



2014 Operational Technical Summary

Submitted to the
Louisiana Department of Education

Louisiana Believes.

September 2014





Louisiana Educational Assessment Program (LEAP)

2014 Technical Summary

The tests used in Louisiana are carefully constructed to fairly assess the progress of Louisiana students. The development process and statistical, or psychometric, work are carried out meticulously. This document provides an overview of the process and summarizes some of the key psychometric information.

Introduction

In May 1997, the State Board of Elementary and Secondary Education (BESE) approved rigorous new content standards for K–12. BESE also approved a new criterion-referenced testing program, called the Louisiana Educational Assessment Program for the 21st Century (LEAP 21) to align with these standards. Students were tested in English language arts (ELA), mathematics, science, and social studies at grades 4 and 8. Beginning in 2001, ELA and mathematics were tested at grade 10, and science and social studies were tested at grade 11 beginning in 2002. The high school level assessment program was called the Graduation Exit Examination for the 21st Century (GEE 21). As of 2006, the names of the assessment programs were changed to LEAP and GEE. In 2012, the assessment requirement for graduation changed for students entering ninth grade. The GEE requirement was replaced by end-of-course exam requirements.

In July 2010, the Board of Elementary and Secondary Education adopted the Common Core State Standards (CCSS) in English Language Arts/Literacy and Mathematics. To ease the transition to assessments based on the CCSS, in spring 2013, transitional assessments for ELA and mathematics were administered. The LEAP English Language Arts transitional tests administered in 2012–2013 were revised in two ways. Although they continued to measure the skills presented in the grade clusters, the items were aligned only to those benchmarks related to the content of the CCSS. Secondly, the writing prompt was replaced by a new type of prompt that asks students to read one or two passages and then write a composition that includes evidence from the text(s) to support the writer’s ideas. The LEAP Mathematics transitional tests administered in 2012–2013 changed to include only items that measured content common to the grade 4 or grade 8 GLEs and the CCSS. The content standards and benchmarks that formed the basis for the LEAP Science and Social Studies tests administered in 2012–2013 did not change. In 2013–2014, the English Language Arts and mathematics tests measured only the content of the CCSS.

While the LEAP English Language Arts tests continued to use the text-based writing prompt, the following changes were made:

- The Using Information Resources session, which focused on skimming and scanning skills to locate selected information, was replaced by the Research to Build Knowledge session, which requires students to demonstrate a thorough understanding of the sources provided.
- A new extended-response item was added to the Reading and Responding session of the grade 4 LEAP.
- The Proofreading session of the LEAP test was replaced by the Language session, which used a similar format as the previous LEAP session but focused on more grade-specific skills.

In 2013–2014, the Mathematics test for LEAP was adjusted so that it is now aligned to the CCSS and reporting is based on CCSS domains or combined domains.

In 2013–2014, the Science tests continued to assess Louisiana’s science content standards. The designs of the multiple-choice and constructed-response sessions of the LEAP tests remained the same as they were in 2012–2013. In the 2013–2014 tests, the LEAP Science tests added a new task aligned to the CCSS for English Language Arts or Literacy in Science and Technical Subjects and Louisiana’s content standards. This task was field tested during spring 2013 operational testing.

Like the Science tests, in 2013–2014 the Social Studies tests added a new task, aligned to the CCSS for English Language Arts or Literacy in History/Social Studies and Louisiana’s content standards. The designs of the multiple-choice sessions of the tests remained the same as they were in 2012–2013 and previous administrations. The new task replaced two constructed-response items.

This technical summary provides item- and form-level results from the 2014 spring operational administration. The configuration of this administration was ELA (including writing), mathematics, science, and social studies at grades 4 and 8. The test forms were administered to a census of the school district students in spring 2014 in two phases. Phase 1 test administration occurred in March. Students were given test materials containing ELA writing prompts and constructed-response (CR) items for mathematics in grades 4 and 8. Phase 2 test window occurred in April. The remaining test items were administered for grades 4 and 8 ELA and mathematics. The entire forms for grades 4 and 8 science and social studies were administered during phase 2.

Development Process

Development and administration of a valid and reliable test form proceeds through a series of steps that include the following activities: (1) alignment analysis and test blueprints development, (2) item development, (3) committee reviews of items, (4) revision of items, (5) field testing, (6) field-test data analysis, (7) operational form construction, (8) operational administration, (9) operational test data analysis, and (10) distribution of reports to schools. This summary primarily concerns activities numbered seven through nine, as performed by DRC in support of the LEAP 2014 operational test administration.

At grades 4 and 8, operational ELA and mathematics tests were first administered in 1999 and science and social studies were initially administered in 2000. For the initial administration of a LEAP content area, two forms were used to provide an item pool from which to draw selected items for future administrations. Each year, a form has some items that were administered in these previous operational administrations. These overlapping items are referred to as anchor items, and are used to equate the forms from one year to the next. In the first year of administration for a content area, Forms 1 and 2 also contain some overlapping items, which are also used for equating purposes. For example, for grade 4 ELA and mathematics, Form 6 was administered in 2003 and contained some items from Forms 1, 2, and 3, which were administered in 1999 and 2000. It is the anchor items in the equating process that ensure the scaled score ranges for each achievement level remain consistent from year-to-year.

