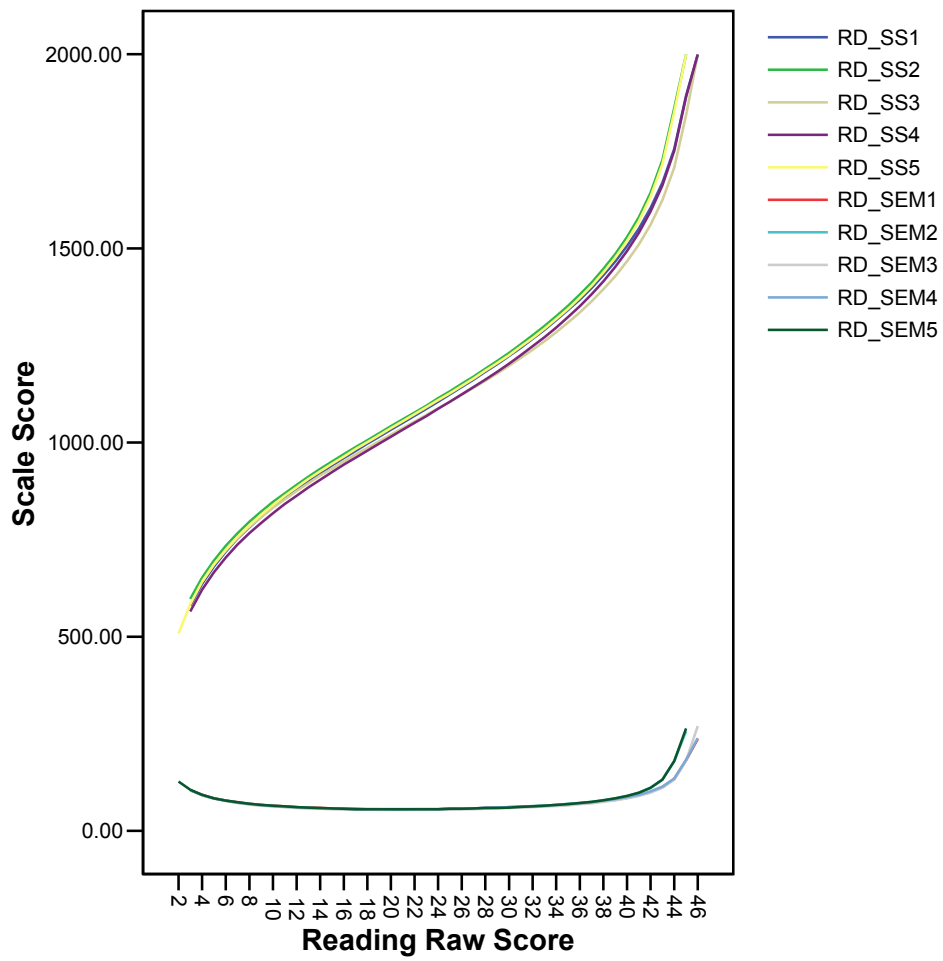


Figure 4
Mathematics scale score and SEM

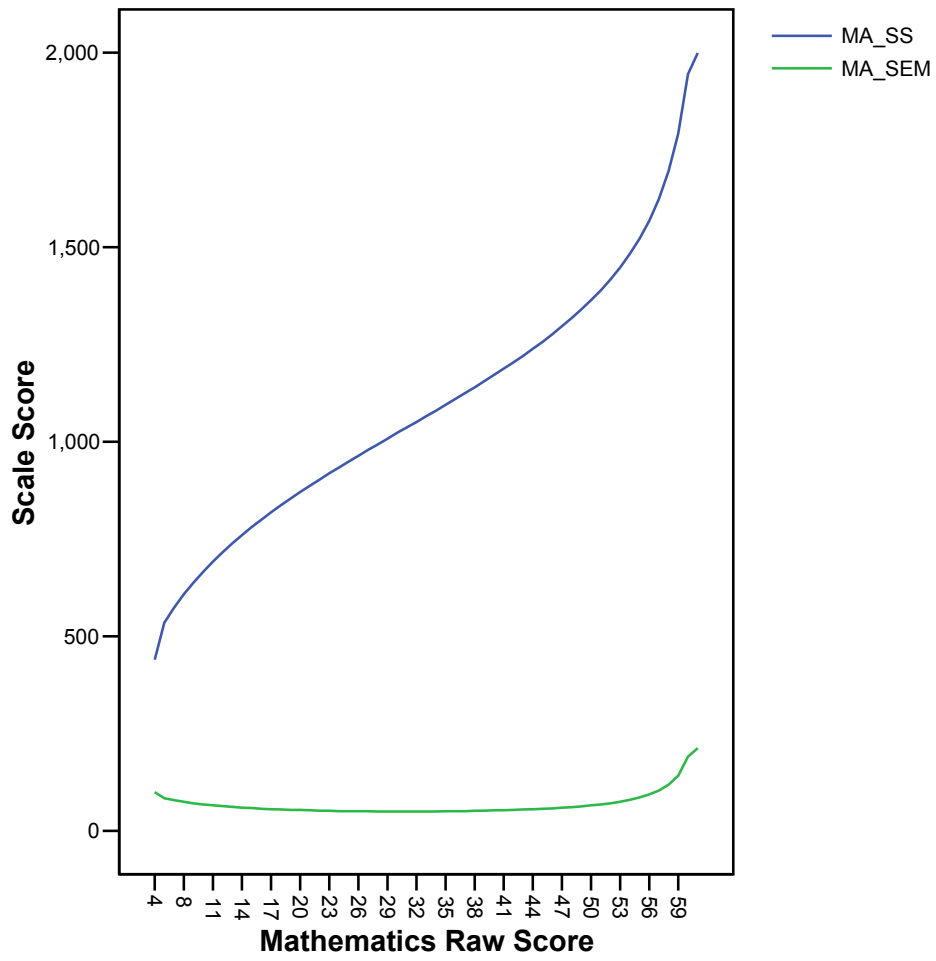


Figure 5
Reading Form 1 Raw Score and Scale Score Distribution

