

**NEW YORK STATE
TESTING PROGRAM**

**ENGLISH AS A SECOND LANGUAGE ACHIEVEMENT TEST
(NYSESLAT)**

**2010 ADMINISTRATION
TECHNICAL MANUAL**

SUBMITTED BY

PEARSON

NOVEMBER 2010



TABLE OF CONTENTS

OVERVIEW OF THIS MANUAL.....	1
<i>Part 1 Introduction</i>	1
<i>Part 2 Test Design and Development</i>	1
<i>Part 3 Scoring</i>	1
<i>Part 4 Classical Item-Level Statistics</i>	1
<i>Part 5 Reliability</i>	1
<i>Part 6 Validity</i>	1
<i>Part 7 Calibration, Equating, and Scaling</i>	2
<i>Part 8 Standard Setting</i>	2
<i>Part 9 Summary of Operational Test Results</i>	2
PART 1: INTRODUCTION.....	3
<i>1.1 Background</i>	3
<i>1.2 Rationale and Purpose</i>	4
<i>1.3 Recommended Test Use</i>	4
<i>1.4 Test Accommodations</i>	4
PART 2: TEST DESIGN AND DEVELOPMENT.....	6
<i>2.1 Overview</i>	6
<i>2.2 Test Blueprint and Specifications by Modality and Grade Span</i>	6
<i>2.3 Test Mapping and Specifications by New York State Learning Standards for English as a Second Language</i>	10
<i>2.4 Item Development and Review Processes</i>	13
<i>2.5 Field Test</i>	13
<i>2.6 Test Construction</i>	14
PART 3: SCORING.....	18
<i>3.1 NYSESLAT Range Finding</i>	18
<i>3.2 Operational File</i>	18
<i>3.3 Rater Training</i>	18
<i>3.4 Inter-Rater and Intra-Rater Reliability</i>	19
PART 4: CLASSICAL ITEM-LEVEL STATISTICS.....	20
<i>4.1 Item-Level Descriptive Statistics</i>	20
<i>4.2 Differential Item Functioning</i>	23
PART 5: RELIABILITY.....	27
<i>5.1 Internal Consistency Reliability</i>	27
<i>5.2 Classical SEM (Based on Classical Test Theory)</i>	27
<i>5.3 Standard Error of the Mean (SEM_n)</i>	28
<i>5.4 Conditional SEM (Based on Item Response Theory)</i>	29
<i>5.5 Summary of Descriptive and Reliability Statistics</i>	30
<i>5.6 Inter-Rater Reliability</i>	35
<i>5.7 Reliability of Classification Decision at Proficient Cut</i>	40
PART 6: VALIDITY.....	44
<i>6.1 Content Validity</i>	44
<i>6.2 Internal Structure</i>	45
<i>6.3 External Structure</i>	51
<i>6.3.1 Relationship with the New York English Language Arts Assessment</i>	51
<i>6.3.2 Relationship with the New York State Regents Examination in Comprehensive English</i>	58

PART 7: CALIBRATION, EQUATING, AND SCALING.....	63
7.1 <i>Item Response Model and Rational for Use</i>	63
7.2 <i>Evidence of Model Fit</i>	63
7.3 <i>Description of Calibration Sample, Process, and Results</i>	64
7.4 <i>Equating of the NYSESLAT</i>	65
7.5 <i>Scaling of the NYSESLAT</i>	80
7.6 <i>Rasch Information</i>	81
PART 8: STANDARD SETTING.....	83
8.1 <i>Introduction</i>	83
8.2 <i>Standard Setting Model</i>	83
8.3 <i>Committees of Panelists</i>	83
8.4 <i>Performance Levels and Cut Scores</i>	84
8.5 <i>Standard-Setting Process</i>	84
8.6 <i>Summary Statistics for the Three Rounds of Ratings</i>	86
8.7 <i>Post-Standard-Setting Analyses</i>	86
8.8 <i>Final Performance-Level Cut Points</i>	86
PART 9: SUMMARY OF OPERATIONAL TEST RESULTS.....	89
APPENDIX A: ITEM MAP FOR 2010 NYSESLAT	94
APPENDIX B: ITEM LEVEL STATISTICS BY GRADE SPAN	139
APPENDIX C: RAW SCORE TO SCALE SCORE CONVERSION TABLES	149
APPENDIX D: IRT STATISTICS	165
APPENDIX E: ITEM INFORMATION AT DIFFERENT CUT POINTS BY GRADES	175
APPENDIX F: CONFIRMATORY FACTOR ANALYSIS FOR THE 2010 NYSESLAT	201
APPENDIX G: SCALE SCORE SUMMARY BY SUBGROUPS.....	211
APPENDIX H: PROFICIENCY PERCENTAGES	237
APPENDIX I: EXIT RATE BY SUBGROUP	276
APPENDIX J: DIF CATEGORIES FOR OPERATIONAL ITEMS	289
APPENDIX K: REFERENCES	292

LIST OF TABLES

Table 1 - 2010 NYSESLAT Operational Test Blueprint.....	7
Table 2 - 2010 Test Specifications by Modality and Grade Span.....	9
Table 3 - Maximum Number of Points by Modality and Grade Span.....	9
Table 4 - Test Specifications for New York State ESL Learning Standards by Grade Span.....	10
Table 5 - Item Mapping by New York State ESL Learning Standards and Performance Indicators.....	11
Table 6 - Four-Year Sampling Matrix for the NYSESLAT Field Test.....	14
Table 7 - Summary of Classical Item Difficulty and Item Discrimination Indices for Each Grade Span.....	21
Table 8 - DIF Classification for Constructed-Response Items.....	24
Table 9 - DIF Classification for Multiple-Choice Items.....	25
Table 10 - Results of DIF Analyses for 2010 Operational Test Items.....	26
Table 11 - Descriptive Statistics and Reliability by Grade Span and Modality.....	31
Table 12 - Descriptive Statistics and Reliability by Grade and Modality.....	32
Table 13 - Audit Sample.....	36
Table 14 - Rater Agreement for Pre-Writing and Writing Prompts.....	38
Table 15 - Percentages of the Score Difference Between Raters.....	39
Table 16 - Comparison Between Local and Pearson Raters.....	40
Table 17 - Classification Accuracy and Consistency by Grade.....	43
Table 18 - Intercorrelation Among the Modalities by Grade Span.....	46
Table 19 - Intercorrelation Among the Modalities by Grade.....	46
Table 20 - Global Fit Indices for the One- and Two-Factor Models (Grades K–1).....	50
Table 21 - Global Fit Indices for the One- and Two-Factor Models (Grades 2–4).....	50
Table 22 - Global Fit Indices for the One- and Two-Factor Models (Grades 5–6).....	50
Table 23 - Global Fit Indices for the One- and Two-Factor Models (Grades 7–8).....	50
Table 24 - Global Fit Indices for the One- and Two-Factor Models (Grades 9–12).....	50
Table 25 - Descriptive Statistics of the NYSESLAT Raw Scores of Listening, Speaking, Reading, and Writing Modalities and Their Correlation with the ELA Test Total Scale Scores.....	53
Table 26 - Descriptive Statistics of the New York ELA Test and NYSESLAT Combined Scale Scores of Listening/Speaking and Reading/Writing.....	54
Table 27 - Grade 3: Percentage of ELL Students Scoring at Each NYSESLAT Proficiency Level.....	57
Table 28 - Grade 4: Percentage of ELL Students Scoring at Each NYSESLAT Proficiency Level.....	57
Table 29 - Grade 5: Percentage of ELL Students Scoring at Each NYSESLAT Proficiency Level.....	57
Table 30 - Grade 6: Percentage of ELL Students Scoring at Each NYSESLAT Proficiency Level.....	57
Table 31 - Grade 7: Percentage of ELL Students Scoring at Each NYSESLAT Proficiency Level.....	58
Table 32 - Grade 8: Percentage of ELL Students Scoring at Each NYSESLAT Proficiency Level.....	58
Table 33 - Descriptive Statistics of the NYSESLAT Raw Scores of Listening, Speaking, Reading, and Writing Modalities and Their Correlations With the Regents English Exam Total Scale Scores.....	59
Table 34 - Descriptive Statistics of the NYSESLAT Scale Scores of Listening/Speaking, Reading/Writing Modalities, and Total Scale Scores and Their Correlations With the Regents English Exam Total Scale Scores.....	59
Table 35 - Grade 10: Percentage of ELL Students Passing/Failing the Regents English Exam at Each NYSESLAT Proficiency Level.....	61
Table 36 - Grade 11: Percentage of ELL Students Passing/Failing the Regents English Exam at Each NYSESLAT Proficiency Level.....	61
Table 37 - Grade 12: Percentage of ELL Students Passing/Failing the Regents English Exam at Each NYSESLAT Proficiency Level.....	62
Table 38 - Common Items as Anchor Items by Modality Combination and Grade Span.....	68
Table 39 - Average Rasch Difficulty by Grade Span and Modality.....	82
Table 40 - Panel Composition for Standard-Setting Committees.....	84
Table 41 - Final Performance-Level Cut Points for Listening and Speaking.....	87
Table 42 - Final Performance-Level Cut Points for Reading and Writing.....	88

Table 43 - Raw Score Summary by Grade	90
Table 44 - Raw Score Summary by Grade Span	90
Table 45 - Scale Score Summary by Grade	91
Table 46 - Scale Score Summary by Grade Span	91
Table 47 - Percentage of Students in Each Proficiency Level by Grade	92
Table 48 - Percentage of Students in Each Proficiency Level by Grade Span	93
Table 49 - Exit Rate by Grade for 2007, 2008, 2009, and 2010 Administrations.....	93
Table 50 - Exit Rate by Grade Span for 2007, 2008, 2009, and 2010 Administrations.....	93
Table 51 - Maximum Number of Points by Modality and Grade Span for the 2010 NYSESLAT	201
Table 52 - Summary Statistics	204
Table 53 - Factor Loading Estimates for the One- and Two-Factor Models	205
Table 54 - Global Fit Indices for the One- and Two-Factor Models	205
Table 55 - Summary Statistics	205
Table 56 - Factor Loading Estimates for the One- and Two-Factor Models	206
Table 57 - Global Fit Indices for the One- and Two-Factor Models	206
Table 58 - Summary Statistics	206
Table 59 - Factor Loading Estimates for the One- and Two-Factor Models	207
Table 60 - Global Fit Indices for the One- and Two-Factor Models	207
Table 61 - Summary Statistics	207
Table 62 - Factor Loading Estimates for the One- and Two-Factor Models	208
Table 63 - Global Fit Indices for the One- and Two-Factor Models	208
Table 64 - Summary Statistics	208
Table 65 - Factor Loading Estimates for the One- and Two-Factor Models	209
Table 66 - Global Fit Indices for the One- and Two-Factor Models	209

LIST OF FIGURES

Figure 1 - Classification Accuracy	41
Figure 2 - Classification Consistency	41
Figure 3 - 2010 TCC L/S Grade Span 1-5	69
Figure 4 - 2010 TCC R/W Grade Span 1-5.....	70
Figure 5 - 2008-09-10 TCC L/S Grade Span 1.....	71
Figure 6 - 2008-09-10 TCC L/S Grade Span 2.....	71
Figure 7 - 2008-09-10 TCC L/S Grade Span 3.....	72
Figure 8 - 2008-09-10 TCC L/S Grade Span 4.....	72
Figure 9 - 2008-09-10 TCC L/S Grade Span 5.....	73
Figure 10 - 2008-09-10 SE Curves L/S Grade Span 1.....	73
Figure 11 - 2008-09-10 SE Curves L/S Grade Span 2.....	74
Figure 12 - 2008-09-10 SE Curves L/S Grade Span 3.....	74
Figure 13 - 2008-09-10 SE Curves L/S Grade Span 4.....	75
Figure 14 - 2008-09-10 SE Curves L/S Grade Span 5.....	75
Figure 15 - 2008-09-10 TCC R/W Grade Span 1.....	76
Figure 16 - 2008-09-10 TCC R/W Grade Span 2.....	76
Figure 17 - 2008-09-10 TCC R/W Grade Span 3.....	77
Figure 18 - 2008-09-10 TCC R/W Grade Span 4.....	77
Figure 19 - 2008-09-10 TCC R/W Grade Span 5.....	78
Figure 20 - 2008-09-10 SE Curves R/W Grade Span 1.....	78
Figure 21 - 2008-09-10 SE Curves R/W Grade Span 2.....	79
Figure 22 - 2008-09-10 SE Curves R/W Grade Span 3.....	79
Figure 23 - 2008-09-10 SE Curves R/W Grade Span 4.....	80
Figure 24 - 2008-09-10 SE Curves R/W Grade Span 5.....	80
Figure 25 - Path Diagram for the One-Factor Model.....	203
Figure 26 - Path Diagram for the Two-Factor Model.....	204

OVERVIEW OF THIS MANUAL

This New York State English as a Second Language Achievement Test (NYSESLAT) Technical Manual for the 2010 administration is organized into nine major parts: Introduction; Test Design and Development; Scoring; Classical Item-Level Statistics; Reliability; Validity; Calibration, Equating, and Scaling (CES); Standard Setting; and Summary of Operational Test Results. An overview of this manual is provided below.

Part 1 Introduction

Part 1 presents the background of the NYSESLAT, its rationale and purpose, recommended test use, and test accommodations. Test accommodations include Large Type and braille.

Part 2 Test Design and Development

Part 2 describes the test development process of the NYSESLAT. It includes test specifications, item development and review processes, item field-testing, and test construction.

Part 3 Scoring

Part 3 provides a description of the scoring process. It includes a description of the range-finding meeting in Albany, New York, in 2010. It also provides information about the audit-scoring process that was conducted on a ten-percent sample, and a description of rater training and inter-rater reliability.

Part 4 Classical Item-Level Statistics

Part 4 provides item-level descriptive statistics based on Classical Test Theory (CTT).

Part 5 Reliability

Part 5 explains the internal consistency reliability, classical Standard Error of Measurement (SEM), conditional SEM based on IRT, and inter-rater reliability. It also provides results of the inter-rater and intra-rater reliability, the rater agreement analyses, the reliability of each of the four modalities, and the reliability of the classification decision at the proficient cut.

Part 6 Validity

Part 6 describes the validity studies that were conducted. It includes evidence of validity based on test content, internal structure, and relationships to other variables.

Part 7 Calibration, Equating, and Scaling

Part 7 explains the Rasch and Partial Credit Models and provides sample-item characteristic curves for a one-step item and a two-step item. It also includes the process of the calibration, equating, and scaling of the 2010 administration of the NYSESLAT. Part 7 also explains the rationale for the use of the IRT model. It includes the IRT model fit statistics and the average Rasch difficulty of the subtests.

Part 8 Standard Setting

Part 8 presents the standard-setting process that was followed to establish the performance level cuts. It includes the standard-setting model, the standard-setting process, summary statistics for the round-by-round ratings, evaluation results, post-standard-setting analyses, and final performance-level cut points.

Part 9 Summary of Operational Test Results

Part 9 presents the raw score summary, scale score summary, percentage of students in each performance category, and exit rate for the 2007, 2008, 2009, and 2010 administrations of the NYSESLAT.

PART 1: INTRODUCTION

1.1 Background

Title III of the Federal *No Child Left Behind* (NCLB) Act of 2001 requires annual assessment of the English language skills of English language learner (ELL) students. Section 3121(d)(1) explains that each state must use evaluation measures designed to assess “the progress of children in attaining English proficiency, including a child’s level of comprehension, speaking, listening, reading, and writing skills in English.” NCLB requires demonstrated annual improvement in English proficiency for such students in order for them to meet “challenging State academic content and student academic achievement standards.” New York State regulations also require annual assessment of ELL students using a state-approved assessment.

NCLB requires that the annual assessment of ELL students be based on specific student achievement objectives. Section 3113(b)(2) explains that the “standards and objectives for raising the level of English proficiency [will be] derived from the four recognized domains of speaking, listening, reading, and writing, and [will be] aligned with achievement of the challenging State academic content and student academic achievement standards” set out in Title I of the Act.

The New York State Education Department (NYSED) developed *Learning Standards for English as a Second Language* to meet the requirements of NCLB. New York’s learning standards and performance indicators (achievement objectives) for English as a Second Language (ESL) students are derived from the domains of Speaking, Listening, Reading, and Writing, and they align with the State’s English Language Arts standards. New York’s ESL learning standards and performance indicators are organized in four grade spans: PreK–1, 2–4, 5–8, and 9–12. To meet Federal and State requirements regarding the assessment of ELL students, NYSED requested test development, research, and scoring based on the State’s *Learning Standards for English as a Second Language*. Accordingly, the New York State English as a Second Language Achievement Test (NYSESLAT) was developed with four subtests—Speaking, Listening, Reading, and Writing—in each of five grade spans: K–1, 2–4, 5–6, 7–8, and 9–12. Individual test items match to specific ESL learning standards and performance indicators. In addition, the test was developed in accordance with the Standards for Educational and Psychological Testing (American Educational Research Association, 1999) and New York State testing requirements. The test is consistent with the principles of Universal Design, which means that the test is as accessible as possible to all populations, including special-needs students. The test is also consistent with applicable federal and state testing requirements.

In response to NYSED’s request, Pearson¹ offered a solution with two distinct phases. With respect to the first phase, custom items were developed and used to create the 2010 NYSESLAT

¹Pearson purchased Harcourt Assessment, Inc. in 2009. Pearson is now the vendor for the NYSESLAT.

operational forms. Anchor items were pulled from the 2009 test and used to produce custom forms for the 2010 test administration. With respect to the second phase, brand new custom items were developed for the 2010 test administration.

1.2 Rationale and Purpose

The New York State Board of Regents has established learning standards for all ELL students attending New York State schools. NCLB mandates that all ELL students from Kindergarten through Grade 12 are to be assessed every year to measure their English language proficiency in speaking, listening, reading, and writing, and that their annual progress toward proficiency is tracked. In compliance with this directive, NYSED developed an annual test that measures student progress toward meeting the State's ESL learning standards. This test is the NYSESLAT. The NYSESLAT helps schools determine which instructional standards teachers must devote time to in order to ensure their ELL students fully acquire the language proficiency that will prepare them for success in the classroom. The purpose of the test is to measure annual student improvement in achieving English language proficiency in order to ultimately exit an ESL or bilingual education program, move into an English Language Arts classroom, and function successfully without any additional support.

1.3 Recommended Test Use

The NYSESLAT is used to make decisions for language instructional programs and accountability determinations. First, the test is used to measure the level of English proficiency of all ELL students in listening, speaking, reading, and writing. The English proficiency level, as measured by the NYSESLAT, is used by school districts to decide the type and amount of instructional services to which the students are entitled. Second, the NYSESLAT is used to measure students' annual progress toward learning English and attaining English language proficiency; it determines whether school districts and the State meet the required progress and attainment targets, as defined in the New York State Title III accountability system. Third, the NYSESLAT is used as the sole criteria for exiting ELL students who score at the proficient level from the ESL/Bilingual programs.

1.4 Test Accommodations

All test items are developed following the guidelines of Universal Design. Adherence to these guidelines ensures that the assessments are accessible and valid for the widest range of students, including students with disabilities. Applying universal test design during the development process helps eliminate the need to address after-the-fact accommodations, and universal test design provides a better assessment for all students. Checklists are used to review every item to ensure that each is built with consideration of the following: equitable use, flexibility in use, simple intuitive design, perceptible information, tolerance for error, low physical effort, and size

and span for approach and use. During forms construction, Pearson utilizes in-house content and fairness experts to ensure that the forms are pulled with concepts of universal design in mind. Pearson stringently reviews forms for special populations—such as visually or hearing-impaired students—to ensure that items are fair, reliable, and accessible to all.

Large Type

Pearson has standardized large-type product specifications that ease the test-taking experience for visually impaired children who require large type. One font size (i.e., a minimum 18-point type for items and passages, and a font size no larger than 24-point type for titles) is produced for each grade span. Pages are printed in black only and on a cream-colored, non-glare vellum stock to ease readability of pages. Covers are printed on heavier stock to provide stiffness to the booklets, which protects interior text pages. Plastic spiral binding makes turning of pages easy to accomplish.

Braille

Pearson produces a Braille version of the NYSESLAT for visually impaired children using certified and experienced transcribers who can work with the multiple codes, rules, and guidelines. Pearson produces Braille forms for each NYSESLAT subtest and grade span. For the K–1 grade span, a checklist is provided rather than a Braille test.

If a content area is difficult for Braille, Pearson and content specialists determine if there are other ways that the construct could be worded or measured. To adapt some items, pictures are described. Care is taken not to convey the correct answer but to give a description that would enable the student to ascertain the correct answer.

PART 2: TEST DESIGN AND DEVELOPMENT

2.1 Overview

To meet the requirements of Title III of the Federal *No Child Left Behind* (NCLB) Act and of New York State regulations regarding the assessment of limited English proficient students, NYSED developed *Learning Standards for English as a Second Language*, a comprehensive set of ESL learning standards and performance indicators in the four domains of Speaking, Listening, Reading, and Writing. NYSED then requested test development, research, and scoring based on these standards. The NYSESLAT consists of a test in each of the five grade spans (K–1, 2–4, 5–6, 7–8, 9–12). The five tests are vertically scaled and each consists of four subtests: Listening, Speaking, Reading, and Writing. The tests are designed to assess the English language skills of ELL students in Grades K–12 and to capture their progress toward achieving full English language proficiency. The tests were developed in accordance with the Standards for Educational and Psychological Testing (American Educational Research Association, 1999) and New York State testing requirements. The test is consistent with the principles of Universal Design, making it as accessible as possible to all populations, including special-needs students, and is consistent with applicable federal and state testing requirements.

2.2 Test Blueprint and Specifications by Modality and Grade Span

The NYSESLAT is divided into four basic domains or modalities—Speaking, Listening, Reading, and Writing—for Grades K–12. It includes multiple-choice, constructed-response, short-response, and extended-response items. The total number of items per grade span varies. For grade span K–1, there are 70 items; for grade span 2–4, there are 80 items; and for grade spans 5–6, 7–8, and 9–12, there are 87 items each.

The Speaking modality has 16 constructed-response items for all grade spans. The Listening and Reading modalities consist of only multiple-choice items. The number of items for the Listening modality varies from 24–25 for the different grade spans. The number of items for the Reading modality varies from 15–27 for the different grade spans. The number of items for the Writing modality ranges from 15–19 for the different grade spans. The Writing modality is composed of the following parts:

- Multiple-choice section that assesses ELL students’ understanding of the principles of written English at the phoneme, word, and sentence levels
- Developmental writing items (K–1 only)
- Pre-writing activity (Grades 2–12)
- Extended response to graphic-based prompt (Grades K–12)

The operational test blueprint for the 2010 administration of the NYSESLAT is presented in Table 1. Test specifications by modality and grade span are shown in Table 2. Table 3 provides the maximum number of points by modality and grade span. Table 4 provides a breakdown of

the number of items based on New York State’s ESL learning standards and grade spans. The 2010 test design consists of all custom-developed items for the NYSESLAT.

Table 1 - 2010 NYSESLAT Operational Test Blueprint

NYSESLAT 2010 Operational Test Blueprint							
Strand	Cluster	Number of Items					Notes
		K-1	2-4	5-6	7-8	9-12	
Listening	Word/Sentence Comprehension	14	10	3	3	3	All items in this strand are multiple-choice items with 0-1 score points.
	Comprehension of Conversational Language	10	9	12	12	12	
	Synthesizing Information (Task-Based Listening)		5	10	10	10	
	Number of Items	24	24	25	25	25	
Number of Points		24	24	25	25	25	
Reading	Word Reading	11					All items in this strand are multiple-choice items with 0-1 score points.
	Sentence Reading	2					
	Short Passages with Questions	2					
	Word/Sentence Reading		7	3	3	3	
	Comprehension		17	24	24	24	
Number of Items		15	24	27	27	27	
Number of Points		15	24	27	27	27	

Table 1 - 2010 NYSESLAT Operational Test Blueprint (continued)

NYSESLAT 2010 Operational Test Blueprint							
Strand	Cluster	Number of Items					Notes
		K-1	2-4	5-6	7-8	9-12	
Writing Conventions	Phonemic Understanding	6	4	2	2	2	All items in this strand are multiple-choice items with 0-1 score points.
	Mechanics & Structure		8	13	13	13	
Writing	Developmental Writing	8					Level K-1 has 8 CR items with 0-2 score points.
	Pre-Writing		3	3	3	3	Levels 2-4, 5-6, 7-8, 9-12 each has 3 CR items with 0-2 score points.
	Extended Response	1	1	1	1	1	Each of the five levels has one ER item. The ER item for Level K-1 has 0-2 score points. The ER items for the Levels 2-4, 5-6, 7-8, and 9-12 each has 0-4 score points.
Number of Items		15	16	19	19	19	
Number of Points		24	22	25	25	25	
Speaking	Sentence Completion	5	5	5	5	5	Each of the five levels has 5 SS items with 0-2 score points in this cluster.
	Storytelling	1	1	1	1	1	Each of the five levels has 1 SE item with 0-4 score points in this cluster.
	Picture Description	5	5	5	5	5	Each of the five levels has 5 SS items with 0-2 score points in this cluster.
	Social Interaction	5	5	5	5	5	Each of the five levels has 5 SS items with 0-2 score points in this cluster.
Number of Items		16	16	16	16	16	
Number of Points		34	34	34	34	34	
Total Number of Items		70	80	87	87	87	
Total Number of Points		97	104	111	111	111	

Table 2 - 2010 Test Specifications by Modality and Grade Span

Number of Items and Passages in NYSESLAT Subtests							
Grade Span	Speaking	Listening	Reading	Writing			Total Number of Items per Grade Span
				Writing Conventions	Pre-Writing	Writing Prompt	
	CR	MC	MC/Passages	MC	SR	CR/ER	
K-1	16	24	15/0	6	0	8 CR (Dev. Writing) + 1 ER	70
2-4	16	24	24/5	12	3	1 ER	80
5-6	16	25	27/5	15	3	1 ER	87
7-8	16	25	27/5	15	3	1 ER	87
9-12	16	25	27/5	15	3	1 ER	87

Table 3 - Maximum Number of Points by Modality and Grade Span

Maximum Number of Points in NYSESLAT Subtests							
Grade Span	Speaking	Listening	Reading	Writing			Total Number of Points per Grade Span
				Writing Conventions	Pre-Writing	Writing Prompt	
	CR	MC	MC/Passages	MC	SR	CR/ER	
K-1	34	24	15/0	6	0	16 CR (Dev. Writing) + 2 ER	97
2-4	34	24	24/5	12	6	4	104
5-6	34	25	27/5	15	6	4	111
7-8	34	25	27/5	15	6	4	111
9-12	34	25	27/5	15	6	4	111

Table 4 - Test Specifications for New York State ESL Learning Standards by Grade Span

Grade Span	Standard*	Number of Items	Percentage of Total Items	Number of Score Points	Percentage of Total Points
K-1	1	37	53	46	47
	2	6	09	11	11
	3	11	16	15	15
	4	8	11	14	14
	5	8	11	11	11
	Total Test		70	100	97
2-4	1	45	56	50	48
	2	12	15	15	14
	3	12	15	22	21
	4	9	11	14	13
	5	2	03	3	03
	Total Test		80	100	104
5-6	1	46	53	52	47
	2	14	16	18	16
	3	9	10	16	14
	4	15	17	21	19
	5	3	03	4	04
	Total Test		87	100	111
7-8	1	45	52	48	43
	2	10	11	13	12
	3	13	15	21	19
	4	14	16	22	20
	5	5	06	7	06
	Total Test		87	100	111
9-12	1	34	39	36	32
	2	12	14	15	14
	3	14	16	19	17
	4	17	20	26	23
	5	10	11	15	14
	Total Test		87	100	111

*Standard 1: Students will listen, speak, read, and write in English for information and understanding. Standard 2: Students will listen, speak, read, and write in English for literary response, enjoyment, and expression. Standard 3: Students will listen, speak, read, and write in English for critical analysis and evaluation. Standard 4: Students will listen, speak, read, and write in English for classroom and social interaction. Standard 5: Students will demonstrate cross-cultural knowledge and understanding.