For spring 2014, eight forms were administered operationally: one form each at grades 4 and 8 for ELA, mathematics, science, and social studies. The forms were constructed from items that had been previously used in an operational administration and field test items. All forms met the content specifications, blueprints and guiding documents, as well as meet psychometric standards for excellence. To construct forms that conformed to content specifications and also meet psychometric standards, the following criteria were followed for item selection.

General Selection Criteria for Anchor and Non-Anchor Items

The following criteria guided selection of items for inclusion in the operational test forms:

- Selected items must match the test blueprint.
- Items should measure the standards and benchmarks specified in the *LEAP Assessment Guides*.
- Items should not have any flaws or distractors that are possible answers.
- Items should cover a variety of questions; similar items on each test form should be minimized.
- Items should cover a range of difficulty levels, but not be extremely easy or hard.
- Items should be selected from the entire item pool. Anchor items must be chosen from operational test items only.
- Anchor items should not be edited. If an anchor item requires editing, it must be removed from the anchor set.
- Constructed-response items in the same form should represent a range of difficulty and a variety of content standards.
- Items should meet the statistical requirements; for example, p -value is between 0.2 and 0.9, and item-total correlation coefficient is greater than 0.2.
- The test characteristic curve (TCC) for the complete test form should closely match the curve from the previous test forms.
- Items with misfit should be avoided.
- Items previously released to the public must not be included in the test form.
- Items should be placed in a similar location, as they were in the field test forms, except for anchor items, which should match the previous item positions.
- Items should be reviewed by content and technical staff members.
- All items and any changes/edits must be reviewed and approved by the head of test development and the lead psychometrician for forms development.

Reliability

Reliability describes the accuracy of the test scores. The more reliable the test, the less measurement error is associated with that test score. Table 1 reports form-level statistics for each content area in grades 4 and 8 for the spring 2014 administration. Results for the test means and standard deviations are based on number-correct (NC) data. NC refers to the total raw score obtained by each student and is used in the calculation of classical test statistics. Of particular importance are the form reliability coefficients (stratified and Cronbach) that are reported in the last two columns of the table. The stratified reliability estimates (Cronbach, Schonemann, & McKic, 1965), define reliability as the sum of the alpha reliability of the component parts (multiple choice items' reliability plus the constructed response items' reliability). This approach provides a better estimate of the overall test reliability (Qualls, 1995).

The standard error of measurement, SEM, calculated from the traditional alpha reliability coefficient and the total score standard deviation are also included in this table. The SEM is useful in making inferences about an individual student's true score. The true score is an individual student's score that is hypothetically attained when there is no measurement error present in a test. It is expected that 68% of the time a student's true score would fall within one SEM around that student's observed score. Correspondingly, the true score is expected to fall within two SEMs of the observed score 95% of the time.

Table 1: Number Correct Test-Level Summary Statistics

Grade and Content	Form	Number of Items	Total Score Points	Mean p -Val	NC Mean	NC Standard Deviation	NC SEM	Reliability	
								Stratified	Cronbach
4 ELA	14T	46	65	0.60	37.09	10.53	3.80	0.90	0.90
4 MA	16T	63	72	0.63	44.00	13.26	3.75	0.92	0.92
4 SC	9S	49	56	0.67	35.83	8.02	3.10	0.85	0.85
4 SS	7R	57	66	0.63	39.85	10.13	3.36	0.88	0.88
8 ELA	14T	49	69	0.64	43.33	9.81	3.40	0.89	0.88
8 MA	16T	64	76	0.54	38.76	13.67	3.87	0.92	0.92
8 SC	9S	49	56	0.66	35.22	8.91	3.21	0.87	0.86
8 SS	8S	67	76	0.60	43.25	12.48	3.53	0.91	0.90

Validity

Validity is the concept that a test measures what it is intended to measure (e.g., grade 4 mathematical concepts and knowledge on the grade 4 test). Validity is thus extended to using test scores for decisions that are supported by what the test purports to measure. Validity is a property of the use of the test information and inferences made using that information, not the test itself.

Content validity, the basis for all LEAP tests, is the basis by which each test is validated. It determines whether a set of items reflects adequate content coverage for each grade, content area, and domain tested.

Content validity begins with the definition of the content domain. For all LEAP tests, the content domain was initially defined by in-state committees. These committees developed the content standards for each subject and grade. The committees were composed of Louisiana educators,

LDOE curriculum and assessment staff, and an outside consultant. The resulting standards were circulated widely across the state for public comment and revised by the committee as necessary.

Content frameworks were then developed and a test blueprint was constructed (see the full technical report for more detail). Thus, the test design was aligned to the content standards established earlier. The content validity was verified by content review committees as well as the LDOE staff and the test contractor. The content review committees were comprised of Louisiana educators and LDOE staff. Some of these individuals had worked on the development of the content standards established earlier. Science and social studies continue to align to these content standards.

The 2013-2014 ELA and mathematics tests were aligned to the CCSS. The CCSS standards were drafted based on state standards already in existence, feedback from teachers, content experts, and education leaders, and comments from the public during two public comment periods. Committees of Louisiana English and mathematics teachers reviewed the CCSS and recommended to BESE that they be adopted.

Items developed for the LEAP tests are reviewed for content and alignment with the standards for each grade and content area. Item field test administrations were conducted and all items were analyzed and verified as to the functioning of the item and the proper coding to the content standards. Thus, the content validity for LEAP tests is built into the test during the development and decisions about students' knowledge or achievement in the various content domains should be valid.