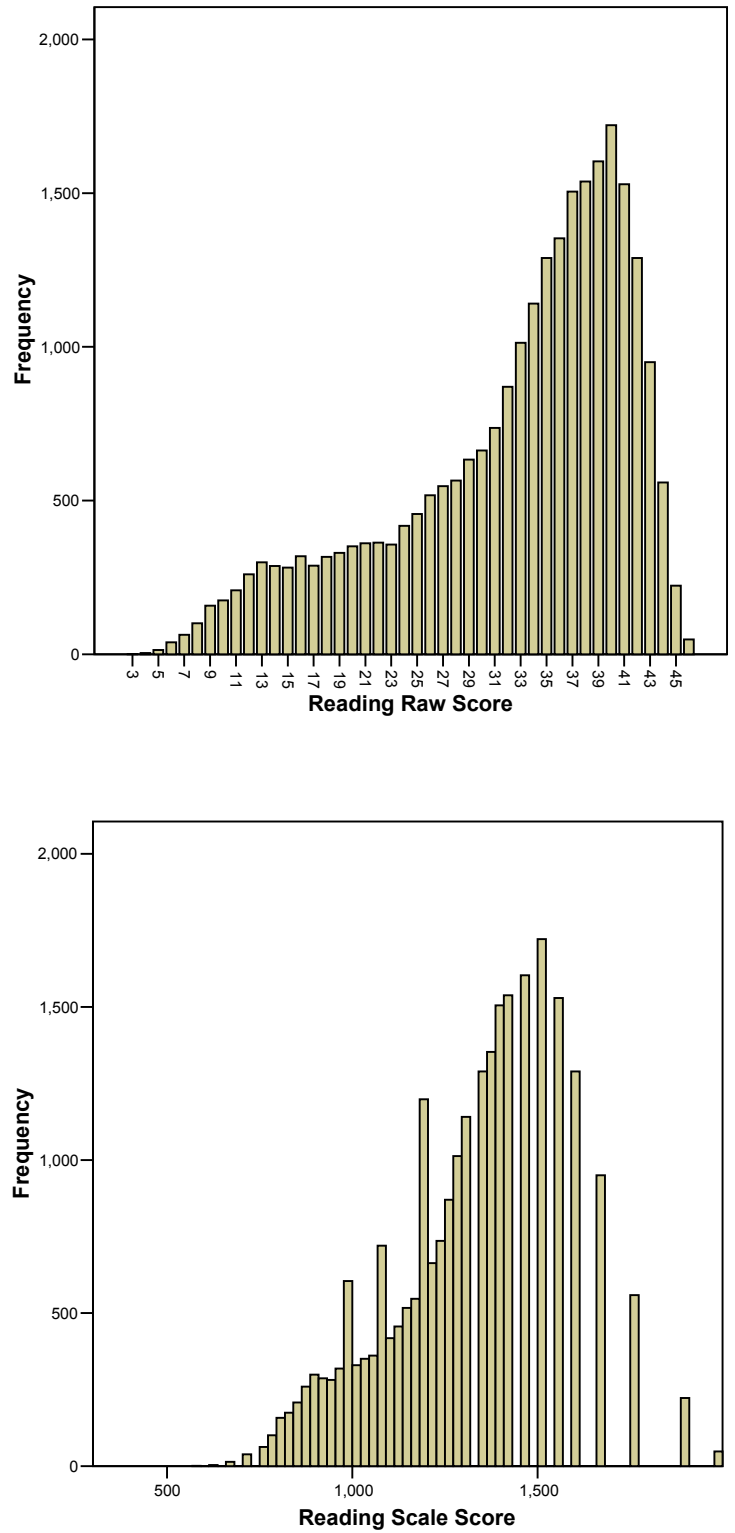


Figure 6
Reading Form 2 Raw Score and Scale Score Distribution

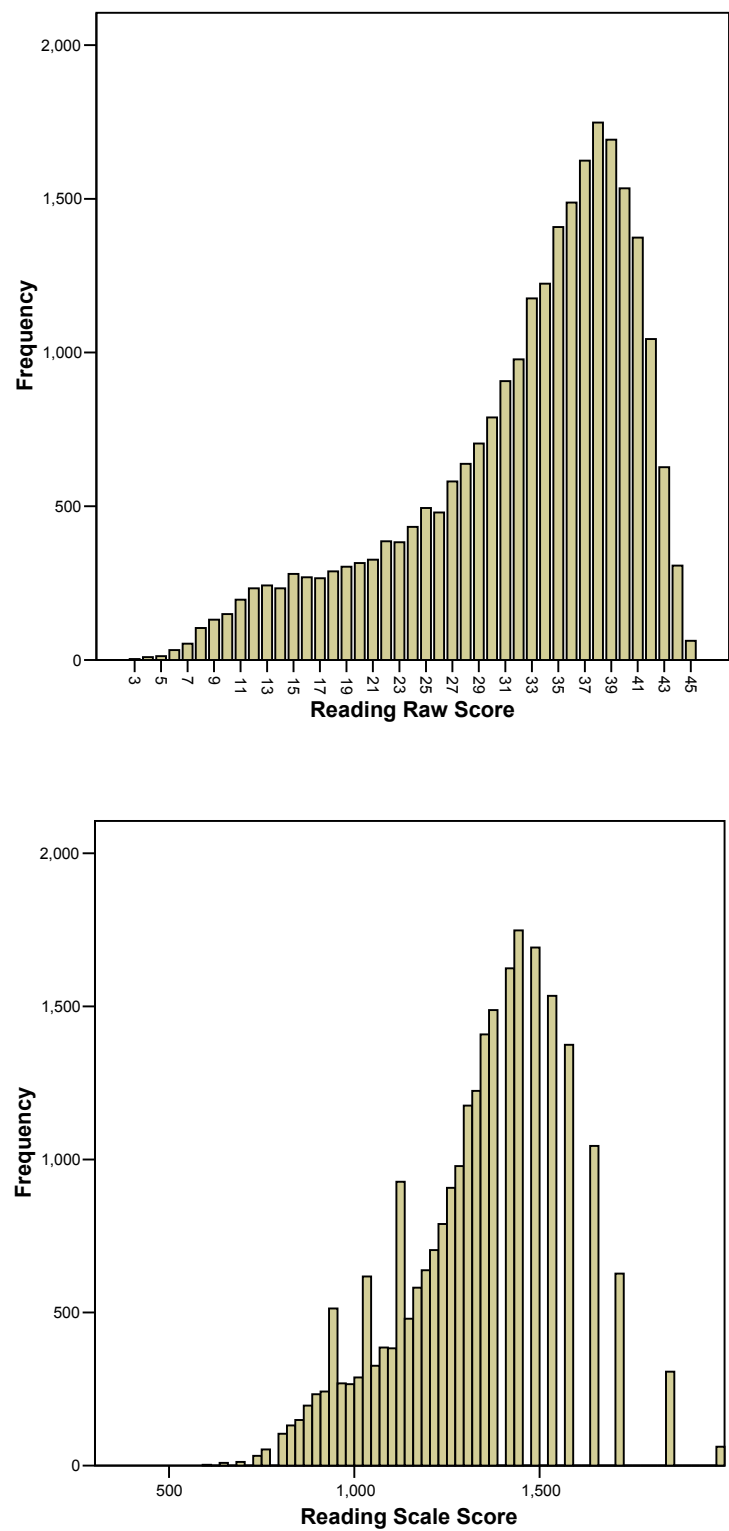


Figure 7
Reading Form 3 Raw Score and Scale Score Distribution