2.3 Test Mapping and Specifications by New York State Learning Standards for English as a Second Language

Table 5 provides an item-mapping summary, and Appendix A.1 provides, in detail, the 2010 item mapping by New York State Learning Standards for each item within each grade span and modality. Item specifications are presented in Appendix A.2.

Table 5 - Item Mapping by New York State ESL Learning Standards and Performance Indicators

Grade Span	Subtest	Standards	Performance Indicators	Items in Subtest	Percentage of Subtest Score
K-1	Speaking	Standard 1	PI 6	7	6
		Standard 2	PI 5 8	2, 6	13
		Standard 3	PI 1 3 9	8, 9, 13, 4,	25
		Standard 4	PI 3 5	1, 10, 11, 12, 15, 16,	38
		Standard 5	PI 3 6	3, 5, 14	19
	Listening	Standard 1	PI 1 3 4 6 9 12	1, 2, 3, 4, 6, 7, 9, 12, 13, 23	42
		Standard 2	PI 2 5	15, 16, 21	13
		Standard 3	PI 1 9	5, 10, 11, 14, 20, 24	33
		Standard 4	PI 3 5	17, 19,	8
		Standard 5	PI 2 3	8, 18, 22	13
	Reading	Standard 1	PI 1, 3, 5, 12	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13	50
		Standard 3	PI 5	15	7
Standard 5		PI 1, 6	7, 14	13	
Writing Conventions Writing	Standard 1	PI 12	1, 2, 3, 4, 5, 6	40	
	Standard 1	PI 8, 11, 12	7, 8, 9, 10, 11, 12, 13, 14,	53	
	Standard 2	PI 8	15	7	
2-4	Speaking	Standard 1	PI 5, 6, 9	1, 2, 3, 5, 7	31
		Standard 2	PI 8	6	6
		Standard 3	PI 1, 3	4, 8, 11, 13,	25
		Standard 4	PI 3, 5, 9	9, 12, 14, 15, 16	31
		Standard 5	PI 4	10	6
	Listening	Standard 1	PI 1, 2, 3, 4, 5, 6, 10, 16	1, 2, 3, 4, 5, 6, 8, 9, 11, 12, 16, 19, 22, 23, 24	63
		Standard 2	PI 2, 3, 4, 5	7, 10, 14, 17 21	17
		Standard 4	PI 3, 5, 7, 8	15, 18, 20, 21	17
		Standard 5	PI 6	13	4
	Reading	Standard 1	PI 1, 3, 4, 5, 6, 16	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 16, 17, 18	58
		Standard 2	PI 2, 4, 5, 12	12, 13, 14, 15, 20, 21, 22	29
		Standard 3	PI 1, 3, 5	19, 23, 24	13
Writing Conventions Writing	Standard 1	PI 12	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12	69	
	Standard 3	PI 1	11	6	
	Standard 3	PI 1, 3	13, 14, 15, 16,	25	
5-6	Speaking	Standard 1	PI 5, 6, 9	1, 3, 4, 5, 7, 11	38
		Standard 2	PI 1, 8	6, 16,	13
		Standard 3	PI 1	13	6
		Standard 4	PI 4, 5	2, 8, 9, 10, 12, 15	38
		Standard 5	PI 3	14	6
	Listening	Standard 1	PI 1, 3, 4	1, 2, 3, 5, 7, 8, 10, 11, 13, 20, 21, 24	44
		Standard 2	PI 1, 2	10, 15	8
		Standard 3	PI 1	12	4
		Standard 4	PI 3, 7, 8	4, 9, 16, 17, 18, 19, 22, 23, 25,	36
		Standard 5	PI 2	6, 14	8

Table 5 - Item Mapping by New York State ESL Learning Standards and Performance Indicators (continued)

Grade Span	Subtest	Standards	Performance Indicators	Items in Subtest	Percentage of Subtest Score
5–6	Reading	Standard 2	PI 2, 4, 5	9, 11, 13, 14, 15, 17, 18, 24, 25, 27	37
		Standard 3	PI 3, 9	6, 8, 12	11
	Writing Conventions Writing	Standard 1	PI 1, 12	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	79
		Standard 3	PI 1, 3, 9	16, 17, 18, 19	21
7–8	Speaking	Standard 1	PI 5, 12	1, 8	11
		Standard 2	PI 8	6	6
		Standard 3	PI 1, 9	9, 10, 14	19
		Standard 4	PI 3, 4, 5, 9	2, 3, 5, 7, 12, 13, 15, 16	50
		Standard 5	PI 3, 6	4, 11	11
	Listening	Standard 1	PI 1, 3, 4, 5, 6, 9, 10	1, 4, 5, 7, 9, 10, 11, 12, 13, 19, 20, 21, 23, 24, 25	60
		Standard 3	PI 2, 9	2, 14	8
		Standard 4	PI 3, 7	6, 8, 16, 17, 18, 22	24
		Standard 5	PI 1, 3	3, 15,	8
	Reading	Standard 1	PI 1, 2, 3, 6, 9, 16	2, 3, 4, 7, 9, 10, 11, 12, 13, 14, 16, 17	44
		Standard 2	PI 2, 4, 5	19, 20, 21, 22, 23, 24, 25, 26, 27	33
		Standard 3 Standard 5	PI 1, 2, 5, 9 PI 2	1, 5, 8, 15, 18 6	19 4
	Writing Conventions Writing	Standard 1	PI 1, 12, 15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	79
		Standard 1	PI 4	16	5
		Standard 3	PI 1, 9	17, 18, 19	16
	9–12	Speaking	Standard 1	PI 6, 9	8, 9
Standard 2			PI 8	6	6
Standard 3			PI 9	11	6
Standard 4			PI 2, 3, 4, 5, 9	2, 3, 7, 10, 12, 13, 14, 15, 16,	56
Standard 5			PI 2, 3, 6	1, 4, 5	19
Listening		Standard 1	PI 1, 2, 3, 4, 6, 16	1, 2, 3, 9, 8, 10, 12, 14, 15, 16, 18, 21, 24, 25	56
		Standard 3	PI 3, 5, 9	4, 5, 13, 20	16
		Standard 4	PI 7, 8	6, 7, 11, 17, 19, 22, 23	28
Reading		Standard 1	PI 1, 5, 6, 10, 16	1, 2, 3, 5, 7, 14, 15, 24	26
		Standard 2	PI 2, 3, 4, 5, 11, 12	8, 16, 19, 20, 21, 22, 23, 24, 25, 26, 27	41
		Standard 3 Standard 5	PI 2, 3, 5, PI 2, 3, 5	10, 11, 13, 17, 18 4, 6, 9, 12	19 15
Writing Conventions Writing		Standard 1	PI 6, 12	1, 2, 6, 8, 9, 10, 11, 12, 13, 14, 15	58
		Standard 3	PI 2, 5	5, 7	11
		Standard 4	PI 7	3	5
		Standard 5	PI 1	4	5
		Standard 3 Standard 5	PI 1, 1 PI 6	16, 19 17, 18	11 11

2.4 Item Development and Review Processes

All items on the 2010 NYSESLAT were written by educators of ELL students, including a number of New York State teachers. An Item Writers' Training Guide was sent out to all item writers to assist them in developing the items. In addition, Pearson specialists were available via phone and/or e-mail to answer any questions the item writers had. Assessment specialists at Pearson reviewed the passages and items, and in accordance with the item specifications, the assessment specialists ensured the following:

- Absence of bias and sensitive topics in passages
- Item soundness can be interpreted as item validity. An item's soundness is based on the idea that the information derived from an item and its consequences are true or valid.
- Absence of bias in items
- Appropriateness of topic, vocabulary, and language structure for each grade span
- Items matched to the intended New York State Learning Standards for English as a Second Language

The Reading passages were then reviewed by a committee of New York State ESL teachers. Each test question was also reviewed by grade-level ESL educators from New York State. Only those passages and test questions judged to be of acceptable quality and fairness to students from the various ethnic/cultural backgrounds of those who presently live in New York State were approved for use. In addition to the committee reviews, a "Final Eyes" review was conducted by a separate group of New York State teachers to ensure items were error-free and appropriate. After items were field tested, New York State teachers were invited to a range-finding session, during which they reviewed potential anchor papers and select practice sets that were used to aid teachers in scoring the 2010 NYSESLAT.

Anchor items serve the purpose of providing statistics that help equate tests from one year to the next, so that the scaled scores between them are comparable. Anchor item selection criteria are described in Section 7.4.

2.5 Field Test

Pearson constructed, and NYSED reviewed and approved, stand-alone field test forms that were administered to a representative sample of limited English proficient students around New York State. There were two field test forms at each grade span, and each form was split into four mini-forms for the four modalities (Listening, Speaking, Reading, and Writing). Each modality mini-form is administered to a representative sample of New York State schools. The field tests are conducted in all New York State schools with ELL populations. To accommodate the field test needs, school districts have been divided into four stratified random samples (S1, S2, S3, S4), using the Need/Resource Capacity Index as the strata. Each of the four samples is assigned to administer a field test in one modality each year. The field test assignments systematically rotate among the field test samples each year (see Table 6). For example, districts in Sample 2 (S2)

were assigned to administer a field test in Listening in 2007. In 2008, the same Sample 2 districts were assigned to administer a field test in Writing.

Table 6 - Four-Year Sampling Matrix for the NYSESLAT Field Test

Field Test Form	2006	2007	2008	2009
Listening	S1	S2	S3	S4
Speaking	S2	S3	S4	S1
Reading	S3	S4	S1	S2
Writing	S4	S1	S2	S3

The 2010 NYSESLAT test administration did not include a field test administration. The data from the 2009 field-test administration informed construction of the 2010 NYSESLAT operational test. The 2009 field-testing sampling plan and statistics presented to NYSED by Pearson are discussed in the *2009 NYSESLAT Fall Field Test Technical Report*.

2.6 Test Construction

Items selected for the 2010 NYSESLAT represented a complete range of difficulty at all grade levels from K–12. Items ranged from easier ones with high p -values that were primarily aimed at measuring the skills of students with very limited ability in English to items with fairly low p -values that were aimed at measuring the skills of students with advanced ability in English. The number of both multiple-choice and constructed-response items increased at each proficiency level, meeting the requirements of the NYSED.

The selection criteria for choosing items for the 2010 NYSESLAT operational forms included the following: item content, skill measured, item difficulty, gender balance, and ethnic balance. Content considerations include having a variety of settings and activities represented in items; having a mix of people, animals, and objects represented; and ensuring that there is no content overlap within a subtest or across subtests in the same grade span. Skills measured varied by subtest and by sections within a subtest. For example, the Word/Sentence Reading items in the first section of the Reading subtest for Grades 2–12 tested the ability of students to read a mixture of nouns, verbs, noun phrases, verb phrases, adjectives and adjective phrases sometimes, and adverbs occasionally. The Reading Comprehension items in the second section of the Grades 2–12 Reading subtest tested the ability of students to identify the main idea or the topic of a passage; to understand details in a passage; to make inferences based on information in a passage; to infer the meaning of a word from context; to infer the author’s purpose; and to understand why information was organized or presented in a particular way in a passage. Similar specific content criteria were applied for the other subtests. As explained above, items ranged from easier ones with high p -values to items with low p -values. The difficulty level of Reading passages, as distinct from the difficulty of the Reading Comprehension items on the passages, was a criterion, as well, for the Reading subtest. Finally, gender and ethnic balance was considered in item selection. The people represented in the items across a subtest represented a balance of males and females and a variety of ethnic backgrounds. Differential Item Functioning (DIF) analyses, described in Section 4.2, were performed to identify potential items that needed

to be scrutinized for item bias. None of the items that showed statistical significant DIF were considered by content specialists to be biased based on either gender or ethnicity.

The process for constructing the 2010 NYSESLAT operational forms was as follows: As soon as Pearson had clean statistics from the spring 2009 Field Test, Pearson ESL assessment specialists reviewed all items from the field test and pulled operational forms based on the above criteria. In addition to the spring 2009 Field Test items, items were pulled from the 2009 Operational Test to be used as anchor items in the 2010 Operational Test. The anchor items were necessary for placing the 2010 Operational Items on the NYSESLAT established scale and constituted, roughly, 25% of the items in each subtest. The previously indicated selection criteria for the non-anchor items (unique items) apply to anchor items as well. Pearson produced draft operational forms that were then presented to NYSED. A joint Pearson-NYSED team reviewed the forms, following all item-selection criteria, and made various changes to the draft forms. Final forms were then reviewed by Pearson ESL assessment specialists, editorial staff, psychometricians, and quality control staff, as well as by NYSED.

Psychometric Guidelines for Operational Item Selection and Form Construction

Statistical considerations included item difficulty and other statistical characteristics of the items. *P*-values were distributed between approximately 0.30 to 0.95, with fewer items at the extremes of difficulty and more items of moderate difficulty. In addition to selecting items with appropriate *p*-values, staff members were given the following instructions during the item selection and form construction process:

- A. Item Statistics:
 1. Check item difficulty
 2. Check the range of item difficulty: Items should be flagged if *p*-value is < 0.30 or > 0.95
 3. Check the point-biserial range: Try to avoid items with pt. biserial < 0.30
 4. Check the omit rate: Watch for items with an omit rate $> 5\%$
 5. Avoid items with a DIF bias flag. If it is necessary to select an item with a DIF flag, then it needs to be reviewed carefully during the operational test construction
- B. There are no changes to items once the item is field-tested:
- C. Tests are built using the following statistical targets
 1. Average *p*-value comparison between the spring 2010 form and spring 2009 form
The average *p*-values should be similar to one another
 2. Average *p*-value comparison between the anchor item set for spring 2010 and the entire spring 2009 test: The average *p*-values should be similar to one another.
- D. The total number of items at each level and the number of items within each strand must follow the test blueprint

In general, flagged items should be avoided. However, the match to blueprint and content considerations should take priority over statistical targets in most circumstances.

Testing Written Language

A fundamental consideration in constructing the NYSESLAT is the language that is being tested. While the test developer's native-speaker intuition provides choices about what language is tested, more rigorous methods in language choice are applied to provide consistency across the forms of the five grade spans and to create a vertical structure within each form wherein language ranges from the most simple—that which is first acquired by non-native speakers—to advanced language that would indicate a level of English proficiency sufficient for participation in regular academic classes.

For the NYSESLAT (a test designed to assess students at all proficiency levels—beginning, intermediate, advanced, and proficient within each grade span), this vertical development of the language tested allows the test to discriminate more finely among students at different stages of language acquisition. Being able to identify accurately students at different levels of language development provides better information to classroom teachers, who must find the most effective way to help their students reach proficiency. It also provides the very important evidence of students' progress toward proficiency, which is required by the NCLB legislation.

To determine the appropriate language for Reading passages, Pearson assessment specialists, editorial staff, and item and passage writers applied the Flesch-Kincaid grade-level readability analyses to all Reading passages. Readability measures are primarily based on factors such as the number of words in the sentences and the number of letters or syllables per word. Additionally, ESL assessment specialists also evaluated the coherence of a passage, the use of anaphora, vocabulary difficulty, sentence and text structure, and concreteness and abstractness. Finally, input from the New York State teacher committee that reviewed the passages was considered. The sum of these analyses and evaluations determines the appropriateness of the language of a passage.

There is a gradual increase in difficulty from passage to passage at every grade span. Each form includes beginner-level passages, as well as passages that are representative of on-grade Reading passages found on English Language Arts reading tests. Pearson uses the *Educational Developmental Laboratory (EDL) Core Vocabularies in Reading, Mathematics, Science, and Social Studies*, published by Steck-Vaughn, to help determine age- and grade-appropriate language for ELL items and stimuli for the oral-language subtests. Furthermore, Pearson ESL assessment specialists and editors ensure that the language in all stimuli and items, from Kindergarten through Grade 12, is both topic- and age-appropriate for test takers.

Testing Oral Language

Recognizing that oral language structure and vocabulary in English differ vastly from the written language, issues of oral language assessment among Kindergarten through Grade 12 ELL students have been the subject of special investigation at Pearson. Pearson's English language proficiency professionals have conducted research on the item types that appear in the NYSESLAT Speaking and Listening subtests by presenting examples of these item types to ELL students in cognitive laboratories. This setting allows for careful observation and recording of

student responses and student reactions to items. Outcomes of this procedure led to the following important design decisions:

- Item types
- Number of items
- Length of pauses between items
- Use of recorded stimuli
- Recording student spoken responses

The Speaking and Listening subtests of the NYSESLAT are based on these decisions. To ensure that the language in the Speaking and Listening stimuli and items reflects current spoken language as much as possible, Speaking and Listening scripts are submitted to a read-aloud proofing process with ELL assessment specialists and editors. Additionally, for the oral components of the NYSESLAT to be relevant, the Speaking and Listening subtests had to have predictive validity for academic achievement. Therefore, both academic and social languages are integral components of the Speaking and Listening subtests of the NYSESLAT.

PART 3: SCORING

This part describes the process for scoring the field test, operational test, and Audit. All field test scoring is done by Pearson at the San Antonio site. For field test scoring, each grade span had at least one team of five readers scoring. A 10 percent check score (2nd reading) was done by team leaders. Anchors, training sets, and rubrics were used as scoring guides. Readers scored Pre-Writing (PW) 1, 2, then 3 successively. If questions arose during scoring, usually the problem was discussed by the group in order to maintain consistency in scoring. The operational test scoring was done by New York State teachers. The final operational test data set was provided by the New York State Data Repository, with rubrics and training material as scoring aids. An audit of the operational test was done by Pearson readers at the San Antonio site. Ten percent of the total operational responses are scored in this audit. The same check-score procedures used in scoring the field test were used in the audit scoring. The details of the scoring process for operational items are described below.

3.1 NYSESLAT Range Finding

During the 2009 administration year, range finding was held in Albany, New York, on July 7–10, 2009, before field-test items were scored by Pearson. The range-finding process was not required because field-test administration was not included in the 2010 administration.

3.2 Operational File

The operational file for the spring 2010 NYSESLAT came from the New York State Data Repository after scoring was completed. The file was verified by Pearson's Assessment and Information Quality (AIQ) Department to ensure data accuracy based on the description values in the file layout. Sections 3.3 through 3.4 indicate the procedures that Pearson conducted to assure reliable and accurate scoring of the items.

3.3 Rater Training

The accuracy of scoring was monitored by training team leaders who were experienced, proficient readers. These team leaders successfully completed a two-day general team leader-training workshop and had the experience of training as Room Directors for many custom projects. These trainers were seasoned Performance Scoring Center (PSC) readers who had vast experience in all facets of scoring. They carefully monitored the scoring and accuracy of their teams of readers. All responses received a single reading with at least 10 percent getting a second reading to monitor for reliability and accuracy. Team leaders independently gave the check score. All readers maintained at least an 88 % agreement rate.

All PSC readers had a minimum of a bachelor's degree and successfully completed generalized workshops in performance assessment scoring before being considered as potential readers for a specific project such as NYSESLAT. Training of readers was based on anchors and training sets

that were part of the sampler incorporated into the training procedures for operational scoring done by New York State teachers.

3.4 Inter-Rater and Intra-Rater Reliability

All readers were trained to score to the same scale to ensure accurate, consistent, reliable scoring. PSC adhered to stringent criteria in its general screening, training, and qualifying procedures as preliminary measures for obtaining high levels of consistency and reliability. Team leaders conducted “read behinds,” reading the same booklets after readers, to check for accuracy of scores. In addition, 10% of all booklets were “second” read by the training team leader to check accuracy. If individuals were not “on track,” retraining ensued, as these readers were being systematically too lenient or too harsh in their ratings, or varied unsystematically and unpredictably from other raters and deviated from the training standards in their scores. The statistics of rater reliability are presented in Section 5.6.

PART 4: CLASSICAL ITEM-LEVEL STATISTICS

4.1 Item-Level Descriptive Statistics

This section presents the raw score summary statistics for all items in the spring 2010 administration of the NYSESLAT within the framework of Classical Test Theory. The compiled file from the New York State Data Repository was used to obtain all raw score statistics. The p -value for each item is defined as the proportion of students that answer an item correctly for the multiple-choice items. A high p -value means that an item is easy; a low p -value means that an item is difficult. For the constructed-response items, the p -value is reported as the average number of points out of the maximum number of possible points.

The point-biserial correlation for each item is an index of the association between the item-score and the total test score. It shows the ability of the item to discriminate between low-ability and high-ability students. An item with a high point-biserial correlation discriminates more effectively between the low- and the high-ability students than a low point-biserial correlation.

The item-level statistics for the operational 2010 NYSESLAT are presented in Appendices B.1–B.5 by grade span. The tables are grouped by Listening/Speaking and Reading/Writing modality combinations. The following item information and statistics are presented for each item:

- Item number
- Item format (multiple-choice, constructed-response, short-response, or extended-response)
- Maximum number of possible points
- N-count (number of students)
- Response options for multiple-choice items and percentage of students obtaining each score point for constructed-response items
- Omits (percentage of students omitting an item)
- P -value for multiple-choice items (percentage of examinees that answered the item correctly)
- Point biserial (index of discrimination between high and low scoring students)

Items that are too easy or too difficult are flagged based on their p -values. Such items do not provide adequate information, and their inclusion serves a limited purpose in the measurement process. However, the NYSESLAT, being a standard-referenced examination, allows exception to the rule if the item is deemed necessary by content experts. Although there is no consensus as to what is an acceptable p -value, the guiding policy during item review and form building is to choose a p -value ≤ 0.90 and ≥ 0.30 .

As explained above, the point biserial is another index that indicates the statistical suitability of an item for inclusion in the examination. Since it is an index of correlation of the item to the total test, one would expect a larger point biserial to indicate a desirable value for item inclusion.

However, in the context of form building, items are selected not only for their adherence in measuring the underlying trait—unidimensionality (high correlation)—, but also based on local independence of the item (low correlation). A very high correlation, in fact, may mean a redundancy of the item; that is, the item performs the same function as some other items. On the other hand, too low a value for the point biserial may indicate a “shaky” association of the item with the underlying trait.

For the NYSESLAT, the upper limit of the point biserial is fixed at 0.80, while the lower limit is fixed at 0.25. Although there is seldom an occurrence of the upper limit violation for the NYSESLAT, the negative value of the point biserial is stringently observed, since this may indicate a wrong key during the scoring process or an item with a very bad distractor or distractors. These items are never used in the form-building process. Once again, the setting of the upper and lower limits is subjective and may be acceptable in many cases, especially if the examination of the response pattern for each distractor shows a negative point biserial with the positive value being only for the correct answer.

As can be seen in Appendix B, with the exception of some high p -values and point-biserial values, all the items fall well within our pre-set level of acceptance, both in terms of the p -value and the point biserial. The mean p -values and point-biserial indices are presented in Table 7. The mean p -values and point-biserial indices at each grade span for the combined modalities of Listening/Speaking and Reading/Writing are in the middle range and are acceptable. The total population had 235,403 records. After removing duplicates and off-level records (N=17,326), the total N= 218,077. Invalid raw scores were also removed from data for the final sample sizes used in Table 7:

	Listening/Speaking					Reading/Writing				
	Initial N	Dup Removed	Off-Level Removed	Invalid RS	Final N	Initial N	Dup Removed	Off-Level Removed	Invalid RS	Final N
K-1	57399	1203	532	4000	53399	57399	1203	532	3604	53795
2-4	69786	1782	691	4433	65353	69786	1782	691	7378	62408
5-6	32996	1040	534	2690	30306	32996	1040	534	4175	28821
7-8	28404	754	549	2177	26227	28404	754	549	3365	25039
9-12	46818	689	2245	4026	42792	46818	689	2245	5884	40934

Table 7 - Summary of Classical Item Difficulty and Item Discrimination Indices for Each Grade Span

Grade Span	Modality	N-Count	Item Difficulty		Item Discrimination	
			Mean	SD	Mean	SD
K-1	Listening	53399	0.85	0.08	0.37	0.08
	Speaking	53399	0.80	0.04	0.69	0.05
	Reading	53795	0.77	0.09	0.49	0.08
	Writing	53795	0.78	0.12	0.57	0.18
	Listening and Speaking	53399	0.83	0.07	0.50	0.07
	Reading and Writing	53795	0.78	0.11	0.53	0.13

Table 7 - Summary of Classical Item Difficulty and Item Discrimination Indices for Each Grade Span (continued)

Grade Span	Modality	N-Count	Item Difficulty		Item Discrimination	
			Mean	SD	Mean	SD
2-4	Listening	65353	0.82	0.11	0.41	0.07
	Speaking	65353	0.84	0.04	0.69	0.04
	Reading	62408	0.74	0.16	0.43	0.10
	Writing	62408	0.73	0.14	0.47	0.11
	Listening and Speaking	65353	0.83	0.08	0.52	0.06
	Reading and Writing	62408	0.74	0.15	0.45	0.10
5-6	Listening	30306	0.75	0.14	0.41	0.10
	Speaking	30306	0.82	0.04	0.74	0.04
	Reading	28821	0.68	0.16	0.44	0.08
	Writing	28821	0.71	0.13	0.44	0.10
	Listening and Speaking	30306	0.78	0.10	0.54	0.08
	Reading and Writing	28821	0.70	0.15	0.44	0.09
7-8	Listening	26227	0.70	0.15	0.43	0.11
	Speaking	26227	0.81	0.03	0.76	0.04
	Reading	25039	0.65	0.18	0.44	0.09
	Writing	25039	0.71	0.13	0.46	0.10
	Listening and Speaking	26227	0.74	0.10	0.56	0.08
	Reading and Writing	25039	0.68	0.16	0.45	0.09
9-12	Listening	42792	0.74	0.12	0.41	0.08
	Speaking	42792	0.80	0.04	0.73	0.03
	Reading	40934	0.62	0.14	0.39	0.10
	Writing	40934	0.72	0.11	0.45	0.10
	Listening and Speaking	42792	0.77	0.09	0.54	0.06
	Reading and Writing	40934	0.66	0.13	0.41	0.10

Note: The data file used in the table above is the final research file. The total n-counts for Grades K-12 are 218,077 after cleaning. The statistics presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

4.2 Differential Item Functioning

DIF analyses are often used to identify potential items that need to be scrutinized for item bias. DIF statistical procedures compute the probability that one demographic group is more likely to answer an item correctly than another group, when the groups are equally able. This information is useful in reviewing items and tests for potential bias in items. However, DIF does not necessarily indicate the existence of bias. As such, the flagging of an item for DIF must be evaluated substantively for bias indication.