Measurement Model

The scaling method used by DRC for the LEAP assessment is the same Item Response Theory (IRT) model as that is used by the National Assessment of Educational Progress (NAEP). Because the characteristics of multiple-choice (MC) and constructed-response (CR) items are different, two item response theory models are used in the analysis of test forms containing both item types. The two-parameter generalized partial credit model (Muraki, 1992) was applied to the CR items and the three-parameter logistic model (Lord & Novick, 1968; Lord, 1980) to the MC items. Both models estimate the item's discrimination and the scale location, which corresponds to the difficulty of the item. The additional parameter for the MC items estimates an intercept parameter, which is interpreted as the likelihood of a low-achieving student answering the item correctly.

These two IRT models (3PL and GPC) are implemented using Scientific Software's PARSCALE software (Muraki and Bock, 2003). PARSCALE estimates parameters simultaneously for dichotomous and polytomous items using marginal maximum likelihood procedures implemented via the EM algorithm (Bock & Aitkin, 1981; Thissen, 1982).

Scaled-Scores

The scaled scores and cut points for LEAP were set in 1999 when Louisiana first created the LEAP assessments. The LEAP tests from the first (baseline) administration are scaled with a mean of approximately 300 and a standard deviation of approximately 50. The lowest obtainable scaled-score (LOSS) is 100, and the highest obtainable scaled-score (HOSS) is 500 for all LEAP test forms. To ensure rigorous achievement levels, Louisiana set these cut scores using the NAEP as guidance. Scaled scores are derived from raw scores (number correct) and maintain their meaning year-to-year.

Raw Scores

When a student takes a LEAP test, a score produced based on his or her responses to the questions (items) on the test is a raw score. The raw score is usually represented by the number-correct score, i.e., the total number of points earned in the test, percent correct, or proportion correct. For example, for LEAP grade 8 mathematics tests, a raw score is calculated as the sum of the number correct earned from all multiple-choice items and the scores earned from the constructed-response questions. On the raw score scale, the minimum score is 0 and the maximum score is 76 for a test with 60 multiple-choice items and four CR items on a 0-4 scale.

Raw scores may be intuitive and easy to calculate. However, they suffer some serious limitations as reporting scores. Because raw scores depend on the items in a particular form of a test, they can be interpreted only in terms of a particular set of test questions. For example, Matt and Johnny both received a raw score of 56 on the grade 8 LEAP mathematics test. However, Matt took the test in 2013 and Johnny was assessed in 2014. The forms administered in these two years consisted of different test questions. Therefore, one cannot say Matt and Johnny have the same proficiency level in math. As such, the form-dependent nature of raw scores makes raw score reporting problematic when multiple forms are used across years/administrations in any large-scale assessment (e.g., ACT, NAEP), which is the case with the LEAP assessments.

In addition, raw scores are not equal-interval scores. The distance between any raw score point and the next raw score point may not be the same. A raw score distance between scores of 12 and 13 may be quite different from the distance between scores of 44 and 45, due to different difficulties of questions. For example, Matt, Johnny, Mary, and Susan received raw scores of 18, 19, 35, and 36, respectively, on the grade 8 LEAP mathematics test in 2014. Due to the non-equal-interval property of raw scores, we cannot say that the difference in mathematic proficiency between Matt and Johnny is the same as that between Mary and Susan. Thus, we cannot compare students' proficiency levels at different points of the raw score scale.

Given the reasons stated above, here arises the need to convert raw scores to scores that can overcome such limitations, and thus allow for valid comparisons over different test forms and over the entire range of the ability scale. This is why raw scores are converted to scaled-scores and scaled-scores are used for reporting purposes.

Converting Raw Scores to Scaled-Scores

Converting raw scores to scaled-scores is a process of mathematical transformation. Two components are required in the process. First, a mathematical function or a measurement model is needed to map raw scores onto a scale that is independent of the actual test questions administered and that is equal-interval in nature. A well-established procedure for this purpose throughout the entire testing industry is to use Item Response Theory (IRT) models and map raw scores onto the IRT scale, called theta scale. Second, the theta scale has to be transformed to a more user-friendly scale. The theta scale ranges from $-\infty$ to $+\infty$ in theory, and -4 to +4 in practice. While this scale is convenient for mathematical and statistical reasons, it includes fractions and negative numbers so it is not convenient for reporting results. Thus, the theta scale is converted to a reporting scale that is easier to understand, and easier to report on (such as the 100-500 scale used for LEAP). For example, on the 2014 grade 8 mathematics test, the corresponding scaled score for raw scores 13, 14, 37, and 38 are 209, 236, 335, and 337 respectively. The difference between raw scores 13 and 14 was 27 scaled points and the difference between raw scores 37 and 38 was two scaled scores. Obviously the difference between raw scores 12 and 13 is greater than between 37 and 38.

Equating of Test Forms

A statistical process called equating is needed to convert the scale of the form administered in the current administration to the scale of the forms in previous administrations. This is to ensure that scores from different administrations have the same meaning. Detailed technical information describing this process can be found in the full technical report.

As a common practice in large-scaled testing programs, the LEAP tests were constructed to have different item sets appear in test forms across years. All forms for a given grade and content area should provide comparable scores and the passing standards across different administrations should be equivalent. Students are not given an unfair advantage or disadvantage because the particular test form they took was easier or harder than a test form taken by other students. Therefore, a form equating procedure is conducted every year to establish score equivalency between across-year forms.