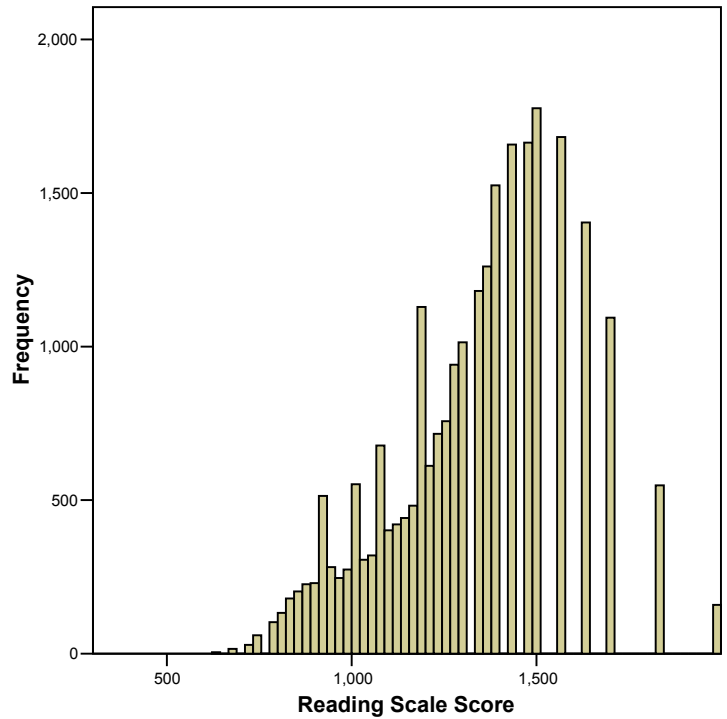
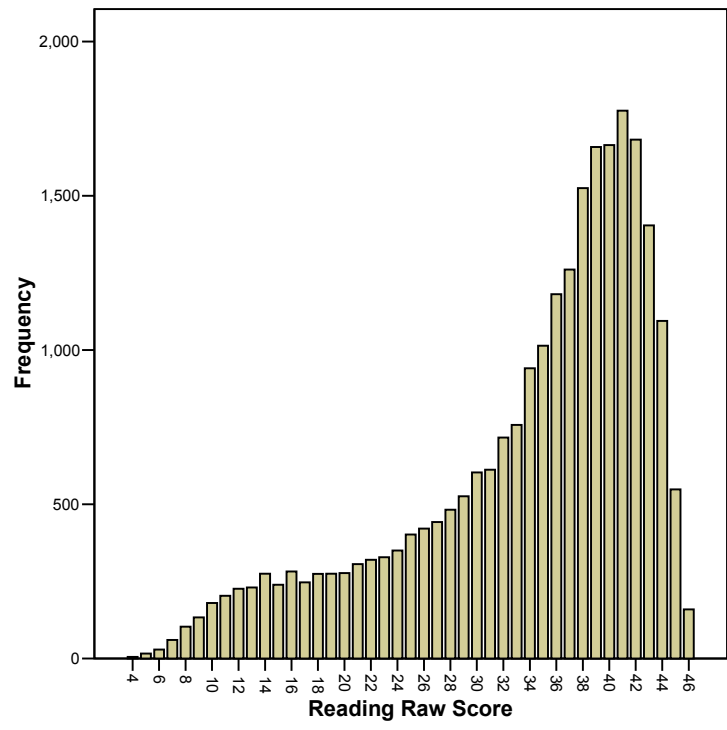


Figure 8
Reading Form 4 Raw Score and Scale Score Distribution

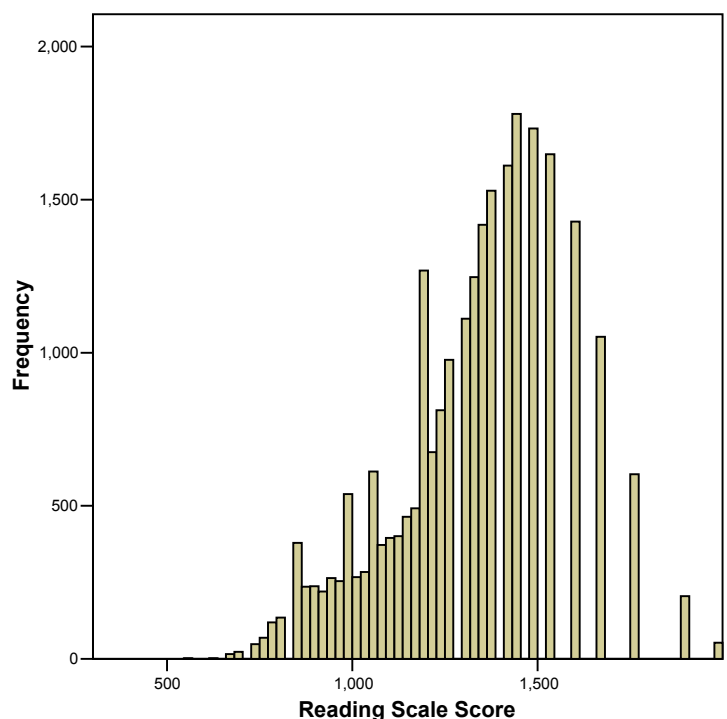
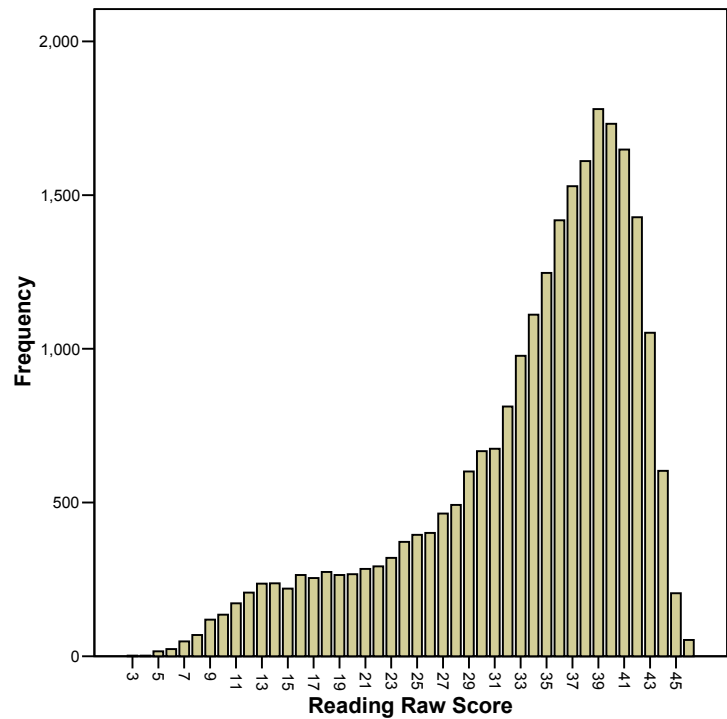