Since the NYSESLAT included constructed-response items that were polytomously scored, the Mantel-Haenszel odds ratio α could not be used as a DIF index for all the items in the form. Instead, a generalization of the Mantel-Haenszel (1959) procedure for ordered categories, the *Mantel Statistic* (Mantel, 1963), was used for the assessment of DIF in the mixed-format examinations. The Mantel Chi-square involves comparing the mean for two groups, conditional on a matching variable. It has one degree of freedom under the null hypothesis of no conditional association between group membership and response. For dichotomous items, the Mantel statistic reduces to the usual Mantel-Haenszel Chi-square statistic (without continuity correction). The Mantel statistic has the following mathematical formulation:

$$\text{Mantel Chi-square} = \frac{\left(\sum_K F_K - \sum_K E(F_K)\right)^2}{\sum_K \text{Var}(F_K)}, \quad (1)$$

where F_K represents the sum of scores for the focal group at the k th level of the matching variable, E represents the expected, and Var represents the variance of F_K .

$$F_K = \sum_T y_T n_{FTK}, \quad (2)$$

where y_T represents the T scores that can be obtained on the item, while n_{FTK} denotes the number of focal group members who are on the k th level of the matching variable and received an item score of y_T . The expectation of F_K under the hypothesis of no association is as follows:

$$E(F_K) = \frac{n_{F+K}}{n_{++K}} \sum_T y_T n_{+TK} \quad (3)$$

The Mantel-Haenszel and the Mantel statistic, however, offer a significance test of the presence of DIF without an indication of the direction of DIF (i.e., whether in favor of the reference or the comparison group). The statistic is less likely to indicate an association in which the pattern of association for some of the strata is in the opposite direction of the patterns displayed by other

strata. On the other hand, as a significance test, its power increases with the number of responses in the two groups of comparison.

Standardized Mean Difference (SMD) is a summary statistic to accompany the Mantel approach (Dorans and Schmitt, 1991). This statistic compares the means of the reference and focal groups, adjusting for differences in the distribution of the reference and focal group members across the values of the matching variable. Mathematically, SMD is defined as follows:

$$SMD = \sum_k p_{Fk} m_{Fk} - \sum_k p_{Rk} m_{Rk}, \quad (4)$$

where

$$p_{Fk} = \frac{n_{F+k}}{n_{F++}} \quad (5)$$

is the proportion of the focal group members who are at the k^{th} level of the matching variable and

$$m_{Fk} = \frac{1}{n_{F+k} (\sum_t y_t n_{Ftk})} \quad (6)$$

is the mean item score of the focal group members at the k^{th} level, and m_{Rk} is the analogous value for the reference group. As can be seen from the equation above, the *SMD* is the difference between the unweighted item mean of the focal group and the weighted item mean of the reference group. The weights for the reference group are applied to compensate for differences in the number of students in the reference and focal groups (within each level of ability). A negative *SMD* value implies that the focal group has a lower mean item score than the reference group, conditional on the matching variable.

For the DIF classification of constructed-response items, the *SMD* is divided by the total group item standard deviation to obtain an effect-size value for the *SMD*. This effect-size *SMD* is then examined in conjunction with the Mantel χ^2 to obtain *DIF* classifications as shown in Table 8 below. DIF is categorized as “no DIF” (AA), “moderate DIF” (BB), or “large DIF” (CC).

Table 8 - DIF Classification for Constructed-Response Items

Category	Description	Criterion
AA	No DIF	Non-significant Mantel χ^2 or Significant Mantel χ^2 and $ SMD/SD \leq .17$
BB	Moderate DIF	Significant Mantel χ^2 and $.17 < SMD/SD \leq .25$
CC	Large DIF	Significant Mantel χ^2 and $.25 < SMD/SD $

Note: SD is the total group standard deviation of the item score in its original metric.

For multiple-choice items, the Mantel-Haenszel Chi-square (M-H χ^2) is used in conjunction with the M-H odds ratio (transformed to what ETS calls the Delta Scale (D)). To calculate the delta, the odds ratio should be obtained first. The odds of a correct response (proportion passing divided by proportion failing) is P/Q (i.e., $P/[1-P]$). The odds ratio is simply the odds of a correct response of the reference group divided by the odds of a correct response of the focal group. For a given item, the odds ratio is defined as follows:

$$\alpha_{M-H} = \frac{P_r / Q_r}{P_f / Q_f} \quad (7)$$

The corresponding null hypothesis is that the odds of getting the item correct are equal for the two groups (the odds ratio is equal to 1):

$$H_0: \alpha_{M-H} = \frac{P_r / Q_r}{P_f / Q_f} = 1 \quad (8)$$

In order to make the odds ratio symmetrical around zero with its range being in the interval $-\infty$ to $+\infty$, the odds ratio is transformed into a log odds ratio, as per the following:

$$\beta_{M-H} = \ln(\alpha_{M-H}) \quad (9)$$

The simple natural logarithm transformation of this odds ratio is symmetrical about zero (zero has the interpretation of equal odds). This DIF measure is a signed index where a positive value signifies DIF in favor of the reference group, while a negative value indicates DIF in favor of the focal group. β_{M-H} is amenable to linear transformations to other interval scale metrics (Camilli & Shepard, 1994). This fact is utilized by ETS to transform β_{M-H} to their Delta Scale metric (D), via:

$$D = -2.35 \cdot \beta_{M-H} \quad (10)$$

The quantity D represents the apparent difference in the difficulty of the items in the Delta metric for the two groups whose performance is being compared. Table 9 depicts DIF classifications for multiple-choice items based on the M-H χ^2 and the item Delta Scale difference value (D):

Table 9 - DIF Classification for Multiple-Choice Items

Category	Description	Criterion
A	No DIF	Non-significant M-H χ^2 or $ D < 1.0$
B	Moderate DIF	Neither A nor C
C	Large DIF	Significant M-H χ^2 and $ D \geq 1.5$

For the NYSESLAT, DIF analysis was completed on both gender and ethnic groups. For the gender DIF analysis, female students were considered the focal group while male students were

considered the reference group. Three ethnic DIF analyses were conducted. Asian, Hispanic, and Caucasian students were considered the focal group in each analysis and the remaining ethnic groups were considered the reference group. Because of low n-counts, DIF analysis was not conducted using other ethnic groups as the focal group (e.g., African-American students). The same focal and reference groups were used in the DIF analysis of field test items, provided that enough n-counts were available for the comparison.

Appendix J of this report provides the same information for the operational items. The > sign next to the DIF category indicates that the item is in favor of the reference group, while the < sign indicates that the item is in favor of the focal group.

Summary of DIF Analyses

Table 10 provides a summary of DIF analysis of operational test items, including n-counts for each DIF analysis group and the number of items flagged by both SMD and/or Mantel for the constructed-response items and ETS Delta Scale metric (D) and M-H criteria for multiple-choice items (see details in Table 8 and Table 9). All items flagged for DIF were carefully reviewed during operational test construction. Only those items that passed the reviews were included in the operational tests. It is important to note that DIF does not indicate bias. None of the items showing moderate or large DIF in Table 10 were identified as “biased” by the content specialists.

Table 10 - Results of DIF Analyses for 2010 Operational Test Items

Grade Span	Modality	N-Counts	Total Number of DIF Items	Total Number of Moderate DIF Items	Total Number of Large DIF Items
K-1	Listening/Speaking	53,399	16	12	4
	Reading/Writing	53,795	4	3	1
2-4	Listening/Speaking	65,353	13	8	5
	Reading/Writing	62,408	12	6	6
5-6	Listening/Speaking	30,306	16	10	6
	Reading/Writing	28,821	19	10	9
7-8	Listening/Speaking	26,227	10	6	4
	Reading/Writing	25,039	18	10	8
9-12	Listening/Speaking	42,792	7	7	0
	Reading/Writing	40,934	21	14	7

Note: Detailed results can be found in Appendix J.

PART 5: RELIABILITY

5.1 Internal Consistency Reliability

The internal consistency of a test is measured by the stability of scores from one sample of content to another. Several methods can be used to estimate the internal consistency of a test. One approach is to split all test questions into two groups and then correlate student scores on the two half-tests. This is known as a split-half estimate of reliability. This method avoids the implications of any changes in the individual by administering only a single test. If scores have a high rate of correlation on the two half-tests, it can be concluded that the test questions complement one another, function well as a group, and measure similar concepts. This also suggests that measurement error is minimal.

However, in the split-half method, the decision about which questions contribute to each score of the half-test can have an impact on the resulting correlation. Therefore, Pearson uses Cronbach's coefficient alpha statistic (Cronbach, 1951) to avoid this concern about the split-half method. The coefficient alpha is the average split-half correlation based on all possible divisions of a test into two parts. The coefficient alpha can be used to estimate the internal consistency of both dichotomously- (right or wrong, 0 or 1 score values) and polytomously- (a wide range of score values) scored test items. Coefficient alpha is computed by the following formula:

$$\alpha = \frac{I}{I-1} \left(1 - \frac{\sum_i s_i^2}{S_x^2} \right) \quad (11)$$

where

I is the number of items on the test,

s_i^2 is the variance of item i , and

S_x^2 is the total test variance.

Cronbach's coefficient alpha reliability statistic was calculated and presented in Tables 10 and 11 in Section 5.5. Reliability values of close to or more than 0.90 are considered high, and the test is considered a reliable test.

5.2 Classical SEM (Based on Classical Test Theory)

Since no assessment measures ability with perfect consistency, it is useful to take into account the likely amount of measurement errors. One way to describe the inconsistency of assessment results is to assess a student on multiple occasions and note how much the scores vary. Repeatedly measuring a student can only be done hypothetically. However, if you could assess a student on multiple occasions, you would obtain a collection of the student's obtained scores. The scores would cluster around an average value. The standard deviation, or spread, of these obtained scores is known as the standard error of measurement (SEM).

The SEM is another index of reliability that provides an estimate of the amount of error in an individual's observed test score. The individual's observed total score is considered the estimate of the person's true score. Because the SEM is inversely related to the reliability of a test, the greater the reliability, the less the standard error of measurement, and the more confidence one may have in the accuracy, or precision, of the observed test score. The measurement error is commonly expressed in terms of standard deviation units; that is, the SEM is the standard deviation of the measurement error distribution. The SEM is calculated using the following equation:

$$SEM = SD\sqrt{1-r_{xx}} \Leftrightarrow s_e = s_x\sqrt{1-\frac{s_t^2}{s_x^2}} \quad (12)$$

where $SEM (=s_e)$ refers to the standard error of measurement,

$SD (=s_x)$ is the standard deviation unit of the scale for a test,

r_{xx} is the reliability coefficient for a sample test (or estimate of ρ_{xx} , which is a population reliability coefficient),

s_t^2 is the estimate of σ_T^2 , and

s_x^2 is the estimate of σ_x^2 .

The SEM s are presented in Tables 11 and 12 in Section 5.5. The smaller the SEM s (close to 0), the higher the quality of the test will be.

5.3 Standard Error of the Mean (SEMn)

The standard error of the mean is an estimate of the magnitude of sampling error associated with the sample mean in the estimation of the population mean. This expected standard mean of sampling errors of the mean is called the standard error of the mean (SEMn) and is defined as follows:

$$SEMn = \frac{\sigma}{\sqrt{n}} \quad (13)$$

where $SEMn$ is the standard error of the mean, σ is the standard deviation of the population, and n is the number of responses in each sample.

The $SEMn$'s are presented in Tables 11 and 12 in Section 5.5. The more accurate the estimation of the population mean, the smaller the $SEMn$ values will be.

5.4 Conditional SEM (Based on Item Response Theory)

Unlike the SEM based on the classical test theory, the SEM based on the item-response theory (IRT) is not the same for all students. For example, if a student gets either a few or a large number of items correct (extreme score), the standard error is greater than if the student gets a moderate number of items correct. This implies that the SEM depends on the total score (Andrich & Luo, 2004). Under the Rasch model, the SEM for each student is as follows:

$$\sigma_{\hat{\beta}} = \frac{1}{\sqrt{\sum_{i=1}^L p_{vi}(1-p_{vi})}} \quad (14)$$

where v is subscript for a person, i is subscript for an item, L is length of the test,

$\hat{\beta}$ is ability estimate, and p_{vi} is the probability that a student answers an item correctly. It is defined as follows:

$$p_{vi} = \frac{e^{\beta_v - \delta_i}}{1 + e^{\beta_v - \delta_i}} \quad (15)$$

where β_v is student v 's ability, and δ_i is the difficulty of the item.

A confidence band can be used in interpreting the ability estimate. For example, an approximate 68% confidence interval for $\hat{\beta}$ is given by:

$$\hat{\beta} \pm SEM \quad (16)$$

Note that the standard error for item difficulty is the smallest when the probability of passing is close to the probability of failing. That is, when an item is near the threshold level for many students in the sample, the standard error is small (Embretson & Reise, 2000).

According to the consensus in measurement, an aspect for the popularity of IRT methods in analyzing data is because classical statistics assume equivalency of students and item measurements for all examinees and items in a test. IRT methods allow for the differentiation of varied student and item performances in estimating the reliability of the measurement (Crocker & Algina, 1986). As evidenced by Equation 13, one reason for the fluctuation in the standard errors of students is that they are a function of the n-counts. As such, the standard errors for each of the ability score estimates are the smallest in the middle of the score distribution (where most examinees perform) and greatest for estimates in the extreme, where subsequently lower numbers of students perform on a test and thus produce less precise estimates. It is for this reason that IRT estimates with individual standard errors at score points, i.e., conditioned on theta (student ability estimates), are preferred to classical SEMs, which do not differentiate between the precision of student estimates at different levels of performances.

The conditional standard errors of measurement are presented in the raw score to scale score conversion tables in Appendix C. Appendix D shows the IRT statistics.

5.5 Summary of Descriptive and Reliability Statistics

Tables 11 and 12 provide the raw score descriptive statistics and reliabilities by grade span and grade level. The tables include the following information:

- Number of items
- Maximum number of possible points
- Number of students
- Means and standard deviations in raw scores
- Mean p -values
- Standard error of the mean (SEM_n)
- Cronbach's alpha internal consistency reliability
- Standard error of measurement (SEM)

In analyzing the information provided in Tables 11 and 12, the measures of reliability should be examined by keeping to the general trend that associates greater reliability with a lower error index for that particular measure. In Table 11, the test reliability values of Listening/Speaking and Reading/Writing across the five grade spans are all above 0.80, which is a strong indication that the test forms were constructed in good quality. The SEM_n values of Listening/Speaking and Reading/Writing across the five grade spans ranged from 0.02 to 0.08, which is very small and acceptable. The SEM values of Listening/Speaking and Reading/Writing across the five grade spans ranged from 1.44 to 2.96, which are also acceptable. Of course, if an item has, for example, a SEM_n greater than the standard deviation of the mean, then the relevancy of the information would be questioned, which might be the result of an outlier or outliers.

In Table 12, most test reliability values of Listening/Speaking and Reading/Writing across the twelve grades are above 0.80. The reliability for Grade K Reading/Writing may seem a little low in comparison to the other reliability values, but this is most likely due to the students' ages, as they just started learning reading and writing, even in their mother language. This may also be explained by the relatively erratic learning of reading and writing English in Kindergarten, which could be based on students being in a new environment, missing their homes, the level of learning to read and write at home (which may not be the same for all students), etc. This kind of reasoning seems plausible because it is the same reason that Kindergarten students are separated from other grade levels and tested on their own by other states (e.g., MI). The SEM_n and SEM values are small and within acceptable ranges at each grade level.

Table 11 - Descriptive Statistics and Reliability by Grade Span and Modality

Grade	Test	Number of Items	Max Points	N-Count	RS Mean	SD	Mean P-value	SEMn	Reliability	SEM
K-1	Listening	24	24	53,399	20.32	3.78	0.85	0.02	0.83	1.57
	Speaking	16	34	53,399	26.86	7.32	0.80	0.03	0.94	1.83
	Reading	15	15	53,795	11.52	3.33	0.77	0.01	0.83	1.39
	Writing	15	24	53,795	18.30	5.36	0.78	0.02	0.88	1.87
	Listening and Speaking	40	58	53,399	47.17	10.08	0.83	0.04	0.93	2.66
	Reading and Writing	30	39	53,795	29.82	8.23	0.78	0.04	0.92	2.38
2-4	Listening	24	24	65,353	19.65	4.09	0.82	0.02	0.84	1.64
	Speaking	16	34	65,353	28.48	6.78	0.84	0.03	0.94	1.67
	Reading	24	24	62,408	17.77	4.50	0.74	0.02	0.83	1.83
	Writing	16	22	62,408	15.75	4.13	0.73	0.02	0.80	1.83
	Listening and Speaking	40	58	65,353	48.13	9.82	0.83	0.04	0.93	2.55
	Reading and Writing	40	46	62,408	33.52	8.08	0.74	0.03	0.90	2.61
5-6	Listening	25	25	30,306	18.73	4.68	0.75	0.03	0.84	1.86
	Speaking	16	34	30,306	27.95	7.79	0.82	0.04	0.95	1.67
	Reading	27	27	28,821	18.45	5.54	0.68	0.03	0.86	2.06
	Writing	19	25	28,821	17.22	4.69	0.71	0.03	0.81	2.06
	Listening and Speaking	41	59	30,306	46.68	11.34	0.78	0.07	0.94	2.77
	Reading and Writing	46	52	28,821	35.66	9.49	0.70	0.06	0.90	2.93
7-8	Listening	25	25	26,227	17.55	5.15	0.70	0.03	0.86	1.94
	Speaking	16	34	26,227	27.47	8.18	0.81	0.05	0.96	1.64
	Reading	26	26	25,039	17.61	5.50	0.68	0.03	0.86	2.08
	Writing	19	25	25,039	17.64	4.89	0.71	0.03	0.83	2.01
	Listening and Speaking	41	59	26,227	45.02	12.32	0.74	0.08	0.95	2.81
	Reading and Writing	45	51	25,039	35.24	9.66	0.69	0.06	0.91	2.92
9-12	Listening	25	25	42,792	18.60	4.94	0.74	0.02	0.85	1.89
	Speaking	16	34	42,792	27.19	7.87	0.80	0.04	0.95	1.70
	Reading	27	27	40,934	16.84	5.27	0.62	0.03	0.82	2.23
	Writing	19	25	40,934	17.75	4.94	0.72	0.02	0.83	2.06
	Listening and Speaking	41	59	42,792	45.79	11.69	0.77	0.06	0.94	2.80
	Reading and Writing	46	52	40,934	34.59	9.36	0.66	0.05	0.89	3.07

Note: The data file used in the table above is the final research file. The total n-counts for Grades K-12 are 218,077. The statistics presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

Table 12 - Descriptive Statistics and Reliability by Grade and Modality

Grade	Test	Number of Items	Max Points	N-Count	RS Mean	SD	Mean P-value	SEMn	Reliability	SEM
K	Listening	24	24	25,927	19.06	3.98	0.79	0.02	0.80	1.77
	Speaking	16	34	25,927	25.06	7.70	0.74	0.05	0.94	1.96
	Reading	15	15	26,225	9.82	3.19	0.65	0.02	0.74	1.62
	Writing	15	24	26,225	15.74	5.33	0.68	0.03	0.86	2.01
	Listening and Speaking	40	58	25,927	44.12	10.53	0.77	0.07	0.92	2.90
	Reading and Writing	30	39	26,225	25.56	7.90	0.67	0.05	0.89	2.64
1	Listening	24	24	27,472	21.51	3.15	0.90	0.02	0.82	1.34
	Speaking	16	34	27,472	28.56	6.51	0.85	0.04	0.93	1.69
	Reading	15	15	27,934	13.11	2.60	0.87	0.02	0.82	1.11
	Writing	15	24	27,934	20.71	4.13	0.87	0.02	0.85	1.58
	Listening and Speaking	40	58	27,472	50.06	8.70	0.88	0.05	0.92	2.39
	Reading and Writing	30	39	27,934	33.82	6.31	0.87	0.04	0.90	1.96
2	Listening	24	24	23,635	18.48	4.14	0.77	0.03	0.81	1.79
	Speaking	16	34	23,635	27.80	6.68	0.82	0.04	0.93	1.75
	Reading	24	24	22,224	16.11	4.47	0.67	0.03	0.81	1.96
	Writing	16	22	22,224	14.19	4.11	0.66	0.03	0.78	1.92
	Listening and Speaking	40	58	23,635	46.28	9.56	0.79	0.06	0.92	2.72
	Reading and Writing	40	46	22,224	30.30	7.94	0.67	0.05	0.88	2.77
3	Listening	24	24	21,889	20.04	3.85	0.83	0.03	0.83	1.59
	Speaking	16	34	21,889	28.74	6.71	0.85	0.05	0.94	1.64
	Reading	24	24	21,040	18.26	4.24	0.76	0.03	0.82	1.80
	Writing	16	22	21,040	16.14	3.86	0.75	0.03	0.79	1.78
	Listening and Speaking	40	58	21,889	48.77	9.59	0.84	0.06	0.93	2.50
	Reading and Writing	40	46	21,040	34.40	7.52	0.76	0.05	0.89	2.55
4	Listening	24	24	19,829	20.61	3.95	0.86	0.03	0.86	1.47
	Speaking	16	34	19,829	29.02	6.90	0.86	0.05	0.95	1.60
	Reading	24	24	19,296	19.15	4.22	0.80	0.03	0.84	1.68
	Writing	16	22	19,296	17.11	3.83	0.80	0.03	0.80	1.72
	Listening and Speaking	40	58	19,829	49.63	10.04	0.86	0.07	0.94	2.40
	Reading and Writing	40	46	19,296	36.26	7.54	0.80	0.05	0.90	2.42

Note: The data file used in the table above is the final research file. The total n-counts for Grades K–12 are 218,077. The statistics presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

Table 12 - Descriptive Statistics and Reliability by Grade and Modality (continued)

Grade	Test	Number of Items	Max Points	N-Count	RS Mean	SD	Mean P-value	SEMn	Reliability	SEM
5	Listening	25	25	16,381	18.83	4.54	0.75	0.04	0.83	1.86
	Speaking	16	34	16,381	28.23	7.49	0.83	0.06	0.95	1.66
	Reading	27	27	15,576	18.45	5.45	0.68	0.04	0.86	2.06
	Writing	19	25	15,576	17.25	4.63	0.72	0.04	0.80	2.05
	Listening and Speaking	41	59	16,381	47.06	10.92	0.78	0.09	0.94	2.74
	Reading and Writing	46	52	15,576	35.70	9.36	0.70	0.07	0.90	2.93
6	Listening	25	25	13,925	18.61	4.85	0.74	0.04	0.85	1.87
	Speaking	16	34	13,925	27.61	8.11	0.81	0.07	0.96	1.68
	Reading	27	27	13,283	18.44	5.65	0.68	0.05	0.87	2.05
	Writing	19	25	13,283	17.17	4.75	0.71	0.04	0.81	2.06
	Listening and Speaking	41	59	13,925	46.22	11.79	0.77	0.10	0.94	2.80
	Reading and Writing	46	52	13,283	35.61	9.65	0.70	0.08	0.91	2.94
7	Listening	25	25	13,230	17.38	5.15	0.70	0.04	0.86	1.96
	Speaking	16	34	13,230	27.55	8.23	0.81	0.07	0.96	1.63
	Reading	26	26	12,607	17.19	5.48	0.66	0.05	0.85	2.11
	Writing	19	25	12,607	17.59	4.95	0.71	0.04	0.84	2.00
	Listening and Speaking	41	59	13,230	44.94	12.39	0.74	0.11	0.95	2.81
	Reading and Writing	45	51	12,607	34.78	9.70	0.68	0.09	0.91	2.94
8	Listening	25	25	12,997	17.72	5.15	0.71	0.05	0.86	1.92
	Speaking	16	34	12,997	27.38	8.12	0.81	0.07	0.96	1.66
	Reading	26	26	12,444	18.03	5.49	0.69	0.05	0.86	2.04
	Writing	19	25	12,444	17.68	4.83	0.71	0.04	0.83	2.02
	Listening and Speaking	41	59	12,997	45.09	12.24	0.75	0.11	0.95	2.80
	Reading and Writing	45	51	12,444	35.71	9.59	0.70	0.09	0.91	2.90
9	Listening	25	25	14,456	17.37	5.37	0.69	0.04	0.86	1.98
	Speaking	16	34	14,456	25.53	9.04	0.75	0.08	0.96	1.77
	Reading	27	27	13,472	15.52	5.28	0.57	0.05	0.81	2.29
	Writing	19	25	13,472	16.56	5.26	0.68	0.05	0.84	2.12
	Listening and Speaking	41	59	14,456	42.89	13.26	0.72	0.11	0.95	2.98
	Reading and Writing	46	52	13,472	32.08	9.65	0.62	0.08	0.89	3.15

Note: The data file used in the table above is the final research file. The total n-counts for Grades K–12 are 218,077. The statistics presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

Table 12 Descriptive Statistics and Reliability by Grade and Modality (continued)

Grade	Test	Number of Items	Max Points	N-Count	RS Mean	SD	Mean P-value	SEMn	Reliability	SEM
10	Listening	25	25	13,032	18.25	4.93	0.73	0.04	0.85	1.93
	Speaking	16	34	13,032	26.65	7.89	0.79	0.07	0.95	1.75
	Reading	27	27	12,523	16.42	5.22	0.61	0.05	0.81	2.25
	Writing	19	25	12,523	17.36	4.99	0.71	0.04	0.83	2.08
	Listening and Speaking	41	59	13,032	44.90	11.67	0.75	0.10	0.94	2.86
	Reading and Writing	46	52	12,523	33.78	9.35	0.65	0.08	0.89	3.10
11	Listening	25	25	9,208	19.63	4.30	0.79	0.04	0.82	1.82
	Speaking	16	34	9,208	28.42	6.59	0.84	0.07	0.94	1.64
	Reading	27	27	8,984	18.00	4.99	0.67	0.05	0.81	2.17
	Writing	19	25	8,984	18.82	4.43	0.76	0.05	0.80	1.98
	Listening and Speaking	41	59	9,208	48.05	9.69	0.81	0.10	0.93	2.65
	Reading and Writing	46	52	8,984	36.83	8.55	0.70	0.09	0.88	2.97
12	Listening	25	25	6,096	20.72	3.67	0.83	0.05	0.79	1.69
	Speaking	16	34	6,096	30.42	4.71	0.90	0.06	0.90	1.46
	Reading	27	27	6,016	18.93	4.73	0.70	0.06	0.80	2.11
	Writing	19	25	6,016	19.60	3.86	0.79	0.05	0.75	1.92
	Listening and Speaking	41	59	6,096	51.14	7.15	0.86	0.09	0.89	2.37
	Reading and Writing	46	52	6,016	38.53	7.69	0.74	0.10	0.86	2.88

Note: The data file used in the table above is the final research file. The total n-counts for Grades K–12 are 218,077. The statistics presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

5.6 Inter-Rater Reliability

As another source of measuring error results in the evaluation of student work, inter-rater reliability investigates the extent to which examinees would obtain the same score if the assessment task were scored two or more times by the same rater or different raters. One way to estimate this type of reliability is to have two raters score each student's paper and then obtain the correlation. In this case, consistency is defined as the similarity of student's rank orderings by two raters. Another way to obtain evidence of inter-rater reliability is to calculate the percent agreement between raters. If raters always agree in their assignment of scores, there is 100% agreement. If raters never agree in their assignment of scores, there is 0% agreement. The choice between using a correlation coefficient or percent agreement depends upon whether students' absolute (actual) or relative (rank order) score level is important for a particular interpretation and use.

Audit scoring of the operational test consists of only open-ended items. Personnel selected as team leaders are usually the same people who have participated in the range-finding process and are familiar with the New York State Test. Training consists of a review of the same scoring materials used by the New York State teachers for scoring the full operational test. Readers are tested by scoring sample papers prior to actual audit scoring. Once qualified, readers are check-scored by team leaders to maintain their accuracy.