The differences in test difficulty are adjusted for by a set of common items across years, the anchor items. For example, a higher number correct score on an easier form is equivalent to a lower number correct score on a more difficult form. This process places the form scores on the same scale such that students performing on an assessment at the same level of (underlying) achievement should receive the same scaled score. On the other hand, students receiving the same number-correct score on two different forms might not have the same level of achievement. For example, a number-correct score of 57 on the 2013 grade 8 mathematics test corresponds to a scaled score of 364 and Basic achievement level, while in 2014 the corresponding scaled score was 378 and the achievement level of Mastery. Receiving a raw score of 57 on the 2014 test is associated with a higher score and achievement level than the 2013 test.

The raw-to-scaled score tables utilized for the 2014 operational administration are presented in Table 2.

Table 2: Raw Score to Scaled Score Conversion with Standard Error of Measurement

No.	Grade 4								Grade 8							
	ELA		Math		Science		Social Studies		ELA		Math		Science		Social Studies	
	Form 14T		Form 16T		Form 9S		Form 7R		Form 14T		Form 16T		Form 9S		Form 8S	
Corr. Score	SS*	SEM	SS*	SEM	SS	SEM	SS	SEM	SS*	SEM	SS*	SEM	SS	SEM	SS	SEM
0	100	20	100	20	100	20	100	20	100	20	100	20	100	20	100	20
1	100	20	100	20	100	20	100	20	100	20	100	20	100	20	100	20
2	100	20	100	20	100	20	100	20	100	20	100	20	100	20	100	20
3	100	20	100	20	100	20	100	20	100	20	100	20	100	20	100	20
4	100	20	100	20	100	20	100	20	100	20	100	20	100	20	100	20
5	100	20	100	20	100	20	100	20	100	20	100	20	100	20	100	20
6	100	20	100	20	100	20	100	20	100	20	100	20	100	20	100	20
7	100	20	100	20	100	20	100	20	100	20	100	20	100	20	100	20
8	106	13	100	20	100	20	100	20	116	11	100	20	100	20	100	20
9	148	15	100	20	100	20	100	20	138	13	100	20	133	17	100	20
10	166	16	100	20	100	21	100	20	149	14	100	20	171	18	100	20
11	179	17	100	20	136	22	100	16	159	14	100	20	192	18	100	20
12	190	17	100	20	162	23	147	17	169	14	100	20	205	19	100	20
13	199	18	100	20	181	24	175	18	179	13	209	13	216	19	100	20
14	207	18	149	21	195	24	194	18	188	14	236	14	224	20	118	16
15	215	18	180	22	207	25	207	19	195	14	251	14	231	20	163	16
16	224	18	200	22	217	25	218	19	201	14	261	15	237	21	186	17
17	232	18	215	23	226	25	227	19	207	15	269	15	243	21	201	17
18	240	17	226	23	234	26	234	20	212	14	276	15	248	21	213	18
19	247	17	236	23	241	26	241	20	219	14	281	15	253	20	222	18
20	254	17	244	23	247	26	246	20	226	14	286	15	257	19	230	18

No.	Grade 4								Grade 8							
	ELA		Math		Science		Social Studies		ELA		Math		Science		Social Studies	
	Form 14T		Form 16T		Form 9S		Form 7R		Form 14T		Form 16T		Form 9S		Form 8S	
Corr. Score	SS*	SEM	SS*	SEM	SS	SEM	SS	SEM	SS*	SEM	SS*	SEM	SS	SEM	SS	SEM
21	261	17	252	24	253	24	252	20	233	14	291	16	261	18	236	18
22	267	17	259	24	259	23	256	20	241	14	295	16	265	17	242	19
23	273	18	265	22	264	21	261	19	248	14	299	16	269	17	247	19
24	278	18	271	21	269	20	265	18	254	14	302	15	273	16	252	19
25	283	17	276	20	274	19	269	17	260	14	305	14	277	15	256	19
26	288	17	281	19	278	19	273	16	265	14	308	13	280	15	261	19
27	292	16	286	18	283	18	277	16	270	15	311	12	284	15	264	18
28	296	16	291	17	287	18	280	15	274	14	314	12	287	14	268	17
29	301	16	295	17	291	17	283	15	278	14	317	11	291	14	272	16
30	304	16	300	16	295	17	287	14	282	13	319	11	294	14	275	15
31	308	16	304	16	299	16	290	14	286	13	322	11	298	14	278	15
32	312	15	308	15	303	16	293	14	289	13	324	10	301	14	281	14
33	315	15	312	15	308	16	296	13	292	13	326	10	305	13	284	14
34	319	15	316	15	312	16	299	13	295	12	328	10	308	13	287	13
35	323	15	319	14	316	16	302	13	299	12	331	10	312	13	290	13
36	326	15	323	14	320	16	305	13	302	12	333	9	316	13	293	12
37	330	15	327	14	325	16	308	12	305	12	335	9	319	13	296	12
38	333	15	330	14	329	16	311	12	308	12	337	9	323	13	299	12
39	337	15	334	14	334	16	314	12	311	12	339	9	327	13	301	12
40	340	15	337	13	339	16	317	12	314	11	341	9	331	14	304	11
41	344	15	341	13	344	16	321	12	317	11	343	9	335	14	307	11
42	348	15	344	13	349	16	324	12	321	11	345	9	340	14	310	11
43	352	15	347	13	355	16	327	12	324	11	346	8	344	14	312	11