Figure 9
Reading Form 5 Raw Score and Scale Score Distribution

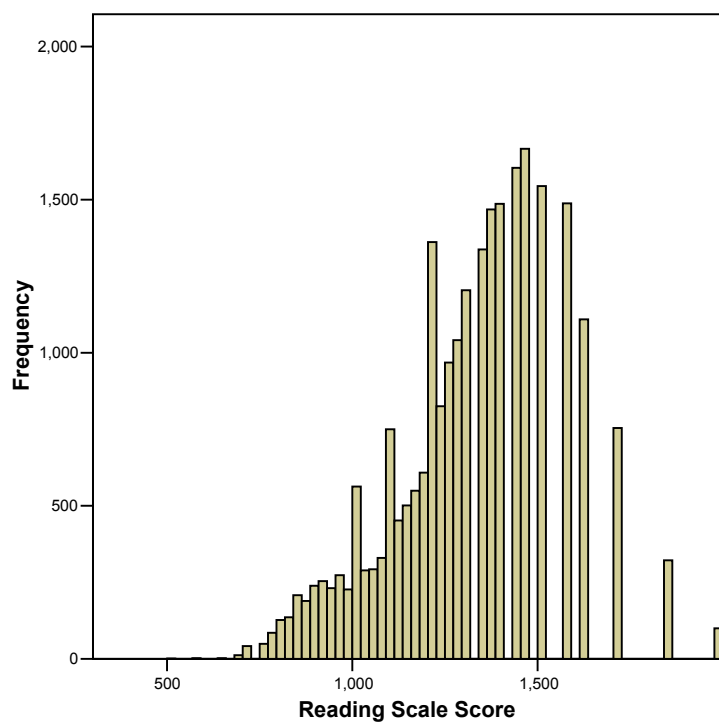
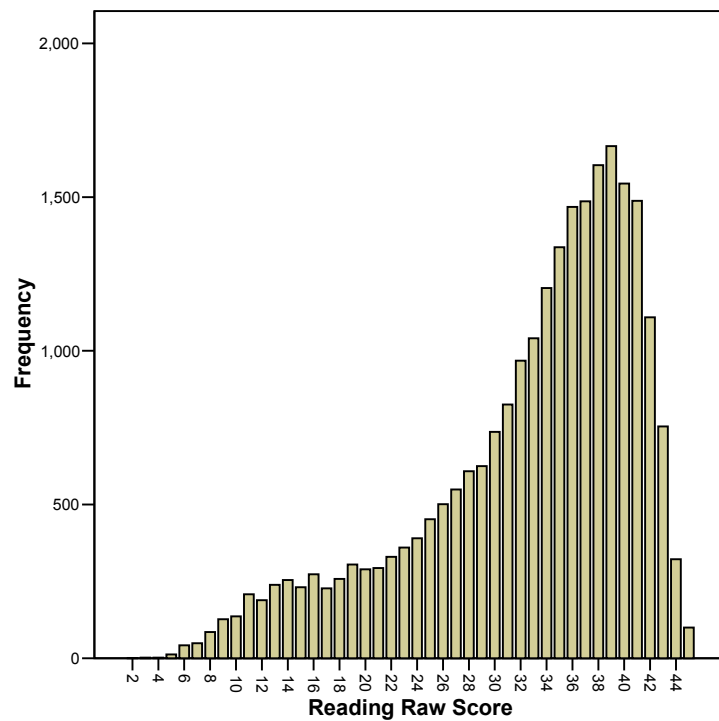


Figure 10
Mathematics Raw Score and Scale Score Distribution

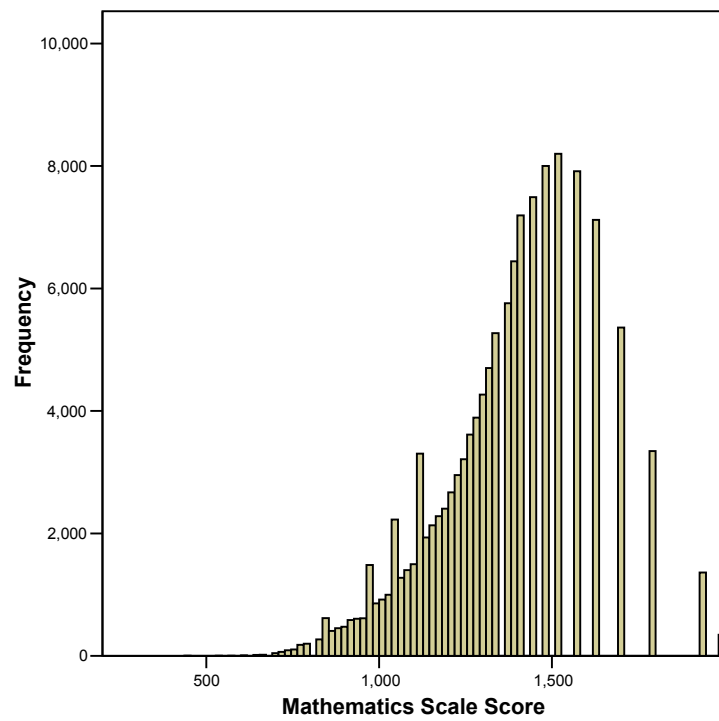
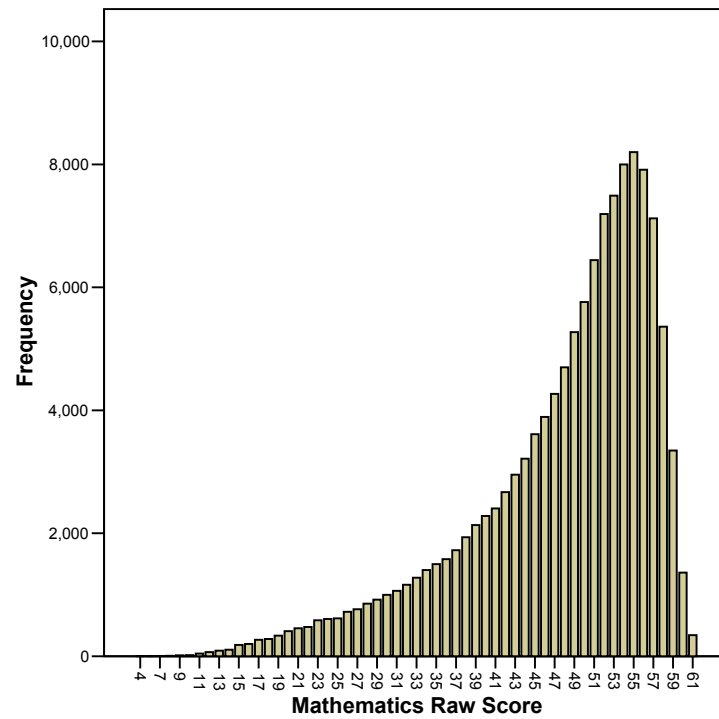