To ensure that local teachers apply the same rigorous scoring standards as intended by the New York State Education Department (NYSED), and to provide evidence of inter-rater reliability, the Department requires that Pearson rescore 10% of all test papers after each test administration. The selection of the 10% audit sample for each grade span followed a stratified random selection procedure based on the Need/Resource Category, which divided the State schools into six categories: New York City, Big 4 Cities (Buffalo, Rochester, Syracuse, and Yonkers), High Need Urban/Suburbs, High Need Rural, Average Need, and Low Need.

Target values were calculated for each level using the percentage of enrollment data for each of the above classifications and was applied to the level targets for "10% Plus 40% Over Sample." Since the sample selection level was by school, enrollment data collected for the spring 2010 administration was used to select schools for the sampling and analysis, with the goal of acquiring at least 95%, but no more than 105%, of the target count for each grade span. These schools returned their booklets to Pearson for rescoring. The following sections (or tables) indicate the procedures that Pearson conducted to ensure reliability and accurate scoring of the items.

Table 13 provides the audit sample information regarding the number of students obtained for the rater analyses after merging the auditing file (Accudata) and final research file. The Accudata and the research file were merged according to student ID (SID). The n-counts in the table are based on 88% of the Accudata file from auditing, matching the research file database. The percentage of the audit sample ranged from 7% to 18% across the five grade spans.

Table 13 - Audit Sample

Grade Span	Number of Students in Audit Sample after Merging ¹	Number of Students in Population by Grade Span	Percentage of Total Population ²
K-1	5,017	56,867	9
2-4	5,152	69,095	7
5-6	5,015	32,462	15
7-8	5,133	27,855	18
9-12	5,116	44,573	11
Total	25,433	230,852	11

Note 1: There is 1 Accudata student that did not merge with the final research file and 313 Accudata students that did merge with the final research file but had missing data. Note 2: Percentage of Total Population = Number of Students in Audit Sample/Number of Students in Final Research File.

Inter-Rater Agreement

Table 14 provides the rater agreement for the Pre-Writing and Writing constructed-response items between local raters and Pearson raters. The Pearson raters scored the items independently. When the two raters assigned the same score to a student's paper, the agreement rating was denoted as *exact* (i.e., perfect agreement). Ratings that differed by exactly one score point were denoted as *adjacent*. Ratings that differed by two or more score points were denoted as *non-adjacent*. The results are presented by grade span and item. For Grades K-1, items 1-9 were scored 2. For Grades 2-12, items 1-3 were scored 0, 1, or 2, and item 4 was scored 0, 1, 2, 3, or 4. The following are the descriptions of the rater-agreement variable:

- **Exact:** 0 Score Point Difference between Local and Pearson Raters
- **Adjacent:** +/-1 Score Point Difference between Local and Pearson Raters
- **Non-Adjacent:** +/-2 Score Point Difference between Local and Pearson Raters

In Table 14, for the 2-point items, the percent of exact rating ranged from 66.24 to 97.57 across all the grade spans. The percent of adjacent rating ranged from 1.54 to 31.07. The total percent of the exact and adjacent ratings ranged from 93.12 to 99.82. The percent of exact rating is relatively low for item 4 (Grade 2-12) because it has a greater range (0-4). The percent of exact rating ranged from 53.96 to 60.24 across the four grade spans. In addition, the percent of adjacent rating ranged from 36.11 to 41.13. The total percent of the exact and adjacent ratings for the four-point items ranged from 95.09 to 96.35.

The table also provides the intraclass correlations between the local raters and Pearson raters, Weighted Kappa coefficient, the asymptotic standard error (ASE) of the Weighted Kappa coefficient, and its lower and upper 95 percent confidence limits.

Intraclass Correlation

The intraclass correlation assesses rating reliability by comparing the variability of different ratings of the same subject to the total variation across all ratings and all subjects. The intraclass correlations in Table 14 were calculated by using the random and fixed effects model introduced by Shrout and Fleiss (1979) to measure the agreement of the local and Pearson raters. The SAS program from Douglas Steinley and Philip Karl Wood (2000) was modified so that large student

samples could be used for the analyses. The intraclass correlations ranged from 0.61 to 0.91 across all the grade spans. The lowest correlations occurred at the 5–6 grade span.

Kappa Coefficient

Kappa coefficient is an index measuring the rater agreement, with the value always less than or equal to 1. A value of 1 implies perfect agreement, and a value less than 1 implies less than perfect agreement. In rare situations, Weighted Kappa can be negative. This is a sign that the two observers agreed less than would be expected just by chance. It is also rare that we get perfect agreement. Different people have different interpretations as to what is a good level of agreement. Here is an interpretation of the Weighted Kappa coefficient (Altman, D.G. Practical Statistics for Medical Research, 1991, page 404).

- Poor agreement = less than 0.20
- Fair agreement = 0.20 to 0.40
- Moderate agreement = 0.40 to 0.60
- Good agreement = 0.60 to 0.80
- Very good agreement = 0.80 to 1.00

In Table 14, the values of the Weighted Kappa across grade spans indicate that the rater agreement ranges within moderate agreement to good agreement, using the interpretation rule above. The asymptotic standard error (ASE) of the Weighted Kappa coefficient and its lower and upper 95 percent confidence limits indicate that the errors of the Weighted Kappa coefficient are very small and within limited range.

Table 14 - Rater Agreement for Pre-Writing and Writing Prompts

Grade Span	Item	Max Point	N-count	Exact	Adjacent	Non-Adjacent	Intraclass Correlation	Weighted Kappa	ASE	95% Confidence Limits	
										Lower	Upper
K-1	1	2	5,017	87.89	11.93	0.18	0.77	0.76	0.01	0.74	0.77
	2	2	5,017	87.39	12.41	0.20	0.74	0.74	0.01	0.72	0.76
	3	2	5,017	97.57	1.54	0.90	0.91	0.90	0.01	0.89	0.92
	4	2	5,017	94.73	4.39	0.88	0.90	0.89	0.01	0.87	0.90
	5	2	5,017	86.77	12.41	0.82	0.91	0.84	0.01	0.83	0.85
	6	2	5,017	88.59	10.57	0.84	0.90	0.84	0.01	0.82	0.85
	7	2	5,017	84.22	15.16	0.62	0.85	0.80	0.01	0.78	0.81
	8	2	5,017	79.15	20.15	0.70	0.83	0.76	0.01	0.75	0.77
	9	2	5,017	73.78	25.38	0.84	0.78	0.69	0.01	0.68	0.71
2-4	1	2	5,152	92.14	7.16	0.70	0.85	0.82	0.01	0.81	0.84
	2	2	5,152	90.90	8.27	0.83	0.87	0.84	0.01	0.82	0.85
	3	2	5,152	79.19	19.84	0.97	0.78	0.73	0.01	0.71	0.74
	4	4	5,152	58.54	37.33	4.13	0.74	0.60	0.01	0.59	0.62
5-6	1	2	5,015	73.02	23.81	3.17	0.61	0.56	0.01	0.54	0.58
	2	2	5,015	75.49	23.03	1.48	0.72	0.67	0.01	0.66	0.69
	3	2	5,015	66.33	26.79	6.88	0.64	0.55	0.01	0.53	0.57
	4	4	5,015	57.88	38.07	4.05	0.74	0.60	0.01	0.58	0.61
7-8	1	2	5,133	89.69	9.33	0.97	0.85	0.82	0.01	0.81	0.84
	2	2	5,133	85.93	13.52	0.55	0.81	0.79	0.01	0.77	0.80
	3	2	5,133	81.72	16.68	1.60	0.81	0.77	0.01	0.75	0.78
	4	4	5,133	60.24	36.11	3.64	0.79	0.67	0.01	0.65	0.68
9-12	1	2	5,116	77.20	21.52	1.27	0.73	0.68	0.01	0.66	0.70
	2	2	5,116	72.02	26.43	1.54	0.68	0.62	0.01	0.61	0.64
	3	2	5,116	66.24	31.07	2.70	0.68	0.58	0.01	0.57	0.60
	4	4	5,116	53.96	41.13	4.91	0.75	0.61	0.01	0.60	0.62

Note: The stats in the table are based on the merged file between Accudata (auditing) and Pearson's database (final research file). The Accudata and research file were merged by SIDs. The n-counts in the table are based on the 88% of the Accudata file from auditing to match to the Pearson's database.

Table 15 provides the frequency distribution of the score point differences between the local raters' and Pearson raters' scores by grade span. For the 2-point items, the percentage of 0 differences ranged from 66.24 to 97.57. For the 4-point items, the percentage of 0 differences ranged from 53.96 to 60.24. The above evidence indicates that the rater-agreement percentage is higher for the 2-point item than for the 4-point item.

Table 15 - Percentages of the Score Difference Between Raters

Grade Span	Item	Percentage of Point Difference Score Difference (Local score minus Pearson score)								
		-4	-3	-2	-1	0	1	2	3	4
K-1	1			0.16	5.86	87.89	6.06	0.02		
	2			0.18	6.78	87.39	5.63	0.02		
	3			0.48	0.82	97.57	0.72	0.42		
	4			0.48	1.10	94.73	3.29	0.40		
	5			0.36	5.07	86.77	7.34	0.46		
	6			0.38	3.77	88.59	6.80	0.46		
	7			0.34	9.10	84.22	6.07	0.28		
	8			0.42	11.75	79.15	8.40	0.28		
	9			0.48	14.01	73.78	11.37	0.36		
2-4	1			0.43	4.10	92.14	3.07	0.27		
	2			0.43	4.31	90.90	3.96	0.41		
	3			0.50	9.49	79.19	10.35	0.47		
	4	0.02	0.17	1.79	16.69	58.54	20.63	2.04	0.12	
5-6	1			2.79	17.73	73.02	6.08	0.38		
	2			0.54	10.95	75.49	12.08	0.94		
	3			0.20	4.09	66.33	22.70	6.68		
	4		0.06	1.68	17.87	57.88	20.20	2.21	0.10	
7-8	1			0.55	5.09	89.69	4.25	0.43		
	2			0.18	6.35	85.93	7.17	0.37		
	3			0.99	8.36	81.72	8.32	0.60		
	4	0.04	0.10	2.22	22.45	60.24	13.66	1.27		0.02
9-12	1			0.65	12.90	77.20	8.62	0.63		
	2			0.55	13.43	72.02	13.00	1.00		
	3			0.84	12.02	66.24	19.04	1.86		
	4		0.14	3.05	21.51	53.96	19.63	1.64	0.04	0.04

Table 16 provides the mean and standard deviation of each item for both the local raters and the Pearson raters. There is a column for the local raters, a column for Pearson raters, and a third column for the difference of the mean scores and ratio of standard deviations between the local and Pearson raters. The largest mean difference (0.32) occurred at grade Spans 5-6 (item 3), and SD ratio ranged from 0.94 to 1.14. For all items, the mean difference is close to 0, and SD ratio is close to 1, which is a strong indication of rater agreement.

Table 16 - Comparison Between Local and Pearson Raters

Grade Span	Item	Local		Pearson		Differences	
		Mean	SD	Mean	SD	Mean	SD Ratio
K-1	1	1.62	0.53	1.62	0.52	0.00	1.02
	2	1.63	0.52	1.65	0.51	-0.02	1.02
	3	1.81	0.57	1.81	0.57	0.00	1.00
	4	1.67	0.66	1.65	0.67	0.02	0.99
	5	1.27	0.83	1.24	0.85	0.02	0.98
	6	1.50	0.76	1.47	0.80	0.03	0.95
	7	1.39	0.80	1.43	0.82	-0.03	0.98
	8	1.16	0.84	1.19	0.85	-0.04	0.99
	9	1.00	0.78	1.03	0.83	-0.03	0.94
2-4	1	1.70	0.61	1.71	0.60	-0.01	1.02
	2	1.61	0.69	1.61	0.69	0.00	1.00
	3	1.30	0.77	1.29	0.74	0.01	1.03
	4	1.80	1.10	1.76	1.01	0.04	1.09
5-6	1	1.41	0.71	1.57	0.64	-0.16	1.11
	2	1.33	0.75	1.31	0.77	0.02	0.97
	3	1.14	0.80	0.82	0.78	0.32	1.02
	4	1.85	1.12	1.82	0.99	0.04	1.14
7-8	1	1.58	0.71	1.59	0.71	-0.01	1.00
	2	1.51	0.71	1.50	0.72	0.01	0.99
	3	1.27	0.81	1.27	0.81	-0.01	1.00
	4	2.12	1.21	2.23	1.16	-0.11	1.04
9-12	1	1.40	0.74	1.44	0.74	-0.04	1.00
	2	1.32	0.74	1.32	0.75	0.00	0.99
	3	1.10	0.80	1.01	0.80	0.09	0.99
	4	2.17	1.23	2.22	1.14	-0.05	1.08

Summary

The evidence of inter-rater reliability statistics (percentage of agreement, intra-class correlation, Weighted Kappa coefficient, percentage of score point difference, and mean difference) indicated that the local raters and Pearson raters generally have good agreement regarding the NYSESLAT scores. However, for a few items (e.g., item 4 in grade span 5-6), the agreement between local raters and Pearson raters was poor. In these cases, the local raters generally gave higher scores than those given by Pearson raters. To improve the agreement between local raters and Pearson raters in later administrations, the scoring guides may need to be improved or the teachers who score the test may need to be trained better, or both.

5.7 Reliability of Classification Decision at Proficient Cut

Based on the NYSESLAT scale scores, student performance is classified into one of four proficiency levels. While it is always important to know the reliability of student scores in any examination, it is of even greater importance to assess the reliability of the decisions based on these scores. Evaluation of the reliability of classification decisions is performed through estimation of the probabilities of correct and consistent classification of student performance. Procedures from Livingston and Lewis (1995) were applied to derive measures of the accuracy

and consistency of the classifications. The brief descriptions of the procedures used and results obtained are presented in this section.

The accuracy of decisions is the extent to which decisions would agree with those that would be made if each student could somehow be tested with all possible forms of the examination. The consistency of decisions is the extent to which decisions would agree with those that would have been made if the students had taken a parallel form of the NYSESLAT, equal in difficulty and covering the same content as the form they actually took. These ideas are shown schematically in Figures 1 and 2. Please note that the term Achieves Proficient Status refers to the proficient category on the Listening/Speaking and Reading/Writing combinations score, and the term Does Not Achieve Proficient Status refers to all categories below proficient status.

		Decision made on a form actually taken	
		Does Not Achieve Proficient Status	Achieves Proficient Status
True status made on all-forms average	Does Not Achieve Proficient Status	Correct Classification	Misclassification
	Achieves Proficient Status	Misclassification	Correct Classification

Note: Adapted from Young and Yoon (1998).

Figure 1 - Classification Accuracy

		Decision made on the 2nd form taken	
		Does Not Achieve Proficient Status	Achieves Proficient Status
Decision made on the 1st form taken	Does Not Achieve Proficient Status	Correct Classification	Misclassification
	Achieves Proficient Status	Misclassification	Correct Classification

Note: Adapted from Young and Yoon (1998).

Figure 2 - Classification Consistency

In Figure 1, accurate classifications occur when the decision made based on the all-forms average agrees with the decision made based on the form actually taken.

Misclassifications occur when, for example, a student who actually accomplished Does Not Achieve Proficient Status based on his or her all-forms average is classified incorrectly as accomplishing Achieves Proficient Status. Consistent classification occurs (Figure 2) when two forms agree on the classification of a student as either Achieves Proficient Status or Does Not Achieve Proficient Status, whereas inconsistent classification occurs when the decisions made by the forms differ.

These analyses make use of the techniques outlined and implemented by Hanson (1991), Haertel (1996), Livingston and Lewis (1995), and Young and Yoon (1998). The software developed by

Hanson (1995) was used for the analyses. Estimates of decision accuracy and consistency were made for the Achieves Proficient Status cut points on the Listening/Speaking and Reading/Writing scores reported in the NYSESLAT.

The table also includes the proportions of False Positive and False Negative classifications. The sum of values of Accuracy, False Positive, and False Negative is equal to 1.00, but due to rounding, the table values may or may not equal 1.00. False Positive and False Negative classifications refer to the mismatch between student true scores and observed scores. The False Positive value is the proportion of student scores misclassified to the category Achieves Proficient Status when student scores do not meet proficient status. The False Negative value is the proportion of student scores misclassified to the category Does Not Achieve Proficient Status when student scores meet proficient status.

Table 17 presents the results of the decision accuracy and consistency of the Achieves Proficient Status cut scores for the Listening/Speaking and Reading/Writing scores. The table contains the following:

- Accuracy
- False positives
- False negatives
- Consistency

Decision accuracy refers to the agreement between the classifications based on the form actually taken and the classifications that would be made if each student could somehow be tested with all possible forms (or true score) of the examination. For example, for K students, 89% of the classifications for Listening/Speaking at the proficiency level were accurate, while 5% percent of the students were classified as proficient when they were not. Six percent of students were classified as not proficient when they were proficient. For Listening/Speaking combination proficiency-level classification, the accuracy of the classification ranged from 78% to 93% across all grade levels. For Reading/Writing combination proficiency-level classification, the accuracy ranged from 83% to 96% across all the grade levels.

Decision consistency refers to the agreement between the classifications based on the form actually taken and the classifications that would be made based on an alternate form. For example, for K students, the chance that students were classified the same was 85% for Listening/Speaking. The Listening/Speaking combination proficiency-level classifications, the consistency ranged from 72% to 90% across all the grade levels. For the Reading/Writing combination proficiency-level classifications, the consistency ranged from 77% to 94% across all the grade levels. In most cases, decision accuracy and consistency for the proficiency level were higher for the Reading/Writing combination than those for the Listening/Speaking. The table also illustrates the general rule that decision consistency is lower than decision accuracy. This is understandable because decision consistency is dependent on a single alternate form, while accuracy relates to the agreement of the assessment classification with the classification on multiple forms, where the variation on a single form would be subsumed under the overall variation based on countless multiple forms to produce true scores. The false positive ranged

from 3% to 9% for Listening/Speaking and 3% to 9% for Reading/Writing across all the grade levels. The false negative ranged from 3% to 13% for Listening/Speaking and 1% to 8% for Reading/Writing across all the grade levels. Relatively speaking, more attention should be paid to the high false positive values because we don't like to see high percentages of students who were not proficient classified as proficient and exit the program. The NYSESLAT shows much higher percentages of accuracy and consistency in comparison to the false positive and negative percentages.

Table 17 - Classification Accuracy and Consistency by Grade

Grade	Test	Accuracy	False Positives	False Negatives	Consistency
K	Listening and Speaking	0.89	0.05	0.06	0.85
	Reading and Writing	0.96	0.03	0.01	0.94
1	Listening and Speaking	0.78	0.09	0.13	0.72
	Reading and Writing	0.88	0.05	0.07	0.82
2	Listening and Speaking	0.91	0.04	0.05	0.87
	Reading and Writing	0.93	0.04	0.03	0.90
3	Listening and Speaking	0.92	0.03	0.05	0.88
	Reading and Writing	0.87	0.07	0.06	0.82
4	Listening and Speaking	0.91	0.03	0.06	0.86
	Reading and Writing	0.83	0.09	0.08	0.77
5	Listening and Speaking	0.86	0.04	0.10	0.79
	Reading and Writing	0.89	0.06	0.05	0.85
6	Listening and Speaking	0.83	0.05	0.12	0.76
	Reading and Writing	0.90	0.05	0.05	0.86
7	Listening and Speaking	0.93	0.03	0.04	0.89
	Reading and Writing	0.90	0.05	0.05	0.86
8	Listening and Speaking	0.93	0.03	0.04	0.90
	Reading and Writing	0.89	0.05	0.06	0.84
9	Listening and Speaking	0.93	0.04	0.03	0.90
	Reading and Writing	0.93	0.04	0.03	0.90
10	Listening and Speaking	0.92	0.04	0.04	0.88
	Reading and Writing	0.93	0.04	0.03	0.90
11	Listening and Speaking	0.90	0.05	0.05	0.86
	Reading and Writing	0.92	0.05	0.03	0.88
12	Listening and Speaking	0.92	0.04	0.04	0.88
	Reading and Writing	0.89	0.06	0.05	0.85

Note: The data file used in the table above is the final research file. The total n-counts for Grades K–12 are 218,077. The stats presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score=0 or >900.) Cases with missing grades are also removed.

PART 6: VALIDITY

Assessments constructed from Pearson support the validity-related standards set forth in the Standards for Educational and Psychological Testing. Our judgments about test validity are based on the following sources of evidence of validity²:

- Test content—“an analysis of the relationship between a test’s content and the construct it is intended to measure” (p. 11)
- Internal structure—“the degree to which the relationships among test items and test components conform to the construct on which the proposed test score interpretations are made” (p. 13)
- Relationships to other variables—“analyses of the relationship of test scores to variables external to the test” (p. 13)

6.1 Content Validity

Evidence of validity based on test content is revealed by the extent to which the material on the test represents an appropriate sampling of skills, knowledge, and understanding of the domain tested. The *New York State Learning Standards for English as a Second Language* defines the skills, knowledge, and understanding expected for ELL students in New York State. There are several steps involved in aligning NYSESLAT items and test forms to these standards in order to ensure the content validity of the test.

Pearson psychometricians and ELL assessment specialists, with input from NYSED assessment and content specialists, developed a test blueprint that includes a range of item types that can measure all those standards and performance indicators from the *New York State Learning Standards for English as a Second Language* that are measurable on a standardized test. (Pearson ELL assessment specialists, working with committees of New York State ESL teachers, identified performance indicators that were designed for classroom use only, and thus could not be measured in a standardized test.)

When NYSESLAT item writers received item-writing assignments, they were instructed to use NYSESLAT item type specifications in conjunction with the *New York State Learning Standards for English as a Second Language* to guide their writing, and then to identify a particular standard and performance indicator that best matches what each item is measuring. When items were reviewed by committees of New York State teachers, an integral part of the review process was the discussion of and agreement on the standard and performance indicator assigned to each item.

² The page number in the parenthesis is the page number in the Standards for Educational and Psychological Testing, 1999.

The assigned standards and performance indicators were one of several criteria (e.g., item statistics, content diversity, gender and ethnic balance, etc.) that were then considered by Pearson ELL assessment specialists and NYSED assessment and content specialists during forms construction to ensure that forms reflected broad coverage of the *Learning Standards*. After forms construction was completed, item maps were developed for each form to show this coverage item by item. ELL assessment specialists at Pearson and content specialists at NYSED reviewed the item maps to confirm the alignment of a given form to the *New York State Learning Standards for English as a Second Language*. The item mapping provided in Table 18 and Appendices A.1–A.2 gives concrete evidence for the alignment to the *New York State Learning Standards for English as a Second Language*.

6.2 Internal Structure

An assessment procedure should not be a random collection of assessment tasks or test questions. Each task in the assessment should contribute positively to the total result. The interrelationship among the tasks on an assessment is known as the internal structure of the assessment. Typical questions that investigate the relationships among assessment parts (Nitko, 2004) include:

- Do all of the assessment tasks “work together” so that each task contributes positively toward assessing the quality of interest?
- If different parts of the assessment procedure are to provide unique information, do the results support this uniqueness?
- If different parts of the assessment procedure are to provide the same or similar information, do the results support this?

In order to investigate the answers to these questions, correlations based on raw scores were obtained from the four modalities. Tables 18 and 19 present the intercorrelation among the four modalities by grade span and grade. The evidence of internal structure of the 2010 NYSESLAT is also depicted by the point-biserial correlation coefficient and fit statistics. Appendices B.1–B.5 and D.1–D.5 provide these statistics for the 2010 NYSESLAT administration.

Table 18 - Intercorrelation Among the Modalities by Grade Span

Grade Span	Test	Correlation Coefficient			
		Listening	Speaking	Reading	Writing
K-1	Listening	1.00			
	Speaking	0.59	1.00		
	Reading	0.56	0.43	1.00	
	Writing	0.57	0.49	0.78	1.00
2-4	Listening	1.00			
	Speaking	0.54	1.00		
	Reading	0.56	0.48	1.00	
	Writing	0.55	0.52	0.75	1.00
5-6	Listening	1.00			
	Speaking	0.57	1.00		
	Reading	0.55	0.51	1.00	
	Writing	0.55	0.60	0.71	1.00
7-8	Listening	1.00			
	Speaking	0.66	1.00		
	Reading	0.66	0.52	1.00	
	Writing	0.71	0.68	0.72	1.00
9-12	Listening	1.00			
	Speaking	0.61	1.00		
	Reading	0.61	0.47	1.00	
	Writing	0.66	0.64	0.68	1.00

Note: The data file used in the table above is the final research file. The total n-counts for Grades K-12 are 218,077. The stats presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900). Additionally, cases with missing grade are also removed.

Table 19 - Intercorrelation Among the Modalities by Grade

Grade	Test	Correlation Coefficient			
		Listening	Speaking	Reading	Writing
K	Listening	1.00			
	Speaking	0.55	1.00		
	Reading	0.45	0.34	1.00	
	Writing	0.48	0.40	0.69	1.00
1	Listening	1.00			
	Speaking	0.56	1.00		
	Reading	0.54	0.43	1.00	
	Writing	0.55	0.50	0.73	1.00
2	Listening	1.00			
	Speaking	0.44	1.00		
	Reading	0.44	0.40	1.00	
	Writing	0.43	0.44	0.71	1.00

Table 19 - Intercorrelation Among the Modalities by Grade (continued)

Grade	Test	Correlation		Coefficient	
		Listening	Speaking	Reading	Writing
3	Listening	1.00			
	Speaking	0.55	1.00		
	Reading	0.55	0.48	1.00	
	Writing	0.53	0.53	0.72	1.00
4	Listening	1.00			
	Speaking	0.63	1.00		
	Reading	0.63	0.57	1.00	
	Writing	0.60	0.62	0.74	1.00
5	Listening	1.00			
	Speaking	0.56	1.00		
	Reading	0.55	0.50	1.00	
	Writing	0.54	0.58	0.71	1.00
6	Listening	1.00			
	Speaking	0.58	1.00		
	Reading	0.56	0.51	1.00	
	Writing	0.56	0.61	0.71	1.00
7	Listening	1.00			
	Speaking	0.67	1.00		
	Reading	0.66	0.52	1.00	
	Writing	0.71	0.68	0.73	1.00
8	Listening	1.00			
	Speaking	0.65	1.00		
	Reading	0.66	0.52	1.00	
	Writing	0.71	0.67	0.72	1.00
9	Listening	1.00			
	Speaking	0.64	1.00		
	Reading	0.58	0.49	1.00	
	Writing	0.65	0.69	0.67	1.00
10	Listening	1.00			
	Speaking	0.61	1.00		
	Reading	0.61	0.46	1.00	
	Writing	0.67	0.63	0.67	1.00
11	Listening	1.00			
	Speaking	0.54	1.00		
	Reading	0.60	0.42	1.00	
	Writing	0.63	0.56	0.64	1.00
12	Listening	1.00			
	Speaking	0.44	1.00		
	Reading	0.55	0.33	1.00	
	Writing	0.54	0.45	0.60	1.00

Note: The data file used in the table above is the final research file. The total n-counts for Grades K–12 are 218,077. The stats presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900). Additionally, cases with missing grade are also removed.

Observations of language proficiency assessment subtests in Table 18:

- Listening and Speaking are moderately correlated for grade spans ranging from 0.55–0.66.
- Listening and Reading are moderately correlated for grade spans ranging from 0.55–0.66.
- Listening and Writing are moderately correlated for grade spans ranging from 0.55–0.71.
- Speaking and Reading are moderately correlated for grade spans ranging from 0.43–0.52.
- Speaking and Writing are moderately correlated for grade spans ranging from 0.49–0.68.
- Reading and Writing are moderately correlated for grade spans ranging from 0.68–0.78.
- The language domain pairs of Listening and Speaking and Reading and Writing are moderately to highly correlated, while Speaking and Reading are not highly correlated.