No.	Grade 4								Grade 8							
	ELA		Math		Science		Social Studies		ELA		Math		Science		Social Studies	
	Form 14T		Form 16T		Form 9S		Form 7R		Form 14T		Form 16T		Form 9S		Form 8S	
Corr. Score	SS*	SEM	SS*	SEM	SS	SEM	SS	SEM	SS*	SEM	SS*	SEM	SS	SEM	SS	SEM
44	356	15	351	13	361	17	330	12	327	11	348	8	350	14	315	11
45	360	15	354	13	368	17	333	12	331	11	350	8	355	15	318	11
46	364	15	357	13	376	18	337	12	334	11	352	8	362	15	320	11
47	369	15	361	13	384	18	340	12	338	11	354	8	369	15	323	11
48	374	15	364	13	395	18	343	12	341	11	356	8	377	14	326	11
49	380	14	368	13	407	17	347	12	345	11	358	8	387	13	328	10
50	386	14	371	13	424	15	351	12	348	11	360	8	399	12	331	10
51	392	13	375	13	449	14	355	12	352	11	363	8	415	11	334	10
52	399	13	379	13	493	13	360	12	357	12	365	8	435	10	337	10
53	407	12	383	13	500	12	365	12	361	12	367	8	464	8	339	10
54	417	11	387	14	500	12	370	12	366	12	370	8	500	6	342	11
55	428	11	391	14	500	12	376	12	371	11	372	8	500	6	345	11
56	441	10	395	14	500	12	384	12	377	11	375	8	500	6	348	11
57	457	9	400	14			393	11	384	10	378	8			352	11
58	479	8	405	14			404	10	392	10	381	8			355	11
59	500	7	411	15			420	9	402	9	384	8			358	11
60	500	5	417	15			444	8	413	9	388	8			362	11
61	500	5	425	15			483	7	426	8	392	8			366	11
62	500	5	433	15			500	6	441	8	397	8			370	11
63	500	5	443	14			500	6	458	7	402	8			375	11
64	500	5	456	14			500	6	477	7	409	8			380	12
65	500	5	473	12			500	6	499	6	418	7			386	12
66			496	11			500	6	500	5	430	7			393	11

No.	Grade 4								Grade 8							
	ELA		Math		Science		Social Studies		ELA		Math		Science		Social Studies	
	Form 14T		Form 16T		Form 9S		Form 7R		Form 14T		Form 16T		Form 9S		Form 8S	
Corr. Score	SS*	SEM	SS*	SEM	SS	SEM	SS	SEM	SS*	SEM	SS*	SEM	SS	SEM	SS	SEM
67			500	10					500	5	445	6			402	10
68			500	8					500	5	468	6			413	9
69			500	8					500	5	498	5			428	9
70			500	8							500	5			449	8
71			500	8							500	5			483	6
72			500	8							500	5			500	5
73											500	5			500	5
74											500	5			500	5
75											500	5			500	5
76											500	5			500	5

*Half-score point increments not shown were defined for actual scoring. Scaled-scores were interpolated.

Table 3 presents the mean scaled-score data based on all test scores valid for reporting purposes from grades 4 and 8 of the spring 2014 LEAP operational administration.

Table 3: Scaled-Score Means and Standard Deviations

Scaled Score				
Grade	Subject	N	Mean	Standard Deviation
4	ELA	54,715	331.18	54.67
	Mathematics	54,794	355.14	67.56
	Science	54,749	323.46	46.59
	Social Studies	54,705	317.97	41.64
8	ELA	51,965	326.82	45.02
	Mathematics	51,930	333.06	52.12
	Science	51,766	316.50	44.12
	Social Studies	51,750	311.02	47.20

Table 4 provides data to compare score distributions from the spring 2014 forms to all previous administrations. These statistics are based on population data. Although these represent rounded values, there are differences in the scaled-score values for a given percentile across all forms. These variations could arise for several reasons: (1) differences in proficiency (i.e., achievement) of the samples or growth in student achievement from the base year to 2014; (2) unevenness in the respective distributions that combine with the number-correct-to-scaled (rounded)-score scoring method, leaving “gaps” in the scale; and (3) other sources of equating error. Other sources of equating error can include subtle content differences between forms, hand-scoring differences, or unusual student samples. Some equating error will always be present between forms. This means that the forms will not always measure identically, even under optimal testing conditions. In general, however, the test characteristic function equating technique is used to level the equated forms through the raw-to-scaled-score adjustment.