Figure 11
Scree Plots for Reading Reporting Categories and Assessment Anchors

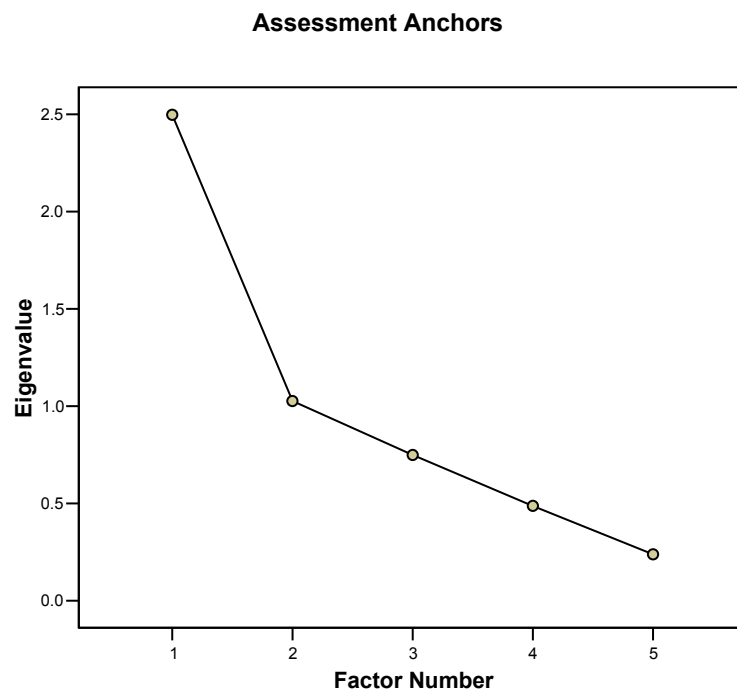
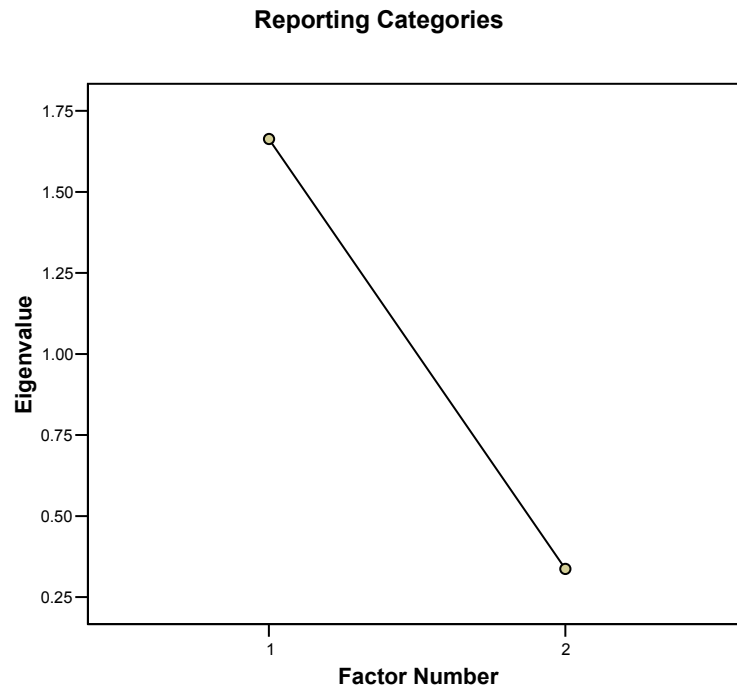


Figure 12
Scree Plots for Mathematics Reporting Categories and Assessment Anchors

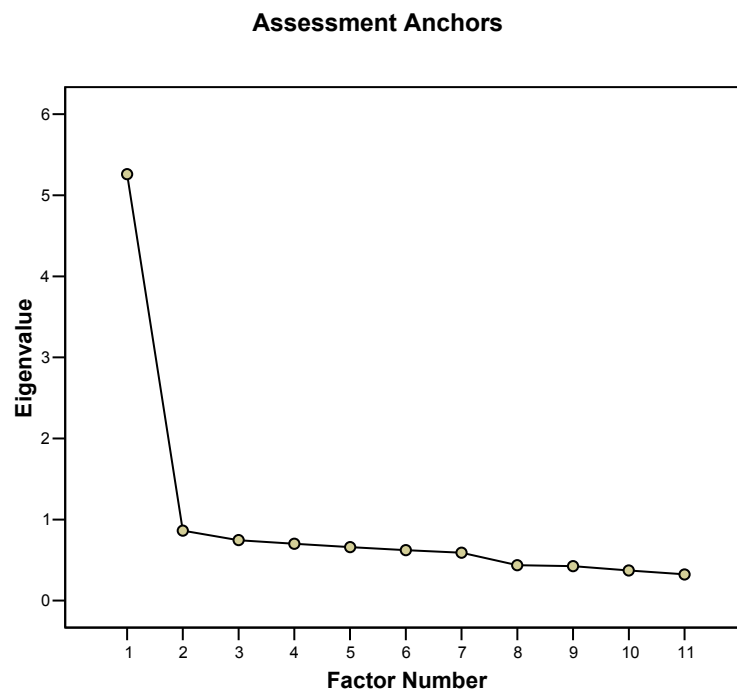
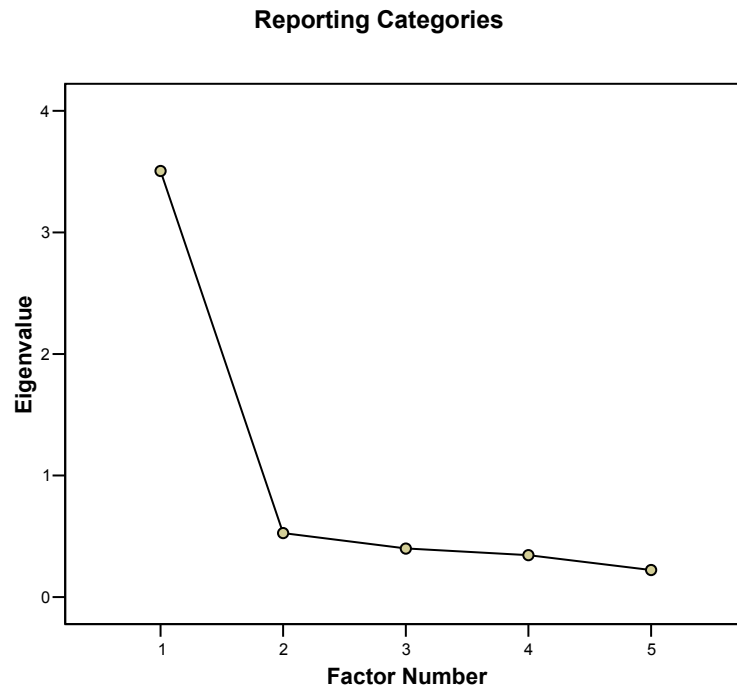
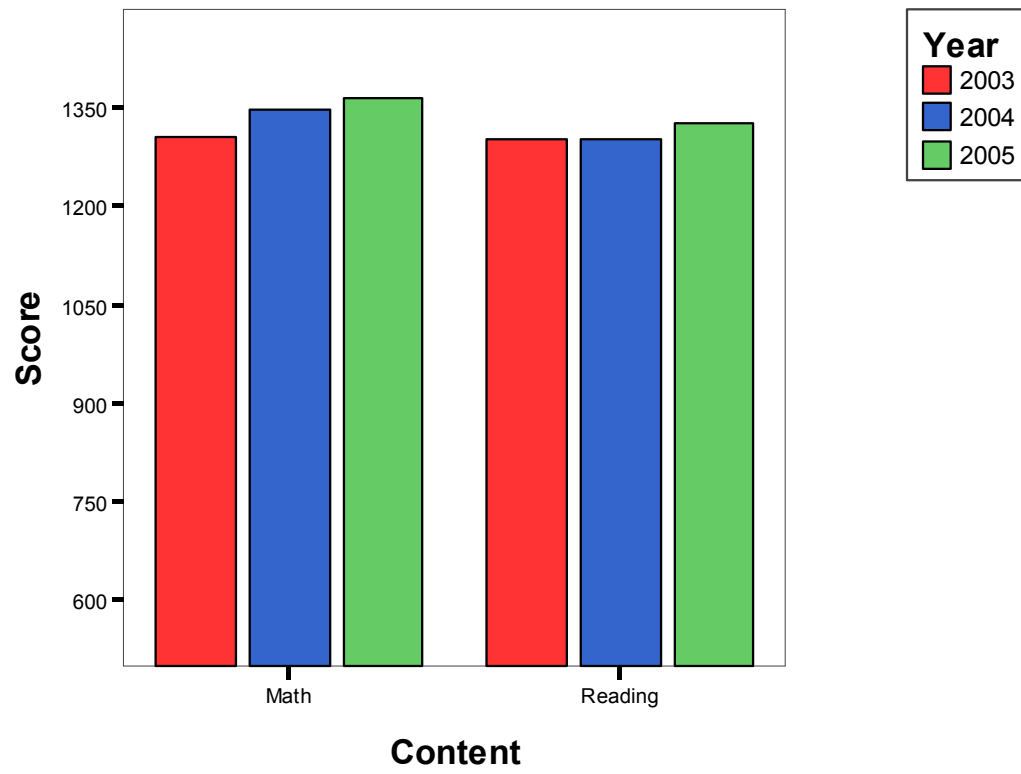