Generally, the internal correlations based on grade spans are within expectation. The Reading and Writing subtests have the highest correlation range (from 0.68 to 0.78) across the five grade spans. Speaking and Reading subtests have the lowest correlation range (from 0.43 to 0.52). Listening and Speaking, Listening and Reading, Listening and Writing, and Speaking and Writing have correlation ranges between 0.49 and 0.71.

Observations of language proficiency assessment subtests in Table 19:

Kindergarten

- Students in this age group do not usually read or write yet, but can have Listening and Speaking skills. The expected outcome is that neither Reading nor Writing will correlate highly with Listening or Speaking.
- In Table 19, for Kindergarten students, Speaking and Listening have a correlation of 0.55, and Reading and Writing have a correlation of 0.69. Speaking and Reading have a low correlation of 0.34, Listening and Reading have a low correlation of 0.45, Speaking and Writing have a correlation of 0.40, and Listening and Writing have a correlation of 0.48.

Grades 1–8

- A trend of increase in the correlation (from 0.54 to 0.71) between Listening and Writing is observed; the lone exception is for Grade 2, which is still in the moderate range (0.43).
- A possible explanation (personal communication with Pearson content specialists) is that, in general, students during this age span experience expanding use and development in their Writing skills. At the same time, demands on the Listening skills of this age group remain static with only moderate development.

Grades 9–12

- A steady decrease in the correlation (from 0.65 to 0.54) between Listening and Writing is observed.
- A possible explanation (personal communication with Pearson content specialists) is that by high school there is an increased focus on the use of writing skills, especially an increased focus on academic content. Listening skill requirements for high school students also increase, but not at nearly as steep of a curve as for writing skills.
- Similar arguments may be made for the correlation behavior between Speaking and Writing in Grades 1–12. There is a trend of increase in the correlation (from 0.44 to 0.67) for Grades 1–8, and there is a steady decrease in the correlation from (0.69 to 0.45) for Grades 9–12.

Generally, the internal correlations based on grade levels are within expectation. The Reading and Writing subtests have the highest correlation range (from 0.60 to 0.74) across the thirteen grades. Speaking and Reading have the lowest correlation range (from 0.33 to 0.57). Listening and Speaking, Listening and Reading, Listening and Writing, and Speaking and Writing have correlation ranges between 0.40 and 0.71.

Validity of the Scoring Structure via Confirmatory Factor Analyses (CFA)

In order to assess the validity of the scoring structure (i.e., Speaking and Listening as one score and Reading and Writing as the other), a confirmatory factor analysis was undertaken with the 2010 full population data in order to compare the two-factor modality combinations with the unidimensional model (i.e., the total test providing a single score).

The 2010 NYSESLAT scoring dimensionality tests produced results that substantiate the use of the two-factor model in reporting IRT scores. With the exception of the 7–8 and 9–12 grade spans, the two-factor model provided a similar fit to the data as the one-factor unidimensional model as measured by global indices of fit. The unidimensional scoring model for the 7–8 and 9–12 grade spans, however, did not outweigh the use of the two-factor model in terms of model-fit and, as such, there was no clear preference of one model over the other for these two grade spans. This pattern, however, was justifiable by content experts as a distinct demarcation of modalities for younger students for whom a marked differentiation did exist between their Speaking and Listening ability as compared to their Reading and Writing ability. This differentiation, according to content experts, blurs for students in the higher grades. (See Tables 20–24 below; Appendix F provides the 2010 CFA analyses in their entirety.)

Table 20 - Global Fit Indices for the One- and Two-Factor Models (Grades K–1)

Fit Indices	One-Factor Model	Two-Factor Model
Chi-square	7617.130	512.309
GFI	0.933	0.995
AGFI	0.663	0.976
RMR	0.068	0.028
RMSEA	0.269	0.070

Note: χ^2 with 2 *df*, *p*=0.00 for the one-factor model and χ^2 with 1 *df*, *p*=0.00 for the two-factor model.

Table 21 - Global Fit Indices for the One- and Two-Factor Models (Grades 2–4)

Fit Indices	One-Factor Model	Two-Factor Model
Chi-square	4309.044	3263.943
GFI	0.966	0.974
AGFI	0.832	0.872
RMR	0.045	0.085
RMSEA	0.186	0.162

Note: χ^2 with 2 *df*, *p*=0.00 for the one-factor model and χ^2 with 1 *df*, *p*=0.00 for the two-factor model.

Table 22 - Global Fit Indices for the One- and Two-Factor Models (Grades 5–6)

Fit Indices	One-Factor Model	Two-Factor Model
Chi-square	1884.852	1820.544
GFI	0.968	0.969
AGFI	0.840	0.845
RMR	0.038	0.094
RMSEA	0.182	0.178

Note: χ^2 with 2 *df*, *p*=0.00 for the one-factor model and χ^2 with 1 *df*, *p*=0.00 for the two-factor model.

Table 23 - Global Fit Indices for the One- and Two-Factor Models (Grades 7–8)

Fit Indices	One-Factor Model	Two-Factor Model
Chi-square	1510.399	998.103
GFI	0.970	0.980
AGFI	0.852	0.901
RMR	0.029	0.026
RMSEA	0.174	0.142

Note: χ^2 with 2 *df*, *p*=0.00 for the one-factor model and χ^2 with 1 *df*, *p*=0.00 for the two-factor model.

Table 24 - Global Fit Indices for the One- and Two-Factor Models (Grades 9–12)

Fit Indices	One-Factor Model	Two-Factor Model
Chi-square	2082.209	1448.785
GFI	0.975	0.983
AGFI	0.875	0.913
SRMR	0.030	0.024
RMSEA	0.160	0.133

Note: χ^2 with 2 *df*, *p*=0.00 for the one-factor model and χ^2 with 1 *df*, *p*=0.00 for the two-factor model.

Overall, there seemed to be a decrease in the precision of model-fit for the two-factor model as one progressed from the lower grade spans to the higher ones.

6.3 External Structure

In order to ascertain the external structure of the NYSESLAT vis-à-vis an external criterion, data from two other tests were used. For ELL students in Grades 3–8, data from those who took both the NYSESLAT in May 2010, and the New York State English Language Assessment (ELA) in January 2010 were analyzed. In New York State, all ELL students, except for the first year ESL students, are required to take the ELA test. For ELL students in Grades 10–12, data from those who took both the NYSESLAT in May 2009 and the New York State Regents Examination in Comprehensive English (Regents English exam) in June 2009 were analyzed. The Regents English exam is an end-of-course exam. All high school students, including ELL students, are required to pass the Regents English exam for the high school graduation requirement. Students can take the exam in any grade, although most students take the exam in Grade 11.

The purpose of the NYSESLAT was to measure annual student progress in achieving English language proficiency in order to exit an ESL or bilingual program. It is also expected that ELL students who score proficient on the NYSESLAT will be able to function successfully in mainstream education programs and have the same chance as their English proficient counterparts to pass the required state examinations. Therefore, it is hypothesized that a relationship exists between the NYSESLAT and the ELA test/Regents English exam—that is, those who perform well on the NYSESLAT are generally expected to perform well on the ELA test or the Regents English exam.

6.3.1 Relationship with the New York English Language Arts Assessment

Correlations and Other Descriptive Statistics

Table 25 gives the sample size, minimum and maximum observed scores, means, standard deviations, and correlation between the New York ELA test and NYSESLAT raw score total of Reading, Writing, and Listening for each grade from Grade 3 to Grade 8. Table 26 gives the sample size, minimum and maximum observed scores, means, standard deviations, and correlation between the New York ELA test and NYSESLAT combined scale scores of Listening and Speaking, and Reading and Writing, for each grade from Grade 3 to Grade 8.

The correlations between the New York ELA test and NYSESLAT raw score total of the three modalities (Reading, Writing, and Listening) ranged from 0.76 to 0.79 across the six grades. The correlations are positive between the two tests, which are within expectation because the modalities tested are Reading, Writing, and Listening for the New York ELA test, and the total raw scores include those three modalities. The correlations between New York ELA test and NYSESLAT combined scale scores of the four modalities ranged from 0.64 to 0.68 across the six grades. The correlations in Table 26 are positive, but slightly lower than the correlations in Table 25, which is within expectation because the NYSESLAT total combined scale scores include the Speaking modality, whereas the New York ELA test does not include the Speaking part.

Generally speaking, the correlation analyses provide evidence that the external structure check of the NYSESLAT with respect to the New York State ELA test ascertains a positive relationship (moderate) between the two examinations (see Table 25 and Table 26). This indicates that higher scores on the NYSESLAT are associated with higher scores on the ELA test for each grade.

Table 25 - Descriptive Statistics of the NYSESLAT Raw Scores of Listening, Speaking, Reading, and Writing Modalities and Their Correlation with the ELA Test

Total Scale Scores.								
Grade	N	Test	Max Points	Mean	SD	Sample Min.	Sample Max.	Correlation with ELA
03	18145	Listening	24	20.84	2.73	1	24	0.60
03	18145	Speaking	34	30.49	4.13	0	34	0.44
03	18145	Reading	24	18.56	4.07	0	24	0.71
03	18145	Writing	22	16.39	3.80	0	22	0.70
03	18145	LI/SP	58	51.33	5.93	2	58	0.59
03	18145	RD/WR	46	34.95	7.31	3	46	0.78
03	18145	LI/RD/WR	70	55.79	9.24	9	70	0.79
04	16108	Listening	24	21.58	2.43	3	24	0.55
04	16108	Speaking	34	31.02	4.03	0	34	0.43
04	16108	Reading	24	19.75	3.66	2	24	0.71
04	16108	Writing	22	17.64	3.42	0	22	0.69
04	16108	LI/SP	58	52.60	5.63	4	58	0.55
04	16108	RD/WR	46	37.39	6.55	5	46	0.76
04	16108	LI/RD/WR	70	58.97	8.25	10	70	0.77
05	12918	Listening	25	19.92	3.19	2	25	0.57
05	12918	Speaking	34	30.50	4.58	0	34	0.43
05	12918	Reading	27	18.94	5.05	0	27	0.71
05	12918	Writing	25	17.67	4.39	0	25	0.66
05	12918	LI/SP	59	50.42	6.74	3	59	0.56
05	12918	RD/WR	52	36.61	8.67	2	52	0.75
05	12918	LI/RD/WR	77	56.53	10.96	10	77	0.76
06	10509	Listening	25	19.92	3.34	1	25	0.58
06	10509	Speaking	34	30.14	5.30	0	34	0.44
06	10509	Reading	27	18.96	5.21	0	27	0.74
06	10509	Writing	25	17.72	4.46	0	25	0.67
06	10509	LI/SP	59	50.06	7.55	1	59	0.57
06	10509	RD/WR	52	36.68	8.86	0	52	0.77
06	10509	LI/RD/WR	77	56.60	11.26	1	77	0.78
07	9950	Listening	25	18.73	4.01	1	25	0.63
07	9950	Speaking	34	30.10	5.71	0	34	0.44
07	9950	Reading	26	17.79	5.19	1	26	0.72
07	9950	Writing	25	18.32	4.60	0	25	0.67
07	9950	LI/SP	59	48.82	8.61	4	59	0.58
07	9950	RD/WR	51	36.11	9.00	1	51	0.75
07	9950	LI/RD/WR	76	54.84	12.14	8	76	0.77
08	9776	Listening	25	18.95	4.16	1	25	0.64
08	9776	Speaking	34	29.72	6.02	0	34	0.48
08	9776	Reading	26	18.60	5.23	0	26	0.70
08	9776	Writing	25	18.35	4.55	0	25	0.69
08	9776	LI/SP	59	48.67	9.10	3	59	0.61
08	9776	RD/WR	51	36.95	8.98	2	51	0.76
08	9776	LI/RD/WR	76	55.90	12.34	12	76	0.77
2-4	34253	Listening	24	21.19	2.62	1	24	0.78
2-4	34253	Speaking	34	30.74	4.09	0	34	0.44
2-4	34253	Reading	24	19.12	3.93	0	24	0.73
2-4	34253	Writing	22	16.98	3.68	0	22	0.70
2-4	34253	LI/SP	58	51.93	5.82	2	58	0.57
2-4	34253	RD/WR	46	36.09	7.07	3	46	0.77
2-4	34253	LIRDWR	70	57.28	8.93	9	70	0.78

Table 26 - Descriptive Statistics of the NYSESLAT Raw Scores of Listening, Speaking, Reading, and Writing Modalities and Their Correlation with the ELA Test Total Scale Scores (Continued).

Grade	N	Test	Max Points	Mean	SD	Sample Min.	Sample Max.	Correlation with ELA
5-6	23427	Listening	25	19.92	3.26	1	25	0.76
5-6	23427	Speaking	34	30.34	4.92	0	34	0.43
5-6	23427	Reading	27	18.95	5.12	0	27	0.72
5-6	23427	Writing	25	17.69	4.42	0	25	0.66
5-6	23427	LI/SP	59	50.26	7.12	1	59	0.55
5-6	23427	RD/WR	52	36.64	8.76	0	52	0.75
5-6	23427	LIRDWR	77	56.56	11.09	1	77	0.76
7-8	19726	Listening	25	18.84	4.09	1	25	0.77
7-8	19726	Speaking	34	29.91	5.87	0	34	0.46
7-8	19726	Reading	26	18.19	5.23	0	26	0.71
7-8	19726	Writing	25	18.33	4.57	0	25	0.68
7-8	19726	LI/SP	59	48.75	8.86	3	59	0.60
7-8	19726	RD/WR	51	36.52	9.00	1	51	0.75
7-8	19726	LIRDWR	76	55.36	12.25	8	76	0.77

Table 27 - Descriptive Statistics of the New York ELA Test and NYSESLAT Combined Scale Scores of Listening/Speaking and Reading/Writing

Grade	N	Test	Mean	SD	Sample Min.	Sample Max.	Correlation
3	18145	NYSESLAT Scale Score	1342.78	75.53	983	1610	0.60
		NY ELA Scale Score	649.76	24.49	475	780	
4	16108	NYSESLAT Scale Score	1372.94	79.75	996	1610	0.68
		NY ELA Scale Score	648.76	27.08	430	775	
5	12918	NYSESLAT Scale Score	1374.05	66.82	1083	1616	0.64
		NY ELA Scale Score	650.69	20.85	495	795	
6	10509	NYSESLAT Scale Score	1374.18	70.69	879	1666	0.67
		NY ELA Scale Score	641.52	16.74	480	785	
7	9950	NYSESLAT Scale Score	1383.81	74.12	1028	1677	0.64
		NY ELA Scale Score	638.54	20.74	470	790	
8	9776	NYSESLAT Scale Score	1388.78	78.04	1100	1677	0.68
		NY ELA Scale Score	623.88	24.11	430	790	

Proficiency Level

In order to know the classification distribution at each proficiency level of both ELA and NYSESLAT tests, the percentage distributions of the ELA proficiency level by the NYSESLAT proficiency level at each grade level are presented in Tables 27–32.

Table 27 shows that for those Grade 3 students who were classified as beginning students on the NYSESLAT, 100% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as intermediate students on the NYSESLAT, 97% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as advanced students on the NYSESLAT test, 77% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 23 % of them were classified as proficient on the ELA test. For those students who were classified as proficient students on the NYSESLAT, 44% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 56% of them were classified as proficient on the ELA test.

Table 28 shows that for those Grade 4 students who were classified as beginning students on the NYSESLAT, 99% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as intermediate students on the NYSESLAT, 98% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as advanced students on the NYSESLAT test, 84% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 16% of them were classified as proficient on the ELA test. For those students who were classified as proficient students on the NYSESLAT, 54% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 46% of them were classified as proficient on the ELA test.

Table 29 shows that for those Grade 5 students who were classified as beginning students on the NYSESLAT, 100% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as intermediate students on the NYSESLAT, 99% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as advanced students on the NYSESLAT test, 90% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 10% of them were classified as proficient on the ELA test. For those students who were classified as proficient students on the NYSESLAT, 60% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 40% of them were classified as proficient on the ELA test.

Table 30 shows that for those Grade 6 students who were classified as beginning students on the NYSESLAT, 100% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as intermediate students on the NYSESLAT, 100% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as advanced students on the NYSESLAT test, 96% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 4% of them were classified as proficient on the ELA test. For those students who were classified as proficient students on the NYSESLAT, 78% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 22% of them were classified as proficient on the ELA test.

Table 31 shows that for those Grade 7 students who were classified as beginning students on the NYSESLAT, 100% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as intermediate students on the NYSESLAT, 100% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as advanced students on the NYSESLAT test, 95% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 5% of them were classified as proficient on the ELA test. For those students who were classified as proficient students on the NYSESLAT, 80% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 20% of them were classified as proficient on the ELA test.

Table 32 shows that for those Grade 8 students who were classified as beginning students on the NYSESLAT, 100% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as intermediate students on the NYSESLAT, 100% of them were also classified as not proficient (Level 1 and Level 2) on the ELA test. For those students who were classified as advanced students on the NYSESLAT test, 98% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 2% of them were classified as proficient on the ELA test. For those students who were classified as proficient students on the NYSESLAT, 86% of them were classified as not proficient (Level 1 and Level 2) on the ELA test, and 14% of them were classified as proficient on the ELA test.

About 98%–100% of students who were classified as beginning and intermediate on the NYSESLAT test were also classified as not proficient (Level 1 and Level 2) by the ELA cut scores across the six grade levels. About 77%–98% of students who were classified as advanced on the NYSESLAT test were also classified as not proficient (Level 1 and Level 2) by the ELA cut scores across the six grade levels. About 0%–2% of students who were classified as advanced on the NYSESLAT test were classified as proficient by the ELA cut scores across the six grade levels. The higher the grade level, the lower the percentage. About 44%–86% of students who were classified as proficient on the NYSESLAT test were classified as not proficient by the ELA cut scores across the six grade levels. About 14%–40% of students who were classified as proficient on the NYSESLAT test were also classified as proficient by the ELA cut scores across the six grade levels. The higher the grade level, the lower the percentage.

Table 28 - Grade 3: Percentage of English Language Learners Scoring at Each ELA Levels at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
	Beginning (N=1070)	Intermediate (N=4039)	Advanced (N=9230)	Proficient (N=3806)
Level 1 on ELA	97%	74%	21%	3%
Level 2 on ELA	3%	23%	56%	41%
Level 3 on ELA	0%	3%	21%	45%
Level 4 on ELA	0%	0%	2%	11%
Total	100%	100%	100%	100%

Table 29 - Grade 4: Percentage of English Language Learners Scoring at Each ELA Levels at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
	Beginning (N=828)	Intermediate (N=3628)	Advanced (N=7419)	Proficient (N=4233)
Level 1 on ELA	95%	60%	13%	2%
Level 2 on ELA	4%	38%	71%	52%
Level 3 on ELA	0%	2%	16%	45%
Level 4 on ELA	0%	0%	0%	1%
Total	100%	100%	100%	100%

Table 30 - Grade 5: Percentage of English Language Learners Scoring at Each ELA Levels at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
	Beginning (N=698)	Intermediate (N=2551)	Advanced (N=6664)	Proficient (N=3005)
Level 1 on ELA	96%	80%	30%	4%
Level 2 on ELA	4%	19%	60%	56%
Level 3 on ELA	0%	1%	10%	37%
Level 4 on ELA	0%	0%	0%	3%
Total	100%	100%	100%	100%

Table 31 - Grade 6: Percentage of English Language Learners Scoring at Each ELA Levels at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
	Beginning (N=714)	Intermediate (N=2610)	Advanced (N=5241)	Proficient (N=1944)
Level 1 on ELA	97%	88%	43%	9%
Level 2 on ELA	3%	12%	53%	69%
Level 3 on ELA	0%	0%	4%	22%
Level 4 on ELA	0%	0%	0%	0%
Total	100%	100%	100%	100%

Table 32 - Grade 7: Percentage of English Language Learners Scoring at Each ELA Levels at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
	Beginning (N=1400)	Intermediate (N=3714)	Advanced (N=3316)	Proficient (N=1520)
Level 1 on ELA	96%	72%	27%	7%
Level 2 on ELA	4%	28%	68%	73%
Level 3 on ELA	0%	0%	5%	19%
Level 4 on ELA	0%	0%	0%	1%
Total	100%	100%	100%	100%

Table 33 - Grade 8: Percentage of English Language Learners Scoring at Each ELA Levels at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
	Beginning (N=1450)	Intermediate (N=3578)	Advanced (N=2832)	Proficient (N=1916)
Level 1 on ELA	96%	73%	29%	10%
Level 2 on ELA	4%	27%	69%	76%
Level 3 on ELA	0%	0%	2%	14%
Level 4 on ELA	0%	0%	0%	0%
Total	100%	100%	100%	100%

Summary

The correlation analyses provide evidence of a positive relationship between the NYSESLAT and the ELA test (ranging from 0.60 to 0.68). This indicates that higher scores on the NYSESLAT are associated with higher scores on the ELA test. The ELA and the NYSESLAT tests were developed to serve different populations of students. Thus, it is not surprising that we see only a moderate, rather than a high, positive correlation between the two tests. The analyses also confirm the validity of the NYSESLAT cut score in that about 86% to 97% (across six grade levels) of those who scored at the proficiency level on the NYSESLAT test were also classified as Level 2 or Level 3 (proficient) by the ELA test. On the other hand, the high percentages of students who were classified as not proficient by both tests is a strong indication that students who cannot do well on one test also did not perform well on the other test.

6.3.2 Relationship with the New York State Regents Examination in Comprehensive English**Correlations and Other Descriptive Statistics**

Table 33 shows the sample size, maximum score points, minimum and maximum observed scores, means, standard deviations, and correlation between the New York Regents English exam and each of the NYSESLAT modalities raw score, and the total raw scores of Reading, Writing, and Listening for Grades 10 through 12. Table 34 presents the sample size, minimum and maximum observed scores, means, standard deviations, and correlation between the New York Regents English exam and the combined scale scores of Listening/Speaking and

Reading/Writing modalities, and total scale scores of the four modalities (Listening/Speaking/Reading/Writing) for Grades 10 through 12.

Table 34 - Descriptive Statistics of the NYSESLAT Raw Scores of Listening, Speaking, Reading, and Writing Modalities and Their Correlations With the Regents English Exam Total Scale Scores

Grade	N	Test	Max Points	Mean	SD	Sample Min.	Sample Max.	Correlation with Regents English
10	2347	Listening	25	20.07	3.92	2	25	0.46
10	2347	Speaking	34	29.47	5.62	4	34	0.39
10	2347	Reading	27	17.69	4.86	2	27	0.55
10	2347	Writing	25	18.85	4.22	1	25	0.55
10	2347	LI/RD/WR Total	77	56.62	11.12	13	77	0.61
11	6549	Listening	25	20.33	3.69	3	25	0.50
11	6549	Speaking	34	29.56	5.48	1	34	0.39
11	6549	Reading	27	18.53	4.73	1	27	0.57
11	6549	Writing	25	19.36	4.02	1	25	0.59
11	6549	LI/RD/WR Total	77	58.22	10.64	17	77	0.65
12	3945	Listening	25	20.46	3.51	2	25	0.40
12	3945	Speaking	34	30.09	4.71	4	34	0.29
12	3945	Reading	27	18.52	4.57	2	27	0.47
12	3945	Writing	25	19.15	3.78	2	25	0.47
12	3945	LI/RD/WR Total	77	58.13	9.87	21	77	0.54
10-12	13211	Listening	25	20.32	3.68	2	25	0.46
10-12	13211	Speaking	34	29.71	5.32	1	34	0.37
10-12	13211	Reading	27	18.35	4.72	1	27	0.54
10-12	13211	Writing	25	19.19	4.00	1	25	0.54
10-12	13211	LI/RD/WR Total	77	57.86	10.54	13	77	0.61

Table 35 - Descriptive Statistics of the NYSESLAT Scale Scores of Listening/Speaking, Reading/Writing Modalities, and Total Scale Scores and Their Correlations With the Regents English Exam Total Scale Scores

Grade	N	Test	Mean	SD	Sample Min.	Sample Max.	Correlation with Regents English
		Regents English	50.17	21.48	0	98	–
10	2347	NYSESLAT Listening and Speaking	723.97	48.28	588	847	0.45
		NYSESLAT Reading and Writing	703.56	34.00	589	862	0.60
		Total NYSESLAT	1428.00	74.56	1202	1709	0.56

Table 34 - Descriptive Statistics of the NYSESLAT Scale Scores of Listening/Speaking, Reading/Writing Modalities, and Total Scale Scores and Their Correlations With the Regents English Exam Total Scale Scores (continued)

Grade	N	Test	Mean	SD	Sample Min.	Sample Max.	Correlation with Regents English
11	6594	Regents English	57.80	19.52	0	97	–
		NYSESLAT Listening and Speaking	725.61	47.73	553	847	0.46
		NYSESLAT Reading and Writing	709.44	34.56	601	862	0.62
		Total NYSESLAT	1435.0	74.06	120	170	0.58
12	3945	Regents English	63.24	15.12	0	98	–
		NYSESLAT Listening and Speaking	727.38	44.17	603	847	0.36
		NYSESLAT Reading and Writing	707.62	31.72	601	862	0.52
		Total NYSESLAT	1435.0	67.01	1237	1709	0.48
10–12	13211	Regents English	57.65	19.46	0	98	–
		NYSESLAT Listening and Speaking	725.88	46.83	553	847	0.42
		NYSESLAT Reading and Writing	707.64	33.68	589	862	0.58
		Total NYSESLAT	1433.5	72.18	1202	1709	0.54

The correlations between the scale scores of the Regents English exam and the NYSESLAT raw score total of the three modalities (Reading, Writing, and Listening) ranged from 0.54 to 0.65 across the three grade levels, which is higher than the correlations between the scale scores of the Regents English exam and any modality of the NYSESLAT raw scores. Additionally, it is within expectation that the correlations between the scale scores of the Regents English exam and the NYSESLAT combined raw scores of Reading/Writing are higher than the correlations between the scale scores of the Regents English exam and the NYSESLAT combined raw scores of Listening/Speaking.

The correlations between the scale scores of the Regents English exam and the NYSESLAT combined scale scores of the four modalities ranged from 0.48 to 0.58 across the three grade levels. The correlations in Table 34 are positive, but slightly lower than the correlations in Table 33, which are within expectation because the NYSESLAT total combined scale scores include the Speaking modality, whereas the Regents English exam does not include the Speaking component. Additionally, it is within expectation that the correlations between the scale scores of the Regents English exam and the NYSESLAT combined scale scores of Reading/Writing are higher than the correlations between the scale scores of the Regents English exam and the NYSESLAT combined scale scores of Listening/Speaking.

Generally speaking, the correlation analyses provide evidence that the external structure check of the NYSESLAT, with respect to the Regents English exam, ascertains a positive relationship (moderate) between the two examinations for Grades 10–12.

Proficiency Level Classification

The classification distribution in relation to passing/failing the Regents English exam by NYSESLAT proficiency level was also examined. The results are presented in Tables 35–37.

Table 35 shows that 71% of the Grade 10 students who scored proficient on the NYSESLAT passed the Regents English exam. The percentage of students who passed the Regents English exam decreased to 45% for students at the advanced level, and 22% for students at the intermediate and beginning levels.