**Table 4: Comparisons of Scaled Scores at Selected Percentiles
Grade 4 English Language Arts**

Percentile	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
	Form 4	Form 5	Form 6	Form 7	Form 8	Form 9	Form 10	Form 11	Form 10	Form 9	Form 12	Form 13	Form 12T	Form 14T
99	435	402	404	430	425	449	447	446	447	458	477	469	483	457
95	394	378	378	394	391	397	400	407	400	408	411	411	424	412
90	372	363	361	378	374	378	381	385	381	381	390	392	395	392
85	358	354	352	367	359	362	368	372	368	368	374	377	382	383
80	353	346	343	356	352	355	360	363	360	360	367	368	371	372
75	342	341	339	349	344	346	350	355	353	353	358	360	362	364
70	332	335	331	342	339	340	346	350	346	344	350	353	356	358
65	328	329	326	335	332	334	339	343	341	340	345	346	349	352
60	319	324	320	328	325	328	333	338	335	332	340	342	341	346
55	311	319	317	321	320	323	327	331	329	328	334	336	337	340
50	307	313	310	314	314	318	322	326	325	321	327	330	333	333
45	299	307	306	305	309	310	316	320	320	318	323	326	327	328
40	293	302	299	299	303	305	311	314	314	310	317	319	321	323
35	289	295	294	292	296	300	306	310	309	305	311	314	315	315
30	283	289	288	282	290	292	300	304	304	300	305	308	310	308
25	275	281	281	275	282	284	293	296	297	292	300	301	303	301
20	265	271	273	266	274	273	286	287	289	282	292	294	296	292
15	257	261	263	250	263	259	274	276	280	270	283	282	288	281
10	224	246	252	232	245	242	257	260	264	253	268	270	275	264
5	184	225	230	202	221	202	221	230	233	219	245	242	252	240
1	100	172	175	137	145	119	154	163	162	138	173	177	194	179

Grade 4 Mathematics

Percentile	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
	Form 4	Form 5	Form 6	Form 7	Form 8	Form 9	Form 10	Form 11	Form 10	Form 11	Form 12	Form 13	Form 14T	Form 16T
99	422	437	441	431	446	492	449	471	485	500	498	500	500	500
95	392	399	398	403	400	418	401	417	412	422	434	448	442	496
90	375	376	380	382	380	396	385	398	385	401	408	412	419	443
85	364	360	370	370	370	382	374	384	374	386	391	397	407	417
80	355	352	360	361	361	370	364	377	367	379	383	386	394	405
75	348	345	353	354	353	360	356	368	360	370	375	376	386	393
70	342	338	346	348	349	355	351	363	353	364	366	370	377	385
65	337	332	340	340	342	346	346	356	346	357	358	364	369	375
60	331	326	334	332	336	340	340	350	342	350	353	357	361	368
55	325	320	328	325	331	334	335	344	336	346	348	350	354	361
50	319	315	324	320	326	327	330	339	330	340	341	345	347	354
45	314	309	318	313	320	320	324	334	324	334	334	339	340	347
40	307	301	312	307	313	315	319	327	319	327	328	333	333	339
35	302	295	307	298	308	308	313	321	313	321	322	326	326	330
30	296	290	300	292	302	301	307	314	307	314	315	320	317	323
25	289	280	292	285	295	295	299	307	300	307	309	312	307	316
20	280	274	283	275	285	284	290	296	292	298	301	304	297	304
15	270	264	275	265	273	273	280	285	282	287	290	295	287	293
10	258	248	261	250	257	259	268	270	268	272	277	280	267	276
5	236	229	234	226	228	235	243	246	246	252	254	262	242	252
1	183	155	157	170	163	170	189	186	194	202	207	213	168	165

Grade 4 Science

Percentile	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
	Form 4	Form 5	Form 6	Form 4R	Form 7	Form 8	Form 6	Form 9	Form 7	Form 8	Form 9R	Form 10	Form 9S
99	430	413	414	445	429	430	428	473	429	430	481	442	449
95	395	374	386	399	389	389	386	411	399	389	430	400	395
90	375	360	373	378	372	374	373	389	380	374	401	382	376
85	361	349	361	363	359	362	361	373	365	362	380	368	368
80	355	343	351	356	353	356	356	366	359	356	372	362	355
75	349	333	347	354	348	347	347	359	353	347	364	357	349
70	340	332	338	340	343	342	342	347	343	342	351	347	344
65	335	319	333	335	333	338	338	341	338	338	345	342	339
60	330	314	329	330	328	334	329	336	333	329	339	338	334
55	325	310	321	325	323	325	325	331	328	325	334	329	329
50	321	300	317	317	314	321	321	326	323	321	328	325	325
45	316	296	313	313	310	317	317	316	314	317	323	321	320
40	312	291	305	308	305	309	309	312	310	309	314	313	316
35	303	282	297	304	296	305	305	307	305	305	309	309	308
30	298	272	292	300	292	297	297	298	296	301	304	305	303
25	289	267	284	292	283	293	292	293	292	293	295	296	295
20	284	263	275	283	273	284	284	284	283	284	286	287	291
15	274	250	264	274	263	275	275	274	273	275	277	277	283
10	263	230	247	265	253	258	259	258	258	264	268	267	269
5	243	204	225	241	226	237	233	235	234	237	247	241	247
1	178	123	165	152	151	161	165	185	151	161	198	172	195