Figure 13
State Mean Scale Scores



References

- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. *Psychometrika*, 46, 443–459.
- Burket, G. R. (1991). PARDUX [Computer program]. Unpublished.
- Fleiss, J.L., Cohen, J., and Everitt, B.S. (1969). Large-Sample Standard Errors of Kappa and Weighted Kappa. *Psychological Bulletin*, 72, 323–327.
- Kolen, J. Michael., and Brennan, L. Robert. (1995). *Test equating*. Springer : New York.
- Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement*, 14, 59–71.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16, 159–176.
- Stocking, M.L., and Lord, F.M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, 7.
- Thissen, D. (1982). Marginal maximum-likelihood estimation for the one-parameter logistic model. *Psychometrika*, 47, 175–186.
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 2, 125–145.
- Yen, W. M. (1991a). *How to evaluate FIT for Maryland*. Unpublished manuscript.
- Zwick, R., Donoghue, J.R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, 30, 233–251.

Appendix

Reading and Mathematics Academic Standards

The Pennsylvania Grade 3 Reading Test measures the following Pennsylvania Assessment

Anchors and Reporting Categories:

- R3.A Comprehension and Reading Skills
 - R3.A.1 Understanding fiction text appropriate to grade level.
 - R3.A.2 Understanding nonfiction text appropriate to grade level.
- R3.B Interpretation and Analysis of Fiction and Nonfiction Text
 - R3.B.1 Identify components within text.
 - R3.B.2 Identify literary devices.
 - R3.B.3 Identify concepts and organization of nonfiction text.

The Pennsylvania Grade 3 Mathematics Test measures the following Pennsylvania Assessment

Anchors and Reporting Categories:

- MA. Numbers and Operations
 - MA.1 Demonstrate understanding of numbers, ways of representing numbers, relationships among numbers and number systems.
 - MA.2 Understand the meanings of operations, use operations and understand how they relate to each other.
 - MA.3 Compute accurately and fluently and make reasonable estimates.
- MB. Measurement
 - MB.1 Demonstrate an understanding of measurable attributes of objects and figures, and the units, systems and processes of measurement.
 - MB.2 Apply appropriate techniques, tools and formulas to determine measurements.

- MC. Geometry
 - MC.1 Analyze characteristics and properties of 2- and 3-dimensional geometric shapes and demonstrate understanding of geometric relationships.
 - MC.2 Identify and/or apply concepts of transformations or symmetry.
- MD. Algebraic Concepts
 - MD.1 Demonstrate an understanding of patterns, relations and functions.
 - MD.2 Represent and/or analyze mathematical situations using numbers, symbols, words, tables, and/or graphs.
- ME. Data Analysis and Probability
 - ME.1 Formulate or answer questions that can be addressed with data and/or organize, display, interpret, or analyze data.
 - ME.3 Understand and/or apply basic concepts of probability or outcomes.

Addendum

Summary Statistics for Raw Score and Scale Score by Four NCLB Subgroups: ELP, Migrant Status, Disability, Economic Disadvantage Status

Table of Contents

Table 38 <i>Raw Score Descriptive Statistics by ELP Status</i>	77
Table 39 <i>Scale Score Descriptive Statistics by ELP Status</i>	78
Table 40 <i>Raw Score Descriptive Statistics by Migrant Status</i>	79
Table 41 <i>Scale Score Descriptive Statistics by Migrant Status</i>	80
Table 42 <i>Raw Score Descriptive Statistics by Disability Status</i>	81
Table 43 <i>Scale Score Descriptive Statistics by Disability Status</i>	82
Table 44 <i>Raw Score Descriptive Statistics by Economic Disability Status</i>	83
Table 45 <i>Scale Score Descriptive Statistics by Economic Disability Status</i>	84

Table 38
Raw Score Descriptive Statistics by ELP Status

Content	Form	English Language Proficient				English Language Learners			
		N	Mean	SD	Mean P-Value	N	Mean	SD	Mean P-Value
Reading	A	24882	32.73	8.93	0.71	861	24.45	9.48	0.53
	B	24667	32.50	8.38	0.71	858	24.62	9.68	0.54
	C	24363	34.03	9.03	0.74	863	24.51	10.30	0.53
	D	24127	33.64	8.48	0.73	851	25.29	9.81	0.55
	E	23879	32.81	8.39	0.71	810	24.39	9.62	0.53
Mathematics (Both Common items and Matrix items)	A	12827	57.98	11.56	0.77	480	49.43	13.86	0.66
	B	12652	58.62	11.31	0.78	464	51.24	13.50	0.68
	C	12456	59.80	10.98	0.80	472	50.57	13.55	0.67
	D	12368	59.07	11.05	0.79	457	49.34	14.17	0.66
	E	12216	59.05	11.26	0.79	413	49.47	13.89	0.66
	F	12086	58.78	11.52	0.78	413	49.52	14.23	0.66
	G	12057	58.61	11.15	0.78	428	49.18	13.97	0.66
	H	11981	59.38	11.35	0.79	422	50.54	13.89	0.67
	I	11818	59.29	11.29	0.79	420	51.93	13.56	0.69
	J	11706	58.96	10.95	0.79	426	49.78	13.35	0.66
Mathematics (Common Items only)	Total	122182	47.92	9.33	0.79	4395	40.61	11.42	0.67

***English Language Learners” includes students who are in their first year of enrollment as those NOT in their first year of enrollment.