Table 36 - Grade 10: Percentage of ELL Students Passing/Failing the Regents English Exam at Each NYSESLAT Proficiency Level

Regents English Exam	NYSESLAT Proficiency Levels			
	Beginning (N=115)	Intermediate (N=1090)	Advanced (N=714)	Proficient (N=428)
Pass	4%	18%	45%	71%
Fail	96%	82%	55%	29%
Total	100%	100%	100%	100%

Table 36 shows that of the Grade 11 students who were classified as proficient on the NYSESLAT, 86% passed the Regents English exam. For those students who were classified as advanced students on the NYSESLAT, 68% passed. For those students who were classified at the beginning and intermediate levels on the NYSESLAT, the percentage of students passing the Regents English were 7% and 33% respectively.

Table 37 - Grade 11: Percentage of ELL Students Passing/Failing the Regents English Exam at Each NYSESLAT Proficiency Level

Regents English Exam	NYSESLAT Proficiency Levels			
	Beginning (N=295)	Intermediate (N=3269)	Advanced (N=2071)	Proficient (N=959)
Pass	7%	33%	68%	86%
Fail	93%	67%	32%	14%
Total	100%	100%	100%	100%

Table 37 shows that of the Grade 12 students who were classified as proficient on the NYSESLAT, 91% passed the Regents English exam. For those students who were classified as advanced students on the NYSESLAT, 80% passed. For those students who were classified at the beginning and intermediate levels on the NYSESLAT, the percentage of students passing the Regents English were 30% and 59% respectively.

Table 38 - Grade 12: Percentage of ELL Students Passing/Failing the Regents English Exam at Each NYSESLAT Proficiency Level

Regents English Exam	NYSESLAT Proficiency Levels			
	Beginning (N=191)	Intermediate (N=2275)	Advanced (N=1039)	Proficient (N=440)
Pass	30%	59%	80%	91%
Fail	70%	41%	20%	9%
Total	100%	100%	100%	100%

Summary

The correlation analyses provide evidence of a moderately positive relationship between the NYSESLAT and the Regents English exam. The evidence suggests that the higher NYSESLAT scores are associated with higher scores on the Regents English exam required for high school graduation.

The analyses also confirm the validity of the NYSESLAT cut score, in that ELL students who were classified as proficient on the NYSESLAT had a higher chance of passing the Regents English exam than ELL students who were below the proficient level on the NYSESLAT.

PART 7: CALIBRATION, EQUATING, AND SCALING

This part introduces the IRT models and the processes that were used to calibrate, equate, and scale the NYSESLAT using these IRT models. The Rasch Model (Rasch, 1960) and the Partial Credit Model (PCM) (Masters, 1982) were used to calibrate, equate, and scale the NYSESLAT. The Rasch Model was used for dichotomous items, and the PCM for polytomous items. These measurement models are regularly used to construct test forms for scaling and equating and to develop and maintain large item banks. All item and test analyses, including item-fit analysis, scaling, equating, diagnosis, and performance prediction, were accomplished within this framework. The statistical software used to calibrate and scale the NYSESLAT was *Winsteps* Version 3.27 (Linacre & Wright, 2000).

7.1 Item Response Model and Rational for Use

In addition to reporting raw score summary statistics and item-level statistics using the classical test theory (CTT), the items on the NYSESLAT test were also analyzed within the framework of Item Response Theory (IRT). The Rasch Model for dichotomous items and the PCM for polytomous items were used for developing, scoring, and reporting the NYSESLAT assessment. These models were recommended for several reasons.

First, the NYSESLAT vertical scale was created based on the pre-existing Stanford English Language Proficiency (SELP) vertical scale that was developed using the Rasch Model. By using SELP items with known Rasch item difficulties, Pearson was able to create the NYSESLAT vertical scale in a timely fashion.

Second, the sample size requirements for calibration, scaling, and equating under the Rasch Model and PCM are significantly smaller than they are for other IRT models. For example, the Rasch Model requires the order of 400 examinees per form for equating versus approximately 1,500 examinees per form under the 3PL IRT model (Kolen and Brennan, 2004, p. 288).

Finally, for the requirements of the NYSESLAT program, the Rasch Model has one characteristic that makes it very useful: There exists a one-to-one relationship between raw scores and scale scores. That is, a student who answers a certain number of items correctly will receive the same scale score as a second student with the same raw score, regardless of which particular items within the test form were answered correctly. These reasons lead Pearson to recommend that the Rasch model be adopted as the IRT methodology for the NYSESLAT.

7.2 Evidence of Model Fit

Fit statistics are used for evaluating the goodness-of-fit of a model to the data. Fit statistics are calculated by comparing the observed and expected trace lines obtained for an item after parameter estimates are obtained using a particular model. *Winsteps* provides two kinds of fit statistics called mean-squares that show the size of the randomness or amount of distortion of the measurement system.

The OUTFIT and the INFIT statistics are used in order to ascertain the suitability of the data for constructing variables and making measures with the Rasch model. These fit statistics are mean square standardized residuals for item by person responses averaged over persons and partitioned between ability groups (OUTFIT) and within ability groups (INFIT). When the observed item characteristic curve (ICC) departs from the expected ICC from a reference value of 1, there is an expectation of high ability students failing on an easy item or low ability students succeeding on a difficult one. The OUTFIT mean square evaluates the agreement between the observed ICC and the best fitting Rasch model curve over the ability sub-groups. It is a standardized outlier-sensitive mean-square fit statistic; more sensitive to unexpected behavior by persons on items far from the person's ability level. The INFIT, on the other hand, is a within-group mean square, which summarizes the degree of misfit remaining within ability groups after the between-group misfit has been removed from the total. The INFIT, therefore, is a standardized information-weighted mean square statistic, which is more sensitive to unexpected responses to items near the person's ability level. OUTFIT mean-squares are influenced by outliers and are usually easy to diagnose and remedy. INFIT mean-squares, on the other hand, are influenced by response patterns and are harder to diagnose and remedy. In general, mean-squares near 1.0 indicate little distortion of the measurement system, while values less than 1.0 indicate that observations are too predictable (redundancy, model overfit). Values greater than 1.0 indicate unpredictability (unmodeled noise, model underfit).

When item-fit indices are lower than 0.6, they do not discriminate well and show a greater than expected degree of consistency. Similarly, a fit value higher than 1.5 indicates inconsistency in examinee scores on the item (e.g., some unexpectedly high scores for low-ability candidates and low scores for high-ability candidates).

To an extent, no data fits the IRT model perfectly. Items that degraded or distorted the measurement system were not included as anchor items and were excluded from the anchor pool. For NYSESLAT, no items that had an INFIT/OUTFIT value greater than or equal to 2.0 were included as anchor items. The percentage of items that were flagged for INFIT and OUTFIT varied depending on grade spans and subtests. In general, around 0–2% of the items were flagged for INFIT. However, the percent of misfit was greater for OUTFIT across grade spans.

The OUTFIT and the INFIT statistics are presented in the item statistics tables in Appendices D.1–D.5.

7.3 Description of Calibration Sample, Process, and Results

Calibration Sample

An equating sample (near 100% of the population) that was used for the equating analyses represented the ELL population in all six need/resource categories: NYC, Large City, and High Need Urban/Suburban, Rural, Average Need, and Low Need districts.

Calibration Process

The separate scales, one for Speaking/Listening and one for Reading/Writing, were then obtained by taking the item parameters for those separate combinations and using them to create raw score

to scale score tables. Finally, when these calibrations and scales were completed, the items tested for the 2010 administration were calibrated to the 2009 vertical scale.

A more detailed outline of the procedure follows:

- The *Winsteps* software program was used to conduct the item calibration, fixing the item parameters for the anchor set items to the 2009 scale values.
- The results of this calibration were used as the operational item parameters to create the preliminary scales for the NYSESLAT spring 2010 administration. Item estimates of the common items from the 2009 administration were used as anchors to place the 2010 scale to the established 2009 vertical scale. The final reporting scale was derived from the vertical scale using a linear transformation.

7.4 Equating of the NYSESLAT

Equating Method and Rationale

In 2010, the 2009 NYSESLAT vertical scale was used to establish the 2010 NYSESLAT scale. The common-item, non-equivalent groups design was used for 2010 and 2009 equating because it is one of the most practical of the various designs used for equating. For example, this design does not require large sample sizes necessary for the random groups design or a concern toward student fatigue by using the single group design (Kolen and Brennan, 1995), which makes the common items non-equivalent groups design a very popular choice for large-scale assessments.

The common-item, non-equivalent groups design necessitates the inclusion of items that are common to both administrations that would allow the linking of the tests and/or forms to be placed on the same scale. There are several methods used for this item response theory (IRT) scale transformation. Kolen and Brennan (2004) provide several IRT scale transformation methods that include the Mean/Mean, Mean/Sigma, Haebara, and the Stocking-Lord methods. For the Rasch Model, the equating constant of scale transformation is often used by researchers and practitioners alike. A fairly common practice among psychometricians is to use a method called “the anchor or the fixed parameter method.” In this method, the common items whose calibrations are known (either through a previous-year calibration or a separate calibration of the form) are anchored or fixed to their known estimates during calibration of the other forms that are to be put on the scale of the first form. In treating these common item parameters as known, they are not estimated, and the remaining item parameters (for the uncommon items) are forced onto the same scale as the anchored (fixed) items (Hanson & Béguin, 1999; Linacre, 2003; Taherbhai & Seo, 2007). For the NYSESLAT, the fixed method of scale transformation was used. This transformation was achieved with the *Winsteps* program that calibrated the items with the use of the two IRT models discussed above.

As Kolen and Brennan (1995) state, “After the item parameters are on the same scale, IRT true score equating can be used to relate number-correct scores on Form X and Form Y. In this process, the true score associated on one form associated with a given θ is considered to be equivalent to the true score on another form associated with that θ ” (p. 175). Furthermore, observed scores can easily be substituted for true scores because doing so has been justified in

IRT by the observation that the true score conversions are similar to observed score conversions (Lord & Wingersky, 1984). True score equating was used for the NYSESLAT by the application of the Newton-Raphson method for finding the roots of the non-linear function associated in the true score equating process. This procedure allows observed scores to be matched to their respective theta values, which in turn are converted to more acceptable scale score values (see scale score transformation to the SELP scale equations on p. 56). These thetas for the different forms/tests are on the same scale as established by the scale transformation method discussed above. By applying the true score equating process, the 2010 NYSESLAT was successfully equated to the 2009 NYSESLAT while maintaining the established 2009 common vertical scale, which was based on the 2006 common vertical scale and the 2005 SELP scale.

In establishing the 2010 vertical scale, the data file used for the calibration and equating of the NYSESLAT had 205,443 cases, which was about 87% of the final operational research file (total N=218,077) that Pearson received from New York. The calibration/equating data was representative of the population covering New York City and the Large City school districts (Buffalo, Syracuse, Rochester, and Yonkers).

Thus by fixing the known parameters of the common set of items between the two years, the items on the 2010 operational form were calibrated, and the newly administered items were then located on the 2009 scale. Once the scale locations of the 2010 NYSESLAT were known, IRT true score equating was used to relate the raw scores on the 2010 NYSESLAT to the 2009 scale. In this process, the true score on the NYSESLAT with a given level of examinee ability was considered an equivalent true score on the 2009 test associated with that level of examinee ability (Kolen and Brennan, 2004, p. 178).

Common Item Linking Design

As explained in the previous section, the common item design is contingent upon the use of items that are common on two administrations (i.e., between the examination that is already placed on the established scale and the one that is to be placed on that scale). Alternatively, one can use items on the current examination that have already been placed on an established scale. These items then act as anchor items to place the new form on the existing scale.

For the NYSESLAT, Pearson used the pre-existing 2009 vertical scale that was used to create the NYSESLAT vertical scale of 2010. For the 2010 administration, the common items between 2009 and 2010 were the anchor items on the 2010 NYSESLAT and were, as explained earlier, fixed to the parameter values from the pre-existing vertical scale of 2009. That is, the common items between 2009 and 2010 were used as anchor items between the two years. Any remaining non-common items on the NYSESLAT were calibrated together with the anchor items using the Rasch and Partial Credit Models. Fixing the values of the anchor items prior to calibration resulted in the item difficulty and step parameters of all the items being placed on the same ability metric.

Anchor Item Evaluation

There are some general guidelines (2010 NYSESLAT Form Construction Specifications) in the use of anchor items. These guidelines are as follows:

- Use an adequate number of anchor items.
- Anchor items selected for linking the two forms need to represent a miniature version of the test in terms of content.
- Anchor items need to appear in the same or similar locations as they do in the test that we are equating.
- Anchor items are expected to have strong statistical properties (e.g., point-biserial > 0.30) and span the range of difficulty including easy, medium, and difficult items. The mean difficulty (p-value/Rasch difficulty values) of the 2010 anchor items should be very similar to the mean difficulty of the 2009 operational test.

Stability of Anchor Items

The stability of the common item parameters was examined prior to their use as anchor items in the equating. There are various methods of evaluating anchor item stability. Most of them are based on how much of a difference exists between the parameters on the established scale and those on the transformed scale. Since the *Winsteps* program provides “displacement” values for the common items that are anchored, a check of these values can be used to exclude items from being anchor items. Items that exceed the displacement value of 0.30, for example, can be excluded as anchor items and the process is repeated until all displacement values are less than or equal to 0.30. Another criterion is that the total number of anchor items needs to be at least 20% of the total number of items for each modality combination. If the 20% anchor item criteria are not met, then a slightly higher displacement value of 0.50 will be used. No true or correct procedures can be applied in excluding common items as anchor items. A balance has to be determined between the minimum items desired in the anchor pool and the number of common items that are excluded from being anchor items. However, as a general rule, since the number of anchor items used in equating is an important criterion in the stability of the linking procedure, many psychometricians recommend no more than 20% of the common items be excluded from being anchor items (e.g., Huynh, 2000).

For the 2010 NYSESLAT, we used the displacement value of 0.30 (as is done by many large-scale assessments) for an item to be considered as an outlier. We used this criterion consistently across all the grade spans. For all grade spans, the overall numbers of anchor items were around 25% of the total items. See Table 38 for the number of items excluded for each modality combination by grade span.

Items that were excluded from being anchor items had their parameters estimated, and the new estimate became the adjusted item parameter in the item bank for future use.

Table 39 - Common Items as Anchor Items by Modality Combination and Grade Span

Grade Span	Modality Combination	Total Number of Items	Common Items ¹ (2009 and 2010)	Percentage of Total Items	Anchor Items (2008 and 2009)	Percentage of Total Items
K-1	Listening/Speaking	40	10	0.25	10	0.25
	Reading/Writing	30	8	0.27	8	0.27
2-4	Listening/Speaking	40	10	0.25	10	0.25
	Reading/Writing	40	10	0.25	9	0.23
5-6	Listening/Speaking	41	10	0.24	10	0.24
	Reading/Writing	46	11	0.24	10	0.22
7-8	Listening/Speaking	41	11	0.27	10	0.24
	Reading/Writing	45	11	0.24	12	0.26
9-12	Listening/Speaking	41	11	0.27	11	0.27
	Reading/Writing	46	11	0.28	13	0.28

Note 1: Common items are the same items appearing in both the 2009 and 2010 operational tests.

TCC and Standard Error Curves Comparison

In order to ascertain the maintenance of the scale across years, comparisons between the 2008, 2009, and 2010 administrations' Test Characteristic Curves (TCC) and Standard Error Curves (SE Curves) were examined with respect to their alignment at each level of administration. For TCCs, the x -axis represents the thetas, or students' abilities, with the y -axis representing the expected score based on the number or proportion of items correctly solved by students. TCCs represent the relative difficulty of a given test form, with TCCs on the right representing more difficult test forms. Measurement error is associated with every test score. The standard error of measurement (SEM) can be used to obtain a range within which a student's true score is likely to fall. The x -axis represents the thetas, or students' abilities, with the y -axis representing the SEM.

The 2010 TCC plots for all grade spans (see Figures 3 & 4) indicate that the TCC shifts to the right with each progression of grade span across the Listening/Speaking and Reading/Writing vertical scales with the exception of grade span 5-6 and grade span 7-8 in Reading/Writing. Overall, the TCCs of 2009 and 2010 align well.

The differences in terms of the SE curves among 2008, 2009, and 2010 for both the Listening/Speaking and Reading/Writing tests are minimal across all the grade spans. The noticeable differences occur only at the tail portion of the SE curves.