Grade 4 Social Studies

Percentile	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
	Form 4	Form 5	Form 6	Form 4	Form 7	Form 8	Form 6	Form 9	Form 7	Form 8	Form 9	Form 10	Form 7R
99	402	400	403	402	399	412	403	419	414	428	419	417	420
95	378	370	377	363	373	381	377	377	380	389	384	381	376
90	358	359	362	349	362	362	362	365	362	367	365	363	365
85	349	350	354	342	353	356	354	355	353	362	360	353	355
80	342	343	346	336	345	347	346	347	349	352	351	349	351
75	336	336	339	330	341	343	339	343	341	347	343	342	343
70	330	329	332	324	334	335	336	335	338	339	339	338	340
65	324	326	325	318	331	331	329	332	331	335	335	335	333
60	318	320	322	316	324	328	322	325	328	331	328	329	330
55	316	314	315	310	321	321	319	322	324	324	325	326	324
50	313	311	312	307	315	317	312	315	318	321	318	320	321
45	307	306	306	302	311	311	309	312	315	317	315	317	314
40	302	300	299	296	305	308	302	305	308	311	309	311	311
35	299	294	295	293	298	304	299	302	302	304	305	305	305
30	293	288	289	287	295	298	292	295	298	301	299	303	302
25	287	282	282	280	288	291	285	288	292	294	292	297	296
20	273	275	274	273	281	284	278	281	284	288	285	287	287
15	260	265	266	264	269	276	270	269	273	276	277	281	280
10	256	257	254	249	255	263	262	255	260	263	260	269	269
5	242	240	232	216	231	231	244	232	238	239	238	248	252
1	170	200	190	100	168	140	201	144	168	140	163	151	207

Grade 8 English Language Arts

Percentile	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
	Form 4	Form 5	Form 6	Form 7	Form 8	Form 9	Form 10	Form 11	Form 10	Form 12	Form 13	Form 12R	Form 13T	Form 14T
99	394	446	404	393	399	429	409	424	409	427	452	442	443	441
95	370	393	379	371	376	388	381	381	378	393	406	402	404	392
90	359	381	365	354	362	369	366	365	363	373	386	382	384	374
85	350	367	357	349	354	356	358	355	356	364	371	367	370	366
80	343	357	349	340	348	348	349	351	349	355	364	361	362	357
75	337	348	342	336	342	342	345	345	345	348	356	353	357	352
70	331	338	337	330	336	335	340	339	340	343	348	346	350	347
65	325	333	331	326	331	331	335	334	337	337	344	342	344	341
60	321	327	326	320	326	326	332	331	332	333	340	338	340	338
55	316	321	321	316	320	321	328	326	329	328	335	334	337	334
50	310	315	316	313	317	318	325	321	326	325	331	329	332	329
45	305	309	311	308	312	314	320	318	322	319	326	326	327	324
40	300	306	304	302	306	309	316	313	317	316	322	321	322	321
35	296	298	299	298	301	305	311	309	313	310	317	317	317	316
30	290	293	294	292	294	300	307	304	308	307	313	312	314	311
25	283	288	289	286	289	294	301	297	302	302	308	308	308	305
20	276	284	280	280	282	287	295	290	296	296	302	302	302	299
15	268	278	268	270	270	279	286	283	288	289	294	294	294	291
10	256	269	256	257	258	264	273	268	277	276	284	284	282	278
5	235	240	237	234	238	230	243	241	250	257	265	265	258	254
1	200	180	192	174	189	134	166	162	169	175	189	195	180	179

Grade 8 Mathematics

Percentile	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
	Form 4	Form 5	Form 6	Form 7	Form 8	Form 9	Form 10	Form 11	Form 10	Form 12	Form 13	Form 11R	Form 14T	Form 16T
99	409	404	421	411	411	419	491	428	500	500	455	500	449	500
95	382	378	383	380	381	378	398	379	425	422	393	402	393	409
90	367	361	370	369	368	364	379	365	388	401	375	377	376	383
85	357	351	359	359	361	355	366	357	373	386	366	365	366	371
80	349	346	351	353	354	349	357	350	363	379	358	359	359	363
75	342	340	344	348	348	344	350	346	354	370	353	353	353	356
70	337	335	337	343	343	339	345	342	349	364	347	348	348	351
65	331	330	333	338	338	335	340	337	343	357	342	344	345	346
60	325	325	327	334	332	330	335	334	339	350	338	340	341	343
55	320	320	323	330	328	325	330	329	333	346	334	335	337	338
50	316	317	317	326	325	321	326	326	329	340	330	331	333	334
45	310	312	312	322	320	316	322	322	324	334	326	328	329	330
40	304	307	307	317	314	311	318	318	320	327	322	323	325	325
35	299	303	301	313	309	307	314	315	316	321	317	319	321	319
30	293	299	295	307	304	300	308	310	311	314	312	315	317	314
25	286	292	289	301	298	294	302	305	305	307	308	310	312	308
20	279	284	280	295	292	288	298	300	299	298	302	305	305	302
15	270	276	270	287	284	279	291	293	292	287	295	298	299	295
10	261	268	254	276	272	267	282	283	282	272	286	289	290	284
5	240	245	226	255	249	249	267	265	267	252	270	275	271	265
1	186	167	100	201	199	162	221	214	221	202	227	243	218	100

Grade 8 Science

Percentile	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
	Form 4	Form 5	Form 6	Form 4	Form 7	Form 8	Form 6	Form 9	Form 7	Form 8R	Form 9R	Form 10	Form 9S
99	402	393	396	405	407	413	406	453	429	421	463	437	435
95	371	364	367	373	368	375	373	391	399	375	395	393	387
90	353	351	356	356	353	358	356	365	380	361	376	375	369
85	343	339	346	347	341	348	346	353	365	350	362	361	355
80	338	332	337	338	332	344	342	342	359	345	350	349	350
75	329	328	329	334	328	335	333	338	353	336	341	344	340
70	324	321	325	329	321	332	329	329	343	332	336	335	335
65	315	314	321	322	317	328	321	325	338	328	332	331	331
60	311	311	313	318	311	320	317	317	333	320	324	327	327
55	306	303	309	310	308	316	309	314	328	316	320	323	319
50	297	300	305	306	301	312	305	310	323	312	317	316	316
45	293	296	297	302	298	305	301	303	314	305	310	312	312
40	288	288	293	294	292	301	293	299	310	301	306	309	308
35	279	284	285	290	289	293	289	292	305	297	299	302	301
30	274	276	280	286	282	289	280	288	296	289	295	299	298
25	263	267	271	277	275	281	276	280	292	281	288	292	291
20	257	262	262	267	267	272	267	273	283	272	280	285	284
15	245	251	252	262	259	262	257	265	273	263	272	275	277
10	231	232	234	252	245	252	241	256	258	248	264	263	265
5	202	205	213	232	226	226	221	237	234	230	246	243	248
1	100	105	148	171	145	170	182	199	151	175	206	184	205