Table 39
Scale Score Descriptive Statistics by ELP Status

Content	Test Form	English Language Proficient					English Language Learners				
		N	Mean	SD	MIN	MAX	N	Mean	SD	MIN	MAX
Reading	A	24882	1328.17	227.09	580	1999	861	1129.74	205.76	681	1893
	B	24667	1330.73	215.60	597	1999	858	1143.11	211.53	652	1999
	C	24363	1341.76	235.31	642	1999	863	1117.50	219.96	686	1843
	D	24127	1335.35	223.23	565	1999	851	1130.00	213.59	666	1751
	E	23879	1333.45	218.46	508	1999	810	1133.17	210.80	686	1851
	Total	121918	1333.86	224.11	508	1999	4243	1130.66	212.46	652	1999
Mathematics	A	12827	1359.29	219.97	608	1999	480	1204.48	215.05	716	1945
	B	12652	1370.85	220.01	574	1999	464	1243.40	226.45	716	1999
	C	12458	1373.10	214.03	638	1999	472	1209.29	215.63	666	1945
	D	12370	1373.66	218.03	440	1999	457	1207.54	227.27	716	1945
	E	12220	1373.15	218.00	574	1999	413	1207.79	214.07	692	1945
	F	12088	1370.89	221.06	535	1999	413	1215.70	229.71	666	1945
	G	12059	1373.21	215.02	608	1999	428	1211.81	222.41	666	1999
	H	11983	1368.77	218.61	535	1999	422	1212.60	227.04	638	1999
	I	11819	1373.12	218.98	638	1999	420	1238.13	222.01	739	1999
	J	11706	1371.01	217.70	608	1999	426	1206.52	213.94	692	1792
Total		122182	1370.65	218.19	440	1999	4395	1215.70	221.54	638	1999

**"English Language Learners" includes students who are in their first year of enrollment as those NOT in their first year of enrollment.

Table 40
Raw Score Descriptive Statistics by Migrant Status

Content	Form	Non-migrant				Migrant			
		N	Mean	SD	Mean P-Value	N	Mean	SD	Mean P-Value
Reading	A	24879	32.64	8.98	0.71	864	27.25	10.13	0.59
	B	24724	32.39	8.48	0.70	801	27.59	9.34	0.60
	C	24328	33.88	9.16	0.74	898	28.90	10.15	0.63
	D	24185	33.54	8.55	0.73	793	27.88	10.05	0.61
	E	23879	32.68	8.51	0.71	810	28.30	9.22	0.62
Mathematics (Both Common items and Matrix items)	A	12842	57.94	11.58	0.77	465	50.34	14.02	0.67
	B	12694	58.61	11.32	0.78	422	50.78	13.43	0.68
	C	12483	59.72	11.08	0.80	445	52.10	12.54	0.69
	D	12430	58.97	11.18	0.79	395	50.90	12.93	0.68
	E	12216	58.96	11.37	0.79	413	52.19	12.69	0.70
	F	12100	58.71	11.61	0.78	399	51.30	13.28	0.68
	G	12098	58.50	11.26	0.78	387	51.74	13.41	0.69
	H	11938	59.34	11.38	0.79	465	52.40	13.71	0.70
	I	11835	59.29	11.31	0.79	403	51.53	13.01	0.69
	J	11727	58.85	11.06	0.78	405	52.50	12.68	0.70
Mathematics (Common Items only)	Total	122376	47.87	9.39	0.78	4201	41.90	10.96	0.69

**"Migrant" includes all students who are migrants at the school, district, and/or state level (i.e. they initially enrolled in the school, district, or state of residence after October 1, 2004).

Table 41
Scale Score Descriptive Statistics by Migrant Status

Content	Test Form	Non-migrant					Migrant				
		N	Mean	SD	MIN	MAX	N	Mean	SD	MIN	MAX
Reading	A	24879	1325.91	227.82	580	1999	864	1195.41	232.83	636	1999
	B	24724	1328.22	217.32	597	1999	801	1207.45	209.68	597	1861
	C	24328	1338.59	237.40	642	1999	898	1212.20	230.50	686	1999
	D	24185	1332.88	224.50	565	1999	793	1190.18	228.23	565	1751
	E	23879	1330.53	220.47	508	1999	810	1219.12	212.69	724	1999
	Total	121995	1331.19	225.65	508	1999	4166	1204.96	223.44	565	1999
Mathematics	A	12842	1358.40	220.30	638	1999	465	1224.14	220.30	608	1945
	B	12694	1370.78	220.21	574	1999	422	1232.71	218.44	716	1945
	C	12485	1371.92	215.11	638	1999	445	1232.53	205.46	666	1792
	D	12431	1372.21	219.20	608	1999	396	1227.32	215.58	440	1999
	E	12220	1371.93	219.23	574	1999	413	1243.67	201.11	692	1792
	F	12101	1370.11	222.18	535	1999	400	1234.41	210.01	666	1945
	G	12100	1371.45	216.35	608	1999	387	1249.59	212.99	716	1945
	H	11940	1368.05	219.44	535	1999	465	1245.47	221.02	608	1945
	I	11836	1373.16	219.36	638	1999	403	1231.28	208.09	692	1945
	J	11727	1369.19	218.91	608	1999	405	1250.49	210.32	666	1792
Total		122376	1369.67	219.07	535	1999	4201	1237.02	212.55	440	1999

***"Migrant" includes all students who are migrants at the school, district, and/or state level (i.e. they initially enrolled in the school, district, or state of residence after October 1, 2004).