Listening and Speaking Test Vertical Scale

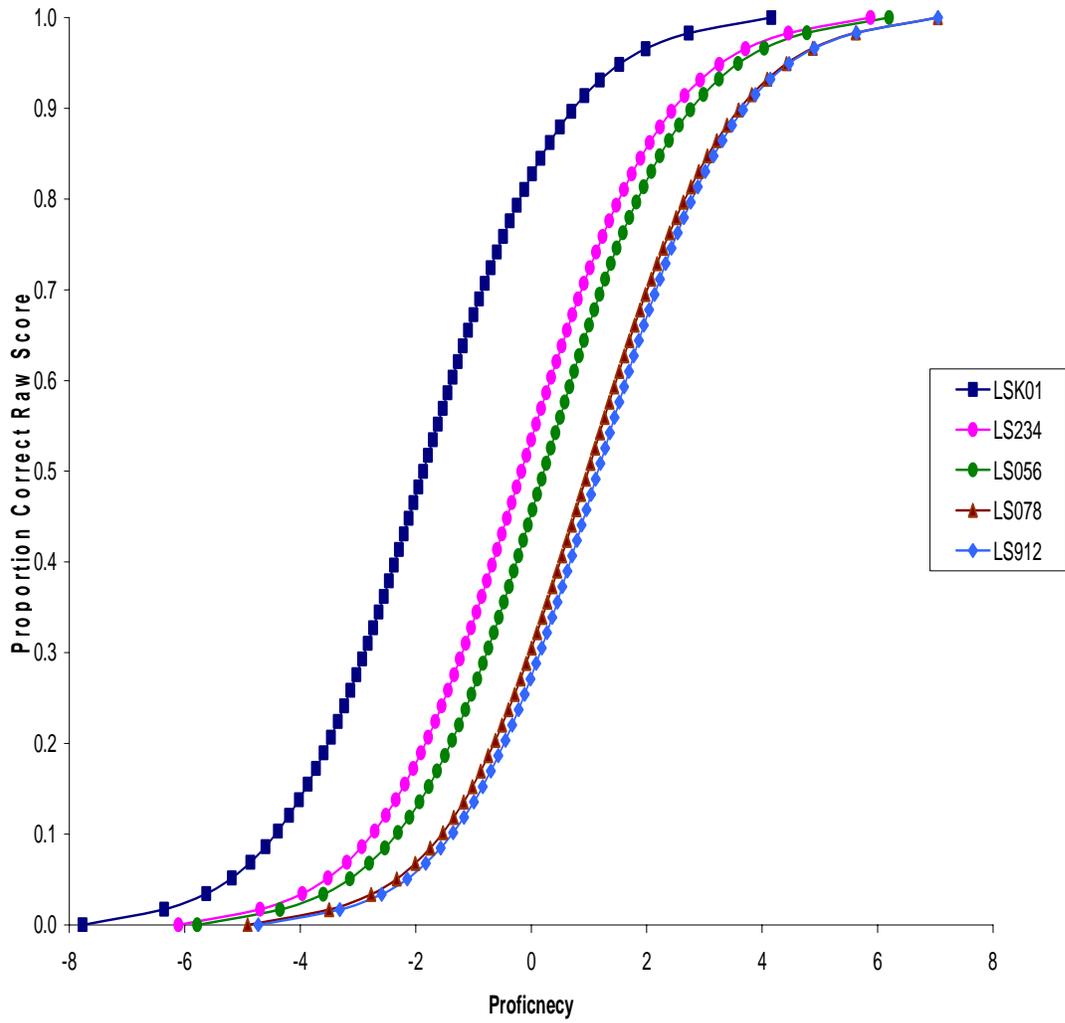


Figure 3 - 2010 TCC L/S Grade Span 1-5

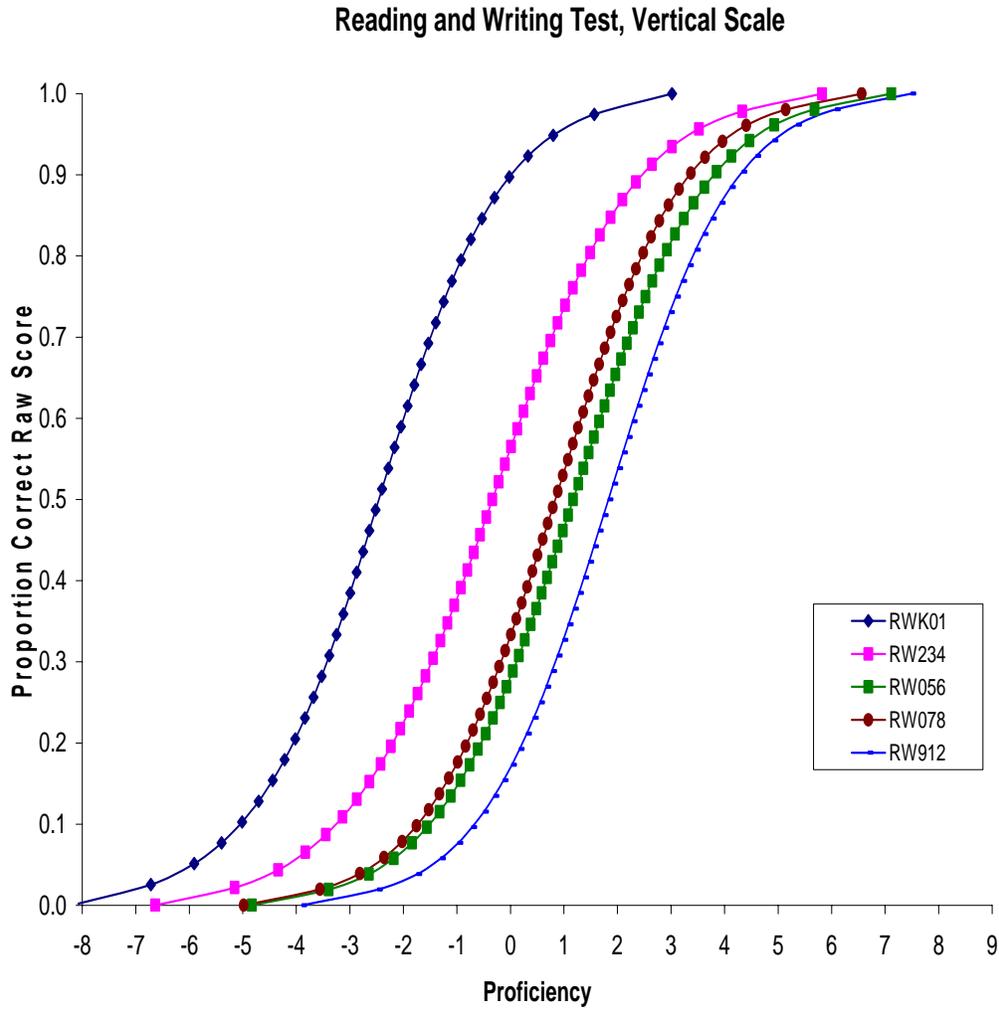


Figure 4 - 2010 TCC R/W Grade Span 1-5

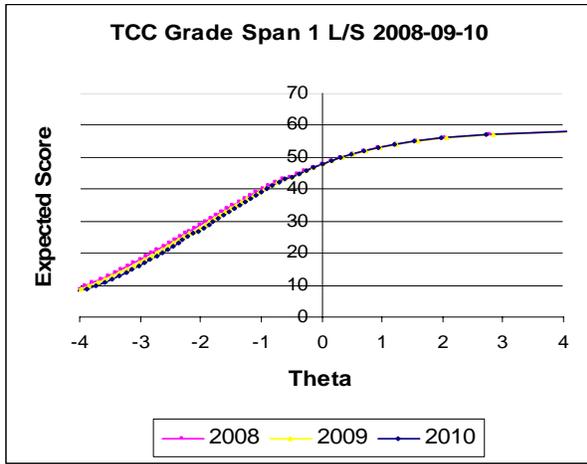


Figure 5 - 2008-09-10 TCC L/S Grade Span 1

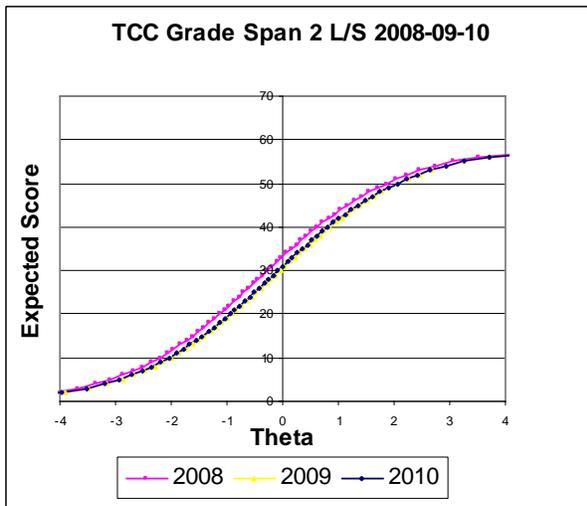


Figure 6 - 2008-09-10 TCC L/S Grade Span 2

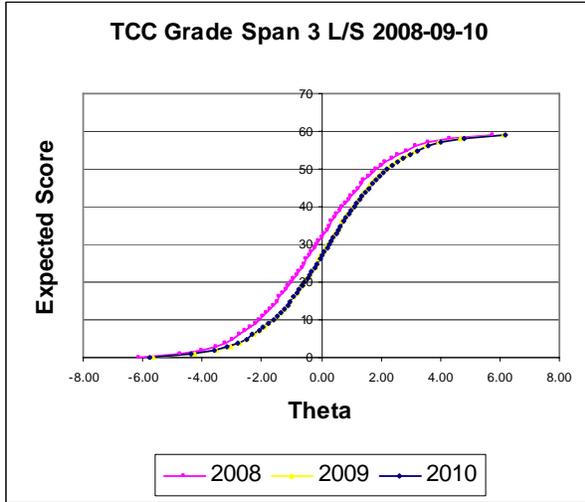


Figure 7 - 2008-09-10 TCC L/S Grade Span 3

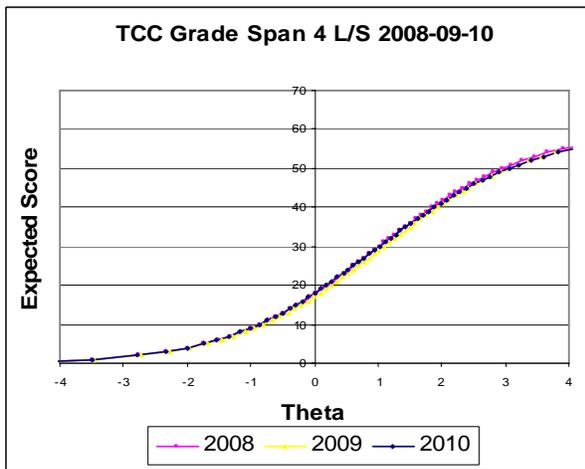


Figure 8 - 2008-09-10 TCC L/S Grade Span 4

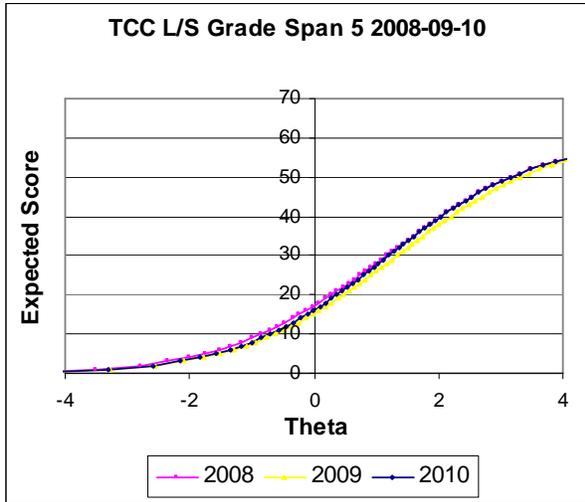


Figure 9 - 2008-09-10 TCC L/S Grade Span 5

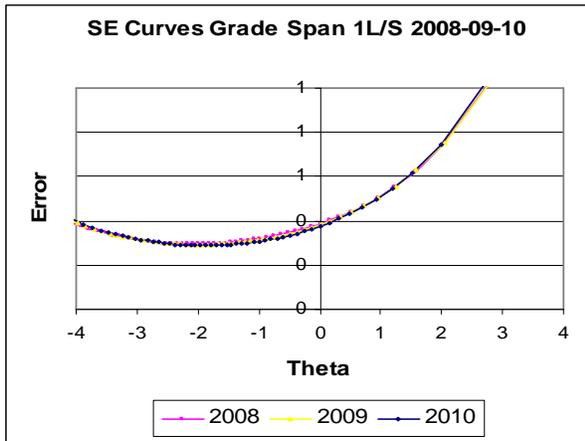


Figure 10 - 2008-09-10 SE Curves L/S Grade Span 1

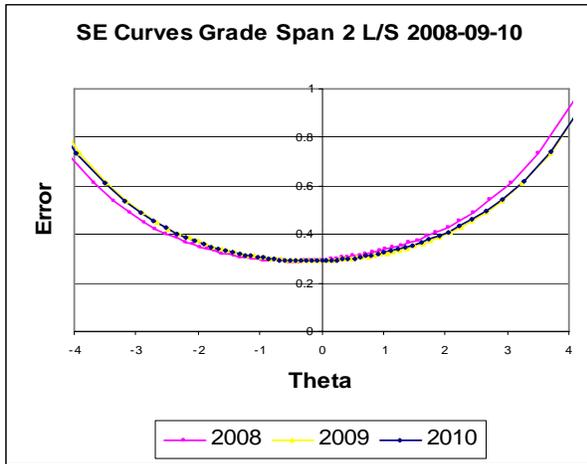


Figure 11 - 2008-09-10 SE Curves L/S Grade Span 2

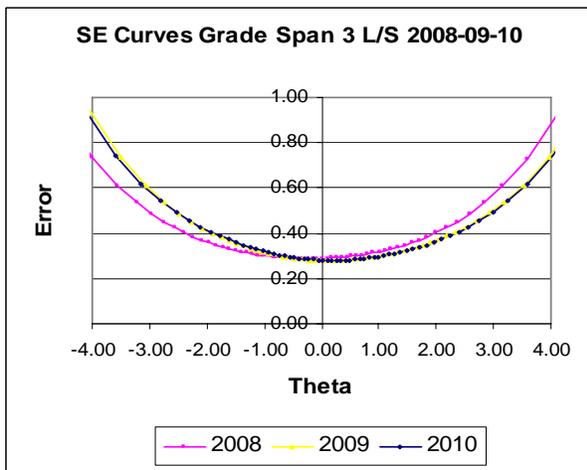


Figure 12 - 2008-09-10 SE Curves L/S Grade Span 3

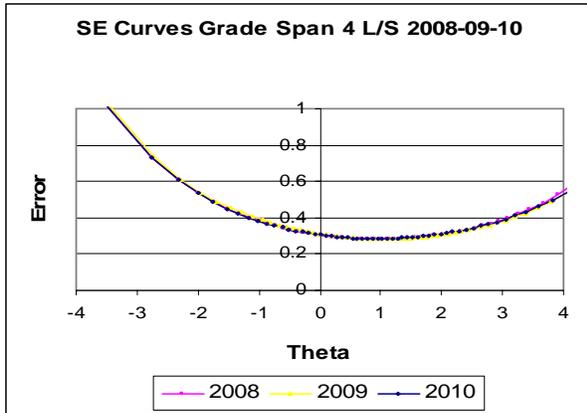


Figure 13 - 2008-09-10 SE Curves L/S Grade Span 4

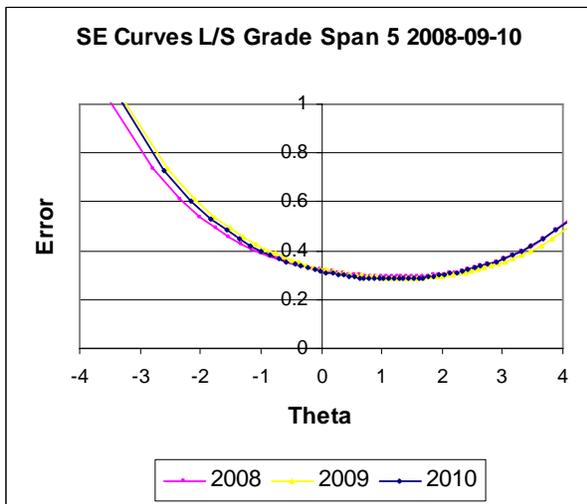


Figure 14 - 2008-09-10 SE Curves L/S Grade Span 5

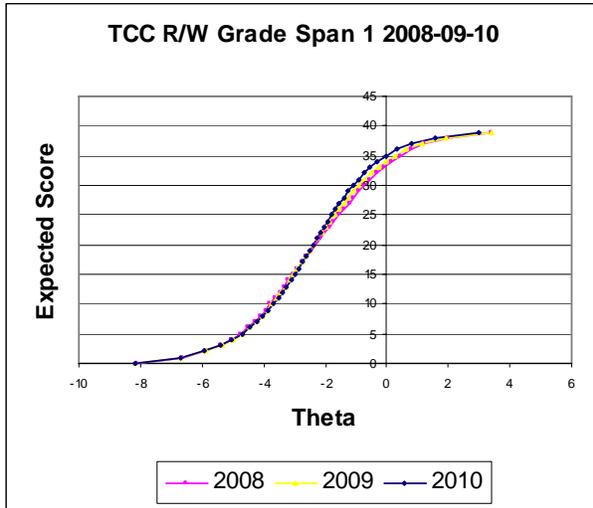


Figure 15 - 2008-09-10 TCC R/W Grade Span 1

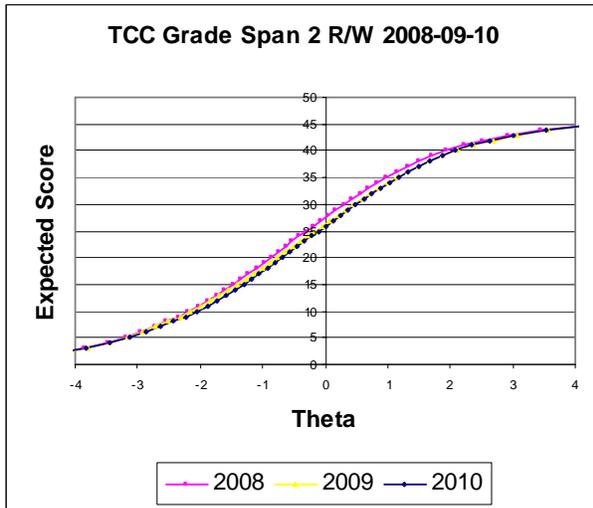


Figure 16 - 2008-09-10 TCC R/W Grade Span 2

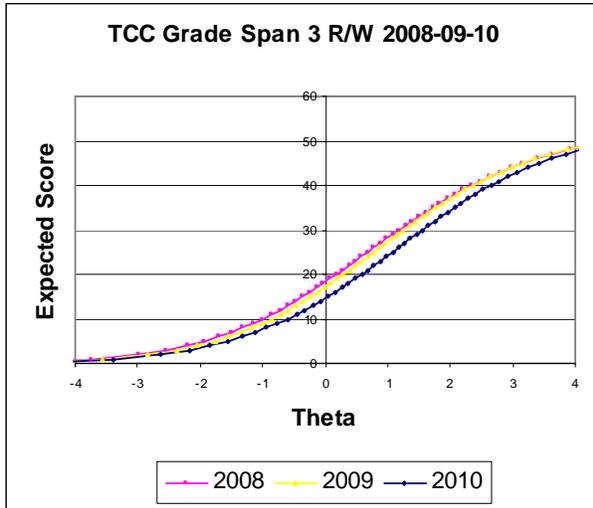


Figure 17 - 2008-09-10 TCC R/W Grade Span 3

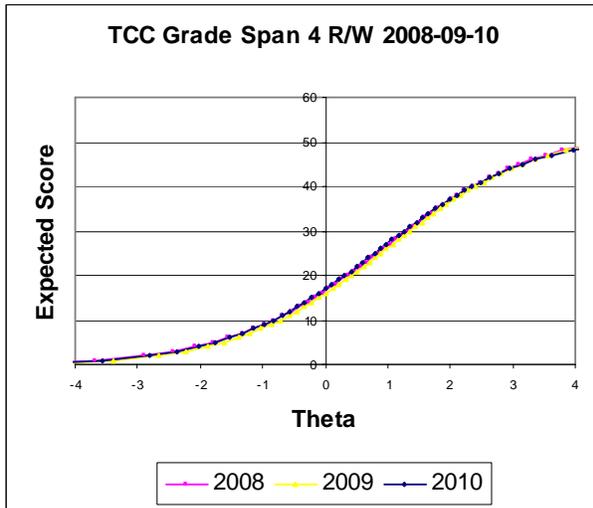


Figure 18 - 2008-09-10 TCC R/W Grade Span 4

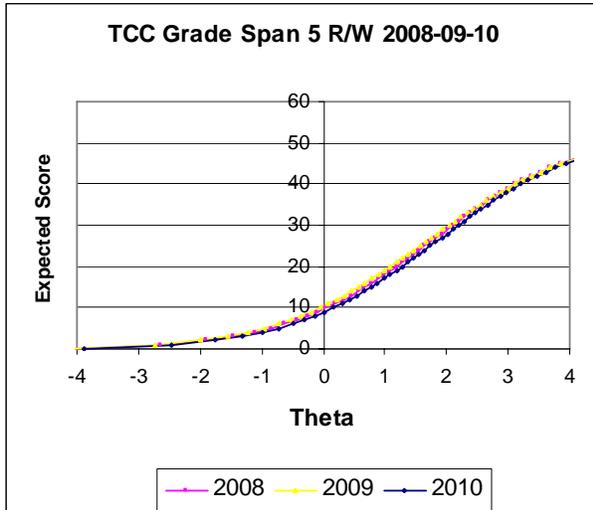


Figure 19 - 2008-09-10 TCC R/W Grade Span 5

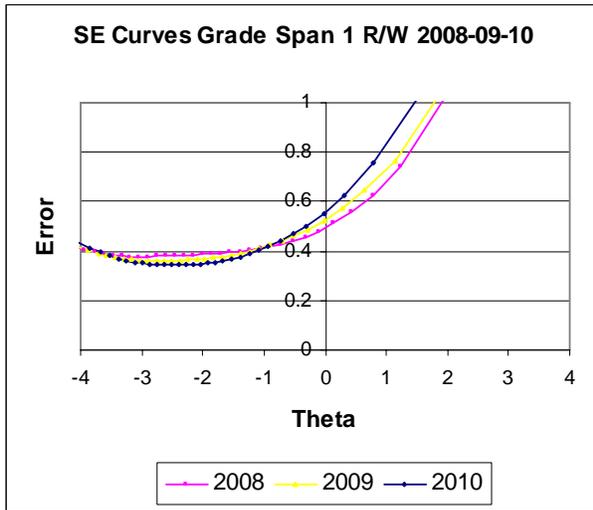


Figure 20 - 2008-09-10 SE Curves R/W Grade Span 1

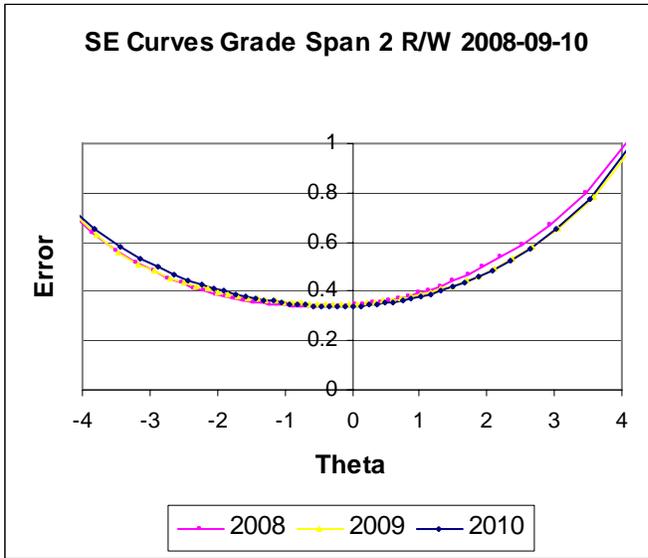


Figure 21 - 2008-09-10 SE Curves R/W Grade Span 2

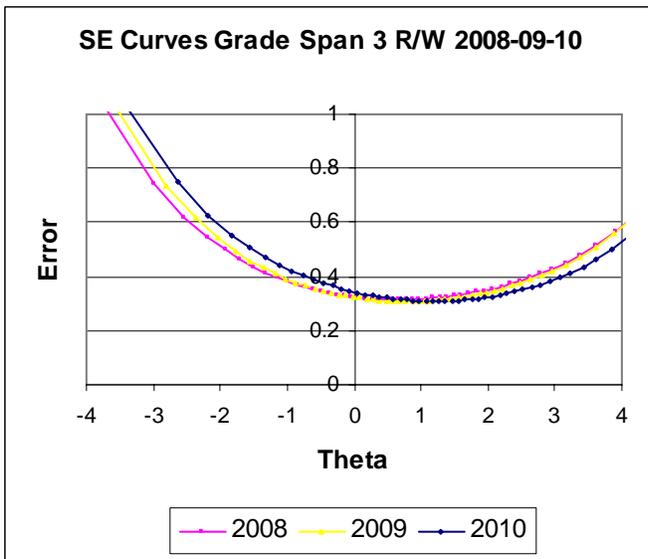


Figure 22 - 2008-09-10 SE Curves R/W Grade Span 3

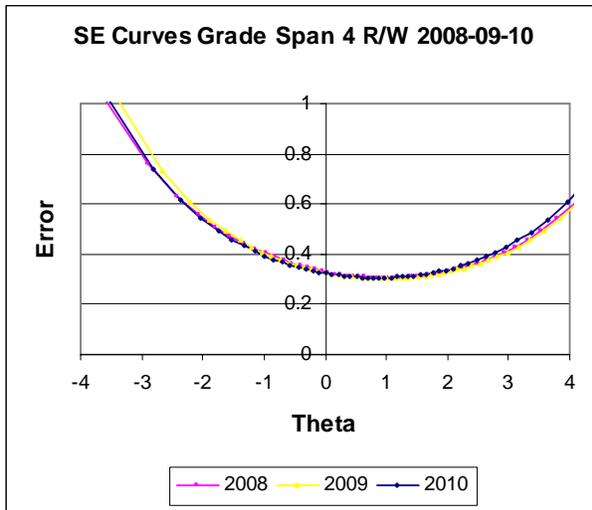


Figure 23 - 2008-09-10 SE Curves R/W Grade Span 4

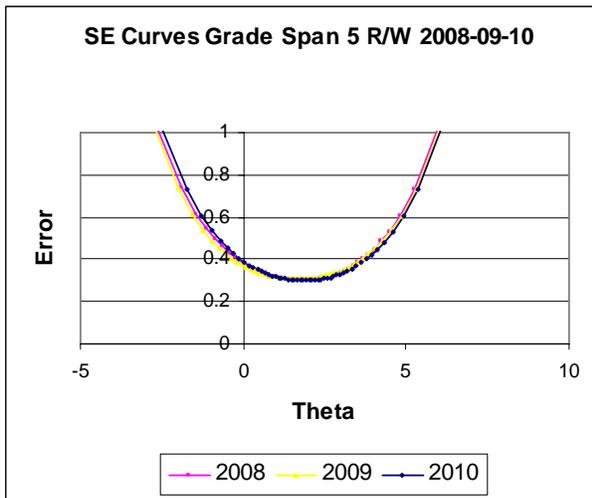


Figure 24 - 2008-09-10 SE Curves R/W Grade Span 5

7.5 Scaling of the NYSESLAT

As explained earlier, the NYSESLAT vertical scale was created based on the pre-existing 2008 vertical scale. The measurement of students' progress across years is augmented by the use of such a vertical scale. Since student abilities and item parameters are placed on an *equated* vertical scale, year-to-year progress in language proficiency for students who are administered the NYSESLAT for the three consecutive years can also be measured and documented. On this scale, a scale score of 500, for example, is greater than 490 and less than 510, irrespective of the students' grade levels and the year of administration. The scale score estimate of students, therefore, becomes an indication of the progress they have made over the previous year's estimate.

The NYSESLAT scaling procedure involves linear transformations of the raw score points into scale score points. These transformations do not give more weight to particular subtests, and they change neither the rank ordering of students nor their performance level classification. Linear transformation constants are utilized. The following equation was used to derive the scale scores:

$$\text{Scale Score} = 35 * (\text{theta}) + 600 \quad (17)$$

In the above equation, theta was derived from item parameters that have been adjusted for the test form and grade span level.

Raw Score to Scale Score Tables

The final reporting scales were used to produce raw score-to-scale score conversion tables for the Speaking/Listening and Reading/Writing combinations. Appendices C.1–C.5 provide the raw-to-scale score conversion tables for the Speaking/Listening and Reading/Writing combinations by grade span.

7.6 Rasch Information

Table 39 presents the grade span, the modality, the number of items in each modality, the maximum number of points attainable for each modality, and the average Rasch difficulty for each modality. Appendix E includes item information at different cut points at each grade level.

Appendix D contains the results of the operational items for the NYSESLAT. The following IRT item parameters are presented for each item grouped by Listening/Speaking and Reading/Writing combinations:

- Number of students
- Rasch difficulty value
- Standard error (SE) of Rasch difficulty
- MNSQ Infit: Standardized information-weighted mean-square statistic, which is sensitive to unexpected behavior affecting responses to items near the person's ability level
- MNSQ Outfit: Standardized outlier-sensitive mean-square fit statistic that is sensitive to unexpected behavior by persons on items far from the person's ability level
- Flag of Misfit

Table 40 - Average Rasch Difficulty by Grade Span and Modality

Grade Span	Modality	Number of Items	Max Points	Average Rasch Difficulty
K-1	Listening	24	24	-2.11
	Speaking	16	34	-1.71
	Reading	15	15	-2.45
	Writing	15	24	-2.65
	Listening and Speaking	40	58	-1.91
	Reading and Writing	30	39	-2.55
	Total	70	97	-2.23
2-4	Listening	24	24	-0.04
	Speaking	16	34	-0.27
	Reading	24	24	-0.50
	Writing	16	22	-0.33
	Listening and Speaking	40	58	-0.16
	Reading and Writing	40	46	-0.41
	Total	80	104	-0.29
5-6	Listening	25	25	0.49
	Speaking	16	34	0.02
	Reading	27	27	1.13
	Writing	19	25	1.04
	Listening and Speaking	41	59	0.25
	Reading and Writing	46	52	1.09
	Total	87	111	0.67
7-8	Listening	25	25	1.49
	Speaking	16	34	0.63
	Reading	26	26	0.91
	Writing	19	25	0.72
	Listening and Speaking	41	59	1.06
	Reading and Writing	46	51	0.82
	Total	86	110	0.94
9-12	Listening	25	25	1.43
	Speaking	16	34	0.93
	Reading	27	27	2.06
	Writing	19	25	1.50
	Listening and Speaking	41	59	1.18
	Reading and Writing	46	51	1.78
	Total	87	111	1.48

Note: The statistics above are based on the equating sample (invalid cases were removed after cleaning procedures).

PART 8: STANDARD SETTING

8.1 Introduction

As the contractor for the NYSESLAT, Pearson organized a performance standard-setting meeting. The standard-setting sessions were conducted in Albany, New York, from March 28 to April 1, 2005. The purpose of this meeting was to provide preliminary recommendations on performance cut scores for the NYSESLAT.

For each group, there was one psychometrics staff member from Pearson to facilitate the technical part of the standard setting. In addition, a content specialist from Pearson and a NYSED official were present to provide support during the standard-setting sessions.

8.2 Standard Setting Model

Item mapping is a well-established method available for establishing performance standards. The item-mapping procedure is capable of incorporating both multiple-choice and constructed-response items into the same process (Mitzel, H.C., Lewis, D.M., & Green, D.R., 2001). It has several other favorable characteristics, including:

- Simplifying the judgment task by reducing the cognitive load required by panelists
- Connecting the judgment task of setting cut scores with the measurement model
- Connecting content with performance level descriptors

The item-mapping procedure orders items for each test into a booklet according to the difficulty of the items, which is determined by IRT scaling techniques. Easy items are placed in the beginning of the booklet, and subsequent items become increasingly more difficult as the booklet progresses. Passages, rubrics, and sample student responses are placed in the booklet for reference purposes. Panelists are also provided with test blueprints and various other materials, including performance level descriptors, scoring rubrics, test booklets, and anchor papers.

8.3 Committees of Panelists

Five standard-setting committees were established to set the cut scores for the five grade spans of the NYSESLAT. As indicated in Table 40, the first group recommended standards on Grades K–1; the second group recommended standards on Grades 2–4; the third group recommended standards on Grades 5–6; the fourth group recommended standards on Grades 7–8; and the fifth group recommended standards on Grades 9–12.

The panel members were New York State certified ESL, ELA, bilingual education, and bilingual special education teachers who were familiar with the content standards. The NYSESLAT panelists were recruited by NYSED to participate in the standard-setting meeting. The panelists possess knowledge of working with students with limited English proficiency and represent all regions of the state of New York. Other educational stakeholders, such as administrators, curriculum specialists, NYSED members, and members of the professional community were also

included, as per recommendations. These additional members tend to provide valuable insights from their area of expertise and help strengthen the consequential validity argument during panel discussions.

Table 41 - Panel Composition for Standard-Setting Committees

Grade	Group	Number of Judges
K-1	1	20
2-4	2	20
5-6	3	18
7-8	4	13
9-12	5	16

8.4 Performance Levels and Cut Scores

For the NYSESLAT, four performance levels are required, and they respond to three cut scores. The four performance levels are:

- Beginning
- Intermediate
- Advanced
- Proficient

The three cut scores are:

- Intermediate (between the Beginning and Intermediate performance levels)
- Advanced (between the Intermediate and Advanced performance levels)
- Proficient (between the Advanced and Proficient performance levels)

To set the three cut points, the item-mapping procedure was utilized. The standard-setting process is described below.

8.5 Standard-Setting Process

The standard setting began with introductions from NYSED, Pearson, and panelists. This was followed by a presentation by the lead facilitator on the role of the panelists in the standard-setting process: setting performance standards and placing cut scores. The goal was to familiarize panelists with the standard-setting process and the item-mapping procedure. This session took place in a large group setting (all five groups together).

After the orientation, the panelists were separated into specific breakout rooms according to their group assignments. Each group/room was led by a facilitator who was an expert in the standard-setting methodology. Assessment specialists rotated from group to group in order to provide content support. In addition, the panel members were further divided into three smaller table groups within their grade spans, each composed of five to seven members. These small groups worked independently but had the opportunity to collaborate with the other table groups in their grade span during the standard-setting process. The following sequences of tasks were followed:

Review of the Assessment

The first task was to review the assessment blueprint. This was done in order for the panelists to gain an understanding of what the assessment was intended to measure. Discussions about the assessment content, the use of different item types, and the number of questions were conducted. The panel members further defined the general performance level descriptors into specific descriptors to help the panel members come to a shared understanding about what it meant to be performing at each of the performance levels. The facilitator led this discussion with support from the assessment specialist who floated between the rooms.

Experience the Assessment

Next, the panel members had an opportunity to experience the assessment administered at the grade span assigned to them. This was an effective way to demonstrate to the panelists the knowledge and skills that students must possess to obtain a high score. It is assumed that panelists are likely to set more realistic performance standards if they experience the assessment themselves.

Scoring the Assessment

After the panelists finished taking the assessment, they were provided with an answer key to grade their test. The panelists scored their own assessments using the scoring rubrics and answer key provided. The scoring process offered an opportunity for the panelists to develop an understanding of the scoring of open-ended responses. They were provided with exemplars of score points. A discussion session followed the scoring of the assessment.

Review of Student Performance Levels

Panelists reviewed the previously established definitions of performance levels (Appendices F.3–F.4, 2006 NYSESLAT Technical Report). Then they discussed the performance levels. The goal was to help panelists clearly distinguish between student performance levels. Panelists' suggestions were related to the performance standards and content frameworks. The suggestions were retained for reference during the standard-setting process. Panelists reviewed definitions and offered illustrative suggestions for the Beginning, Intermediate, Advanced, and Proficient performance standards. After all the performance levels were reviewed, a discussion session was held. The focus was on the characteristics and interrelationships of performance standards.

Three Rounds of Ratings

The actual standard setting proceeded in three rounds. Each round was designed to foster increased consensus among panelists, although reaching consensus was not necessary. Panelists expressed their cut score judgment by placing a marker on the item that a student at that threshold of a performance level should master. One marker was placed for each cut score. There were three cut points.

During the Round 1 ratings, each panelist began by setting his/her three cut scores. The data were captured for each panelist. Before the Round 2 ratings, panelists were provided feedback on the Round 1 cut score positions of all panelists and their group medians. The panelists then discussed the Round 1 results. After the discussions, the Round 2 cuts were made, followed by further discussions. At this point, the panelists were provided with information about the percentage of students who would be classified in each of the performance levels, if those cuts were to be implemented.

In order to promote consistency across the grade spans, the groups came together to discuss the process and results of their assigned grades among all grade spans. Panelists then got back into their breakout groups and proceeded to make their Round 3 ratings. The median cut scores of the panelists then served as the starting point for the decision makers on establishing the cut scores for the assessment.

8.6 Summary Statistics for the Three Rounds of Ratings

Appendix F.2 (2006 NYSESLAT Technical Report) provides the summary statistics for the round-by-round results by grade of the three performance level cuts. The statistics include the range, the raw score mean and standard deviation, the standard error of the mean, the median, standard error of the median, and the interquartile range.

8.7 Post-Standard-Setting Analyses

The median scores from the standard-setting committees were used as the recommended cuts. The cut scores were based on the total NYSESLAT score. After the standard-setting meetings, Pearson performed several post-standard-setting analyses. The first step was to look up the equivalent scale scores corresponding to the raw score cuts recommended by the committees. Graphs were then plotted using the grades as the independent variable and scale score as the dependent variable. The three cut points were then plotted on the same graph to show that the cuts were monotonically increasing from the lower cuts to the higher cuts. Some technical adjustments were made to the recommended cuts. Any adjustments made were within one standard error of the median. A quadratic equation was then applied to the scaled cut scores across all 13 grades, K–12. Further analyses were performed to smooth the cuts. This was done in order to ensure that there are no reversals.

Next, the proportion of the three cut points in raw score points based on the total test was applied to the Listening/Speaking and Reading/Writing combinations for each of the grades. Impact analysis was conducted on the 2005 operational data. The percentage of students falling into each of the performance levels was calculated for each grade if those cut points were adopted. This information was provided to the NYSED to help them make their final decisions on the cut points for the Listening/Speaking and Reading/Writing combinations.