Grade 8 Social Studies

Percentile	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
	Form 4	Form 5	Form 6	Form 4R	Form 6	Form 7	Form 6	Form 9	Form 7	Form 8R	Form 9	Form 10	Form 8S
99	393	386	401	401	413	411	401	428	414	433	454	451	428
95	359	359	367	368	367	371	372	375	380	381	389	386	380
90	352	345	353	355	352	355	353	359	362	360	369	366	362
85	340	338	342	346	340	345	345	347	353	352	355	352	352
80	332	329	335	337	334	338	335	341	349	342	347	345	345
75	327	323	329	329	327	332	329	332	341	336	341	339	339
70	321	316	321	324	322	326	323	326	338	330	335	333	334
65	316	310	318	318	316	320	318	321	331	324	329	328	328
60	311	304	312	313	311	314	312	316	328	318	324	323	323
55	306	298	307	307	305	312	307	311	324	315	319	318	318
50	300	291	299	301	300	306	302	306	318	310	314	313	312
45	294	285	294	298	294	300	296	301	315	304	311	309	310
40	287	281	291	291	289	297	291	296	308	299	304	304	304
35	280	274	285	284	283	291	285	291	302	293	299	297	299
30	277	267	279	280	277	285	279	282	298	287	293	292	293
25	268	258	273	272	271	279	273	276	292	282	288	284	284
20	259	249	263	262	261	269	266	269	284	272	279	279	278
15	248	239	252	252	250	262	256	258	273	263	269	269	268
10	235	227	238	239	236	250	243	244	260	251	258	252	256
5	208	193	215	215	211	229	221	219	238	233	233	228	236
1	100	120	165	153	133	182	178	160	168	181	177	110	163

The scaled-score ranges for each grade and subjects' achievement levels are summarized in Table 5. These ranges remain the same from year-to-year. The final percentages of students in each achievement level for the 2014 operational administration are also reported in Table 5. Percentages may not add up to 100 due to rounding.

**Table 5: Scaled-Score Ranges and Percentages of Students in Achievement Levels
Grade 4 English Language Arts**

Achievement Level	Scaled-Score Range	Percentage of Students
Advanced	408–500	5.7
Mastery	354–407	28.1
Basic	301–353	42.1
Approaching Basic	263–300	14.6
Unsatisfactory	100–262	9.5

Grade 4 Mathematics

Achievement Level	Scaled-Score Range	Percentage of Students
Advanced	419–500	14.8
Mastery	370–418	24.6
Basic	315–369	35.7
Approaching Basic	282–314	13.2
Unsatisfactory	100–281	11.7

Grade 4 Science

Achievement Level	Scaled-Score Range	Percentage of Students
Advanced	405–500	4.0
Mastery	360–404	15.7
Basic	306–359	48.3
Approaching Basic	263–305	24.5
Unsatisfactory	100–262	7.4

Grade 4 Social Studies

Achievement Level	Scaled-Score Range	Percentage of Students
Advanced	399–500	1.9
Mastery	353–398	15.2
Basic	301–352	52.9
Approaching Basic	272–300	19.2
Unsatisfactory	100–271	10.7

Grade 8 English Language Arts

Achievement Level	Scaled-Score Range	Percentage of Students
Advanced	402–500	3.8
Mastery	356–401	18.1
Basic	315–355	43.6
Approaching Basic	269–314	27.4
Unsatisfactory	100–268	7.2

Grade 8 Mathematics

Achievement Level	Scaled-Score Range	Percentage of Students
Advanced	398–500	6.3
Mastery	376–397	6.3
Basic	321–375	52.3
Approaching Basic	296–320	19.5
Unsatisfactory	100–295	15.6

Grade 8 Science

Achievement Level	Scaled-Score Range	Percentage of Students
Advanced	400–500	2.6
Mastery	345–399	18.2
Basic	305–344	43.1
Approaching Basic	267–304	25.6
Unsatisfactory	100–266	10.6

Grade 8 Social Studies

Achievement Level	Scaled-Score Range	Percentage Students
Advanced	404–500	1.8
Mastery	350–403	15.2
Basic	297–349	48.3
Approaching Basic	263–296	22.4
Unsatisfactory	100–262	12.2

References

Bock, R. D. & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 66, 443-459.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. New York: Erlbaum.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16, 159-176.

Muraki, E., and Bock, R. D. (2003). *PARSCALE 4: IRT item analysis and test scoring for rating-scale data* [computer program]. Chicago, IL: Scientific Software.

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

Thissen, D. (1982). Marginal maximum-likelihood estimation for the one-parameter logistic model. *Psychometrika*, 47, 175-186.