Table 42
Raw Score Descriptive Statistics by Disability Status

Content	Form	Not Disabled				Disabled			
		N	Mean	SD	Mean P-Value	N	Mean	SD	Mean P-Value
Reading	A	21854	33.83	8.05	0.74	3889	24.73	10.50	0.54
	B	22008	33.50	7.55	0.73	3517	24.33	10.07	0.53
	C	21675	35.03	8.15	0.76	3551	25.58	11.17	0.56
	D	21582	34.59	7.62	0.75	3396	25.54	10.56	0.56
	E	21337	33.77	7.58	0.73	3352	24.64	10.13	0.54
Mathematics (Both Common items and Matrix items)	A	11098	59.28	10.40	0.79	2209	49.61	14.55	0.66
	B	11249	59.82	10.25	0.80	1867	49.58	14.22	0.66
	C	11112	60.78	10.04	0.81	1816	51.42	14.29	0.69
	D	11151	59.96	10.31	0.80	1674	50.47	14.00	0.67
	E	10885	60.12	10.30	0.80	1744	50.10	14.33	0.67
	F	10808	59.73	10.73	0.80	1691	50.44	14.45	0.67
	G	10823	59.49	10.38	0.79	1662	50.44	14.19	0.67
	H	10639	60.46	10.38	0.81	1764	50.73	14.40	0.68
	I	10507	60.34	10.33	0.80	1731	51.09	14.34	0.68
	J	10511	59.91	10.10	0.80	1621	50.38	13.94	0.67
Mathematics (Common Items only)	Total	108792	48.77	8.58	0.80	17785	40.92	11.86	0.67

**"Disabled" refers to students with any of the following disabilities: autism, deaf-blindness, deafness, emotional disturbance, hearing impairment, mental retardation, multiple disabilities, orthopedic impairment, other health impairment, specific learning disability, speech or language impairment, traumatic brain injury, visual impairment including blindness.

Table 43
Scale Score Descriptive Statistics by Disability Status

Content	Test Form	Not Disabled					Disabled				
		N	Mean	SD	MIN	MAX	N	Mean	SD	MIN	MAX
Reading	A	21854	1353.60	212.30	580	1999	3889	1141.32	237.10	636	1999
	B	22008	1353.94	202.37	652	1999	3517	1139.72	222.28	597	1999
	C	21675	1364.87	221.86	686	1999	3551	1146.23	248.84	642	1999
	D	21582	1357.66	209.38	666	1999	3396	1142.11	238.61	565	1999
	E	21337	1355.95	205.50	641	1999	3352	1141.81	227.32	508	1999
	Total	108456	1357.19	210.42	580	1999	17705	1142.23	235.11	508	1999
Mathematics	A	11098	1380.74	209.58	638	1999	2209	1217.88	230.70	608	1999
	B	11249	1391.46	210.09	608	1999	1867	1214.95	228.00	574	1999
	C	11113	1389.28	205.39	716	1999	1817	1231.60	231.29	638	1999
	D	11152	1389.40	210.85	608	1999	1675	1223.50	229.07	440	1999
	E	10888	1391.26	208.62	574	1999	1745	1221.01	231.07	638	1999
	F	10810	1387.01	213.53	666	1999	1691	1229.97	234.71	535	1999
	G	10824	1388.26	207.62	638	1999	1663	1233.68	230.62	608	1999
	H	10639	1387.50	209.23	608	1999	1766	1218.62	232.46	535	1945
	I	10508	1391.54	210.00	666	1999	1731	1228.53	230.42	638	1945
	J	10511	1387.56	209.55	638	1999	1621	1220.44	228.47	608	1999
Total		108792	1388.40	209.46	574	1999	17785	1223.78	230.71	440	1999

***"Disabled" refers to students with any of the following disabilities: autism, deaf-blindness, deafness, emotional disturbance, hearing impairment, mental retardation, multiple disabilities, orthopedic impairment, other health impairment, specific learning disability, speech or language impairment, traumatic brain injury, visual impairment including blindness.

Table 44

Raw Score Descriptive Statistics by Economic Disability Status

Content	Form	Not Disabled				Economically Disabled			
		N	Mean	SD	Mean P-Value	N	Mean	SD	Mean P-Value
Reading	A	16033	35.16	7.53	0.76	9710	27.98	9.61	0.61
	B	15879	34.57	7.17	0.75	9646	28.38	9.21	0.62
	C	15698	36.39	7.63	0.79	9528	29.27	9.94	0.64
	D	15680	35.93	7.01	0.78	9298	29.02	9.42	0.63
	E	15586	34.98	7.10	0.76	9103	28.34	9.22	0.62
Mathematics (Both Common items and Matrix items)	A	8208	60.98	9.57	0.81	5099	52.35	12.94	0.70
	B	8117	61.61	9.27	0.82	4999	53.09	12.69	0.71
	C	8028	62.72	8.94	0.84	4900	54.12	12.46	0.72
	D	8052	61.98	8.97	0.83	4773	53.21	12.66	0.71
	E	7895	61.91	9.35	0.83	4734	53.45	12.69	0.71
	F	7837	61.97	9.28	0.83	4662	52.60	13.03	0.70
	G	7794	61.34	9.33	0.82	4691	53.22	12.62	0.71
	H	7702	62.36	9.21	0.83	4701	53.71	12.91	0.72
	I	7663	62.16	9.36	0.83	4575	53.81	12.66	0.72
	J	7718	61.72	8.87	0.82	4414	53.25	12.66	0.71
Mathematics (Common Items only)	Total	79022	50.32	7.65	0.82	47555	43.26	10.59	0.71

Table 45
Scale Score Descriptive Statistics by Economic Disability Status

Content	Test Form	Not Disabled					Economically Disabled				
		N	Mean	SD	MIN	MAX	N	Mean	SD	MIN	MAX
Reading	A	16033	1389.90	207.17	636	1999	9710	1208.65	218.97	580	1999
	B	15879	1383.86	199.54	597	1999	9646	1226.59	212.01	597	1999
	C	15698	1403.79	218.25	642	1999	9528	1219.25	225.15	642	1999
	D	15680	1395.69	202.01	666	1999	9298	1214.80	218.76	565	1999
	E	15586	1389.86	201.60	584	1999	9103	1219.04	211.17	508	1999
	Total	78876	1392.59	205.93	584	1999	47285	1217.66	217.38	508	1999
Mathematics	A	8208	1415.84	203.46	666	1999	5099	1253.69	213.04	608	1999
	B	8117	1426.84	202.93	574	1999	4999	1268.10	215.07	608	1999
	C	8029	1428.56	195.44	666	1999	4901	1266.47	210.98	638	1999
	D	8053	1430.00	200.29	608	1999	4774	1262.72	213.12	440	1999
	E	7896	1428.00	201.97	666	1999	4737	1267.29	211.57	574	1999
	F	7839	1429.88	203.44	535	1999	4662	1257.95	212.86	608	1999
	G	7795	1423.34	200.17	608	1999	4692	1275.20	212.99	638	1999
	H	7703	1424.66	200.82	692	1999	4702	1263.18	215.11	535	1999
	I	7664	1427.20	203.35	666	1999	4575	1270.13	212.83	638	1999
	J	7718	1423.70	198.96	638	1999	4414	1262.99	216.72	608	1999
Total		79022	1425.78	201.13	535	1999	47555	1264.71	213.47	440	1999