8.8 Final Performance-Level Cut Points

The final cut points adopted by NYSED for the 2010 administration of the NYSESLAT for the Listening/Speaking and Reading/Writing combinations in raw score points, scale score, and theta metric (from standard setting) are presented in Table 41 and Table 42. There are three cut points that correspond to four performance levels. Any score below the Intermediate cut point is the Beginning performance level.

Table 42 - Final Performance-Level Cut Points for Listening and Speaking

Grade	Listening & Speaking								
	Raw Score Cut, 2010			Scale Score			Theta		
	I	A	P	I	A	P	I	A	P
K	21	42	51	511	573	616	-2.55	-0.70	0.50
1	24	44	55	520	582	649	-2.29	-0.49	1.53
2	14	32	48	543	602	661	-1.55	0.09	1.74
3	17	35	51	555	611	675	-1.23	0.35	2.23
4	20	38	52	567	621	684	-0.95	0.62	2.43
5	20	38	52	579	631	686	-0.56	0.92	2.56
6	23	41	53	587	639	696	-0.30	1.19	2.76
7	17	36	50	597	651	703	-0.09	1.53	3.05
8	20	39	50	605	661	703	0.19	1.79	3.05
9	24	42	51	623	678	715	0.71	2.23	3.31
10	25	45	52	628	686	721	0.79	2.54	3.48
11	28	47	53	635	694	728	1.04	2.76	3.67
12	30	49	53	640	704	728	1.20	3.02	3.67

Table 43 - Final Performance-Level Cut Points for Reading and Writing

Grade	Reading & Writing								
	Raw Score Cut, 2010			Scale Score			Theta		
	I	A	P	I	A	P	I	A	P
K	25	34	36	537	582	609	-1.79	-0.52	0.25
1	30	36	38	559	609	655	-1.17	0.25	1.58
2	19	32	39	571	623	663	-0.82	0.64	1.81
3	21	33	41	580	630	679	-0.57	0.86	2.25
4	24	36	42	589	642	691	-0.33	1.19	2.61
5	20	32	43	620	660	704	0.56	1.73	2.98
6	21	34	44	623	668	710	0.66	1.94	3.15
7	26	39	45	629	674	710	0.83	2.13	3.15
8	27	40	45	632	682	710	0.92	2.34	3.15
9	20	37	43	643	698	726	1.21	2.80	3.60
10	21	38	44	646	702	732	1.30	2.91	3.78
11	22	40	46	649	711	739	1.39	3.16	3.97
12	23	41	46	652	715	739	1.48	3.30	3.97

Note: The statistics in the above two tables were based on the valid cases in the equating sample. I = Intermediate, A = Advanced, and P = Proficient.

PART 9: SUMMARY OF OPERATIONAL TEST RESULTS

This part presents both the raw score and scale score summaries for each of the reporting combinations (Listening/Speaking and Reading/Writing). Table 43 presents the raw score summary by grade. Table 44 presents the raw score summary by grade span. Table 45 presents the scale score summary by grade. Table 46 presents the scale score summary by grade span. Tables 43 through 46 include the sample size (n-count), the mean, median, interquartile range (IQR), range, and the standard deviation (SD). Table 47 presents the percentage of students in each of the proficiency levels by grade. Table 48 presents the percentage of students in each of the proficiency levels by grade span. The overall proficiency level is defined by the lower of the two proficiency level designations (e.g., Listening/Speaking level=3, Reading/Writing level=4, Overall level=3). Table 49 presents the exit rate by grade for 2007, 2008, 2009, and 2010, administrations. Table 50 presents the exit rate by grade span for 2007, 2008, 2009, and 2010 administrations. These two tables include n-counts and the percentage of ELL students receiving exit status after scoring at the Proficient Level (level 4) on both the Listening/Speaking and Reading/Writing modality combinations. Additionally, similar types of analyses (See Appendices G, H, and I) were done according to the following subgroups:

- Gender
- Ethnicity (American Indian or Alaska Native, Black or African American, Asian, Native Hawaiian/other Pacific Islander, Hispanic or Latino, White, and Multi-Racial [not of Hispanic origin])
- Needs/Resource Group (NYC, Large City, High Need Urban/Suburban, High Need Rural, Average Need, Low Need, Charter Schools)
- Five major languages background (Arabic, Bengali, Chinese, English, Spanish), language missing, and other languages
- Number of years as an ELL
- Students with disabilities (Autism, emotional disturbance, learning disability, mental retardation, deafness, hearing impairment, speech or language impairment, visual impairment, orthopedic impairment, other health impairment, multiple disabilities, traumatic brain injury)

Table 44 - Raw Score Summary by Grade

Grade	Test	N- count	Mean	Median	IQR	Range	SD
K	Listening and Speaking	25,927	44.12	47	13	2~58	10.53
	Reading and Writing	26,225	25.56	26	12	2~39	7.90
1	Listening and Speaking	27,472	50.06	53	8	3~58	8.70
	Reading and Writing	27,934	33.82	36	6	2~39	6.31
2	Listening and Speaking	23,635	46.28	49	10	2~58	9.56
	Reading and Writing	22,224	30.30	31	12	5~46	7.94
3	Listening and Speaking	21,889	48.77	52	8	2~58	9.59
	Reading and Writing	21,040	34.40	36	9	2~46	7.52
4	Listening and Speaking	19,829	49.63	53	8	3~58	10.04
	Reading and Writing	19,296	36.26	39	9	3~46	7.54
5	Listening and Speaking	16,381	47.06	51	10	4~59	10.92
	Reading and Writing	15,576	35.70	38	13	3~52	9.36
6	Listening and Speaking	13,925	46.22	51	12	4~59	11.79
	Reading and Writing	13,283	35.61	38	14	5~52	9.65
7	Listening and Speaking	13,230	44.94	50	14	3~59	12.39
	Reading and Writing	12,607	34.78	37	15	5~51	9.70
8	Listening and Speaking	12,997	45.09	50	14	3~59	12.24
	Reading and Writing	12,444	35.71	38	14	7~51	9.59
9	Listening and Speaking	14,456	42.89	47	20	3~59	13.26
	Reading and Writing	13,472	32.08	33	15	4~52	9.65
10	Listening and Speaking	13,032	44.90	48	16	3~59	11.67
	Reading and Writing	12,523	33.78	35	14	6~52	9.35
11	Listening and Speaking	9,208	48.05	51	12	4~59	9.69
	Reading and Writing	8,984	36.83	38	13	5~52	8.55
12	Listening and Speaking	6,096	51.14	53	8	4~59	7.15
	Reading and Writing	6,016	38.53	40	10	4~52	7.69

Table 45 - Raw Score Summary by Grade Span

Grade	Test	N- count	Mean	Median	IQR	Range	SD
K-1	Listening and Speaking	53,399	47.17	50	12	2~58	10.08
	Reading and Writing	54,159	29.82	32	13	2~39	8.23
2-4	Listening and Speaking	65,353	48.13	51	10	2~58	9.82
	Reading and Writing	62,560	33.52	35	11	2~46	8.08
5-6	Listening and Speaking	30,306	46.68	51	11	4~59	11.34
	Reading and Writing	28,859	35.66	38	13	3~52	9.49
7-8	Listening and Speaking	26,227	45.02	50	14	3~59	12.32
	Reading and Writing	25,051	35.24	38	14	5~51	9.66
9-12	Listening and Speaking	42,792	45.79	50	15	3~59	11.69
	Reading and Writing	40,995	34.59	36	14	4~52	9.36

Note: The data file used in Tables 43 and 44, above, is the final research file. The total n-counts for Grades K-12 are 218,077. The stats presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

Table 46 - Scale Score Summary by Grade

Grade	Test	N-count	Mean	Median	IQR	Range	SD
K	Listening and Speaking	25,927	595.35	596	59	403~746	50.04
	Reading and Writing	26,225	547.70	542	58	393~706	47.92
1	Listening and Speaking	27,472	633.52	632	68	418~746	56.48
	Reading and Writing	27,934	616.40	611	81	393~706	61.19
2	Listening and Speaking	23,635	664.41	666	54	461~806	46.27
	Reading and Writing	22,224	623.08	622	56	490~804	40.37
3	Listening and Speaking	21,889	682.20	685	58	461~806	52.34
	Reading and Writing	21,040	645.71	646	51	448~804	42.74
4	Listening and Speaking	19,829	692.15	693	69	477~806	57.88
	Reading and Writing	19,296	659.43	665	62	466~804	47.74
5	Listening and Speaking	16,381	677.05	684	52	502~817	47.99
	Reading and Writing	15,576	680.35	684	53	524~849	40.30
6	Listening and Speaking	13,925	674.61	684	59	502~817	51.27
	Reading and Writing	13,283	680.54	684	57	545~849	42.08
7	Listening and Speaking	13,230	696.82	707	68	518~847	52.66
	Reading and Writing	12,607	666.62	669	59	539~830	41.09
8	Listening and Speaking	12,997	698.44	707	68	518~847	53.68
	Reading and Writing	12,444	671.01	673	56	554~830	41.68
9	Listening and Speaking	14,456	695.32	697	83	525~847	57.66
	Reading and Writing	13,472	686.84	686	51	566~862	36.98
10	Listening and Speaking	13,032	702.99	701	71	525~847	54.25
	Reading and Writing	12,523	693.42	693	50	583~862	37.21
11	Listening and Speaking	9,208	717.18	716	63	536~847	51.30
	Reading and Writing	8,984	705.54	704	52	575~862	36.74
12	Listening and Speaking	6,096	732.01	728	55	536~847	45.58
	Reading and Writing	6,016	711.91	712	42	566~862	33.96

Table 47 - Scale Score Summary by Grade Span

Grade	Test	N-count	Mean	Median	IQR	Range	SD
K-1	Listening and Speaking	53,399	614.99	611	74	403~746	56.75
	Reading and Writing	54,159	583.13	574	95	393~706	64.97
2-4	Listening and Speaking	65,353	678.79	678	67	461~806	53.30
	Reading and Writing	62,560	641.90	641	60	448~804	46.06
5-6	Listening and Speaking	30,306	675.93	684	56	502~817	49.54
	Reading and Writing	28,859	680.44	684	53	524~849	41.13
7-8	Listening and Speaking	26,227	697.62	707	68	518~847	53.17
	Reading and Writing	25,051	668.80	673	56	539~830	41.44
9-12	Listening and Speaking	42,792	707.59	711	74	525~847	55.17
	Reading and Writing	40,995	696.63	697	51	566~862	37.73

Note: 1. Generally speaking, the mean for each grade/grade span should increase from one grade/grade span to the next higher grade/grade span in a similar manner, as shown in Table 41 and Table 42 of this manual, which depicts increases across the grade levels. However, due to artifacts of the population whereby some grades/grade spans may have a greater percentage of higher scoring students than the next higher grade/grade span, the mean for the lower grade/grade span can be higher than the next higher grades/grade spans. 2. The statistics in Tables 45 and 46 are based on the final research file. The total n-counts for grade K-12 are 218,077. The stats presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

Table 48 - Percentage of Students in Each Proficiency Level by Grade

Grade	Test	Proficiency Levels			
		Beginning	Intermediate	Advanced	Proficient
K	Listening and Speaking	5.47	26.37	35.73	32.43
	Reading and Writing	46.68	26.73	11.91	14.69
	Overall	45.79	29.98	14.33	9.89
1	Listening and Speaking	3.04	12.42	46.62	37.91
	Reading and Writing	18.27	18.37	27.96	35.41
	Overall	17.46	21.33	39.04	22.17
2	Listening and Speaking	1.20	7.21	33.83	57.76
	Reading and Writing	13.31	37.77	32.70	16.21
	Overall	13.08	38.11	34.88	13.92
3	Listening and Speaking	2.29	6.38	31.82	59.50
	Reading and Writing	9.56	21.89	47.07	21.48
	Overall	9.36	22.27	49.33	19.05
4	Listening and Speaking	3.49	7.27	27.53	61.71
	Reading and Writing	9.74	23.28	40.96	26.01
	Overall	9.49	23.61	43.65	23.24
5	Listening and Speaking	5.14	9.72	38.78	46.36
	Reading and Writing	12.71	25.38	39.81	22.10
	Overall	12.42	25.97	44.60	17.00
6	Listening and Speaking	8.11	14.17	46.16	31.55
	Reading and Writing	16.01	29.65	35.42	18.93
	Overall	15.88	31.68	40.64	11.79
7	Listening and Speaking	5.45	13.14	30.26	51.14
	Reading and Writing	19.45	32.47	32.65	15.43
	Overall	19.20	32.78	34.62	13.40
8	Listening and Speaking	6.06	16.13	26.48	51.33
	Reading and Writing	19.66	35.91	24.87	19.56
	Overall	19.39	36.22	27.30	17.09
9	Listening and Speaking	10.67	25.59	20.65	43.09
	Reading and Writing	19.20	47.44	20.48	12.89
	Overall	20.01	47.98	20.97	11.04
10	Listening and Speaking	7.48	26.98	23.17	42.36
	Reading and Writing	16.60	46.94	23.34	13.12
	Overall	17.57	47.59	23.52	11.32
11	Listening and Speaking	4.17	23.94	29.26	42.64
	Reading and Writing	10.13	49.73	23.98	16.16
	Overall	10.88	51.01	24.86	13.26
12	Listening and Speaking	1.71	20.31	23.28	54.71
	Reading and Writing	6.07	52.46	23.17	18.30
	Overall	6.55	54.45	23.75	15.25

Table 49 - Percentage of Students in Each Proficiency Level by Grade Span

Grade	Test	Proficiency Levels			
		Beginning	Intermediate	Advanced	Proficient
K-1	Listening and Speaking	4.22	19.20	41.33	35.25
	Reading and Writing	32.02	22.42	20.19	25.38
	Overall	31.11	25.50	27.14	16.25
2-4	Listening and Speaking	2.26	6.95	31.24	59.54
	Reading and Writing	10.95	27.96	40.08	21.01
	Overall	10.72	28.31	42.45	18.52
5-6	Listening and Speaking	6.51	11.77	42.17	39.56
	Reading and Writing	14.23	27.34	37.79	20.64
	Overall	14.01	28.59	42.79	14.61
7-8	Listening and Speaking	5.75	14.63	28.39	51.24
	Reading and Writing	19.55	34.18	28.79	17.48
	Overall	19.29	34.49	30.99	15.23
9-12	Listening and Speaking	7.02	24.91	23.64	44.43
	Reading and Writing	14.49	48.53	22.51	14.47
	Overall	15.27	49.48	23.02	12.24

Note: The statistics in Tables 47 and 48 are based on the final research file. The total n-counts for Grades K-12 are 218,077. The stats presented in the table are based on the n-counts after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

Table 50 - Exit Rate by Grade for 2007, 2008, 2009, and 2010 Administrations

Grade	2007		2008		2009		2010	
	N-Count	Percent	N-Count	Percent	N-Count	Percent	N-Count	Percent
K	1,069	4.41	1,191	4.94	1,663	7.04	2,515	9.89
1	3,265	12.62	4,132	16.05	5,148	19.02	6,059	22.17
2	2,514	11.53	3,256	14.12	2,875	12.80	3,071	13.92
3	3,567	18.51	3,708	19.12	3,455	16.75	3,982	19.05
4	3,486	21.69	3,314	19.88	3,386	19.49	4,446	23.24
5	1,935	14.72	2,442	17.71	3,489	23.64	2,624	17.00
6	1,447	12.63	1,890	15.52	2,172	17.45	1,545	11.79
7	866	7.98	1,192	10.10	1,873	16.29	1,675	13.40
8	1,473	12.77	1,608	14.57	2,607	22.71	2,104	17.09
9	1,589	12.23	1,612	12.29	1,896	15.46	1,470	11.04
10	1,222	10.66	1,142	9.68	1,644	13.83	1,407	11.32
11	725	10.58	897	11.60	1,032	12.80	1,187	13.26
12	407	12.45	726	14.51	836	14.78	917	15.25
Total	23,565	12.48	27,110	13.85	32,076	16.10	33,002	15.80

Table 51 - Exit Rate by Grade Span for 2007, 2008, 2009, and 2010 Administrations

Grade	2007		2008		2009		2010	
	N-Count	Percent	N-Count	Percent	N-Count	Percent	N-Count	Percent
K-1	4,334	8.65	5,323	10.68	6,811	13.44	8,574	16.25
2-4	9,567	16.74	10,278	17.38	9,716	16.07	11,499	18.52
5-6	3,382	13.75	4,332	16.68	5,661	20.81	4,169	14.61
7-8	2,339	10.45	2,800	12.26	4,480	19.50	3,779	15.23
9-12	3,943	11.40	4,377	11.63	5,408	14.28	4,981	12.24
Total	23,565	12.48	27,110	13.85	32,076	16.10	33,002	15.80

APPENDIX A: ITEM MAP FOR 2010 NYSESLAT

A. 1: 2010 NYSESLAT Operational Test Information Grades K–1

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
Speaking					
1	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Sentence Completion
2	SS	2	2	Make predictions and inferences, and discuss the meaning of literary works to understand text presented orally and in written form	Sentence Completion
3	SS	2	5	Share cross-cultural experiences and ideas with others	Sentence Completion
4	SS	2	3	Apply learning strategies to explore a variety of materials	Sentence Completion
5	SS	2	5	Learn about and demonstrate an appreciation of some commonalities and distinctions across cultures and generations, including students' own	Sentence Completion
6	SE	4	2	Create personal stories using appropriate vocabulary and elements of the literature students have read or heard	Storytelling
7	SS	2	1	Formulate and share opinions about information and ideas with reference to features in oral and written text such as details and facts	Picture Description
8	SS	2	3	Form and express responses to ideas through reading, listening, viewing, discussing, and writing	Picture Description
9	SS	2	3	Form and express responses to ideas through reading, listening, viewing, discussing, and writing	Picture Description
10	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Picture Description
11	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Picture Description
12	SS	2	4	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction
13	SS	2	3	Recognize personal point of view in self and others in discussing information	Social Interaction
14	SS	2	5	Share cross-cultural experiences and ideas with others	Social Interaction
15	SS	2	4	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction

A.1: 2010 NYSESLAT Operational Test Information Grades K–1 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
16	SS	2	4	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction
Listening					
1	MC	1	1	Become familiar with some conventions of American English	Word/sentence comprehension
2	MC	1	1	Become familiar with some conventions of American English	Word/sentence comprehension
3	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word/sentence comprehension
4	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word/sentence comprehension
5	MC	1	3	Apply learning strategies to explore a variety of materials	Word/sentence comprehension
6	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word/sentence comprehension
7	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word/sentence comprehension
8	MC	1	5	Demonstrate familiarity with some U.S. cultural referents at the local and national levels	Word/sentence comprehension
9	MC	1	1	Select information appropriate to the purpose of the investigation	Word/sentence comprehension
10	MC	1	3	Form and express responses to ideas through reading, listening, viewing, discussing, and writing	Word/sentence comprehension
11	MC	1	3	Apply learning strategies to explore a variety of materials	Word/sentence comprehension
12	MC	1	1	Formulate and share opinions about information and ideas with reference to features in oral and written text such as details and facts	Word/sentence comprehension
13	MC	1	1	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/sentence comprehension
14	MC	1	3	Form and express responses to ideas through reading, listening, viewing, discussing, and writing	Word/sentence comprehension
15	MC	1	2	Make predictions and inferences, and discuss the meaning of literary works to understand text presented orally and in written form	Comprehension of conversational language

A.1: 2010 NYSESLAT Operational Test Information Grades K–1 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
16	MC	1	2	Use basic reading and listening strategies to make literary text comprehensible and meaningful	Comprehension of conversational language
17	MC	1	4	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Comprehension of conversational language
18	MC	1	5	Share cross-cultural experiences and ideas with others	Comprehension of conversational language
19	MC	1	4	Explain actions, choices, and decisions in social and classroom situations	Comprehension of conversational language
20	MC	1	3	Apply learning strategies to explore a variety of materials	Comprehension of conversational language
21	MC	1	2	Make predictions and inferences, and discuss the meaning of literary works to understand text presented orally and in written form	Comprehension of conversational language
22	MC	1	5	Demonstrate familiarity with some U.S. cultural referents at the local and national levels	Comprehension of conversational language
23	MC	1	1	Convey information using some organizational patterns and structures	Comprehension of conversational language
24	MC	1	3	Apply learning strategies to explore a variety of materials	Comprehension of conversational language
Reading					
1	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
2	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
3	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
4	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading

A.1: 2010 NYSESLAT Operational Test Information Grades K–1 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
5	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
6	MC	1	1	Formulate, ask, and respond to questions to obtain and provide information and meaning	Word Reading
7	MC	1	5	Learn about and demonstrate an appreciation of some commonalities and distinctions across cultures and generations, including students' own	Word Reading
8	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
9	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
10	MC	1	1	Select information appropriate to the purpose of the investigation	Word Reading
11	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
12	MC	1	1	Select information appropriate to the purpose of the investigation	Sentence Reading
13	MC	1	1	Become familiar with some conventions of American English	Sentence Reading
14	MC	1	5	Demonstrate familiarity with some cultural and language patterns and norms in American English	Short Passage with Question
15	MC	1	3	Recognize how structural features affect readers' and listeners' understanding and appreciation of text	Short Passage with Question
Writing					
1	MC	1	1	Become familiar with some conventions of American English	Phonetic Understanding
2	MC	1	1	Become familiar with some conventions of American English	Phonetic Understanding
3	MC	1	1	Become familiar with some conventions of American English	Phonetic Understanding
4	MC	1	1	Become familiar with some conventions of American English	Phonetic Understanding
5	MC	1	1	Become familiar with some conventions of American English	Phonetic Understanding
6	MC	1	1	Become familiar with some conventions of American English	Phonetic Understanding
7	CR	2	1	Present ideas clearly in written form	Developmental Writing

A.1: 2010 NYSESLAT Operational Test Information Grades K–1 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
8	CR	2	1	Present ideas clearly in written form	Developmental Writing
9	CR	2	1	Express and develop ideas and understanding, using some elements of the "writing process"	Developmental Writing
10	CR	2	1	Express and develop ideas and understanding, using some elements of the "writing process"	Developmental Writing
11	CR	2	1	Become familiar with some conventions of American English	Developmental Writing
12	CR	2	1	Become familiar with some conventions of American English	Developmental Writing
13	CR	2	1	Express and develop ideas and understanding, using some elements of the "writing process"	Developmental Writing
14	CR	2	1	Express and develop ideas and understanding, using some elements of the "writing process"	Developmental Writing
15	CR	2	2	Create personal stories using appropriate vocabulary and elements of the literature students have read or heard	Extended Response

A.1: 2010 NYSESLAT Operational Test Information Grades 2–4

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
Speaking					
1	SS	2	1	Formulate, ask, and respond to questions to obtain, clarify, and extend information and meaning	Sentence Completion
2	SS	2	1	Make and support inferences about information and ideas with reference to features in oral and written text	Sentence Completion
3	SS	2	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Sentence Completion
4	SS	2	3	Recognize personal point of view in self and others in discussing, interpreting, and evaluating information	Sentence Completion
5	SS	2	1	Convey information using a variety of organizational patterns and structures	Sentence Completion
6	SE	4	2	Create personal stories, poems, and songs, including those that reflect traditional and popular American culture; use appropriate vocabulary and elements of the literature students have read or heard	Storytelling
7	SS	2	1	Make and support inferences about information and ideas with reference to features in oral and written text	Picture Description
8	SS	2	3	Apply learning strategies to examine, interpret, and evaluate a variety of materials	Picture Description
9	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Picture Description
10	SS	2	5	Interpret and demonstrate knowledge of nonverbal communication, and understand the contexts in which they are used appropriately	Picture Description
11	SS	2	3	Recognize personal point of view in self and others in discussing, interpreting, and evaluating information	Picture Description
12	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Social Interaction
13	SS	2	3	Recognize personal point of view in self and others in discussing, interpreting, and evaluating information	Social Interaction
14	SS	2	4	Use appropriate vocabulary, language, and interaction styles for various audiences and formal or informal social or school situations	Social Interaction

A.1: 2010 NYSESLAT Operational Test Information Grades 2–4 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
15	SS	2	4	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction
16	SS	2	4	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction
Listening					
1	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Word/sentence comprehension
2	MC	1	1	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/sentence comprehension
3	MC	1	1	Select information appropriate to the purpose of the investigation, and relate ideas from one written or spoken source to another	Word/sentence comprehension
4	MC	1	1	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/sentence comprehension
5	MC	1	1	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/sentence comprehension
6	MC	1	1	Read, gather, view, listen to, organize, discuss, and interpret information related to academic content areas from various sources	Word/sentence comprehension
7	MC	1	2	Make predictions, inferences, and deductions, and discuss the meaning of literary works with some attention to meaning beyond the literal level, to understand and interpret text presented orally and in written form	Word/sentence comprehension
8	MC	1	1	Distinguish between fact and opinion and relevant and irrelevant information	Word/sentence comprehension
9	MC	1	1	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Word/sentence comprehension
10	MC	1	2	Recognize some features that distinguish some genres and use those features to aid comprehension	Word/sentence comprehension
11	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language
12	MC	1	1	Formulate, ask, and respond to questions to obtain, clarify, and extend information and meaning	Comprehension of conversational language

A.1: 2010 NYSESLAT Operational Test Information Grades 2–4 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
13	MC	1	5	Recognize and demonstrate an appreciation of some commonalities and distinctions across cultures and groups (differentiated by gender, ability, generations, etc.)	Comprehension of conversational language
14	MC	1	2	Locate and identify key literary elements in texts and relate those elements to those in other works and to students' own experiences	Comprehension of conversational language
15	MC	1	4	Request and provide information and assistance orally or in writing, for personal, social, and academic purposes	Comprehension of conversational language
16	MC	1	1	Distinguish between fact and opinion and relevant and irrelevant information	Comprehension of conversational language
17	MC	1	2	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language
18	MC	1	4	Negotiate and manage interactions to accomplish social classroom tasks	Comprehension of conversational language
19	MC	1	1	Select information appropriate to the purpose of the investigation, and relate ideas from one written or spoken source to another	Comprehension of conversational language
20	MC	1	4	Follow oral and written directions to participate in classroom and social activities	Task-based
21	MC	1	4	Explain actions, choices, and decisions in social and classroom situations	Task-based
22	MC	1	1	Select information appropriate to the purpose of the investigation, and relate ideas from one written or spoken source to another	Task-based
23	MC	1	1	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Task-based
24	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Task-based
Reading					
1	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word/Sentence Reading
2	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Word/Sentence Reading

A.1: 2010 NYSESLAT Operational Test Information Grades 2–4 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
3	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Word/Sentence Reading
4	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Word/Sentence Reading
5	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word/Sentence Reading
6	MC	1	1	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Word/Sentence Reading
7	MC	1	1	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/Sentence Reading
8	MC	1	1	Formulate, ask, and respond to questions to obtain, clarify, and extend information and meaning	Comprehension
9	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
10	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Comprehension
11	MC	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Comprehension
12	MC	1	2	Make predictions, inferences, and deductions, and discuss the meaning of literary works with some attention to meaning beyond the literal level, to understand and interpret text presented orally and in written form	Comprehension
13	MC	1	2	Apply learning strategies to comprehend literature and make inferences about literature and produce literary responses	Comprehension
14	MC	1	2	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Comprehension
15	MC	1	2	Make predictions, inferences, and deductions, and discuss the meaning of literary works with some attention to meaning beyond the literal level, to understand and interpret text presented orally and in written form	Comprehension
16	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension

A.1: 2010 NYSESLAT Operational Test Information Grades 2–4 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
17	MC	1	1	Select information appropriate to the purpose of the investigation, and relate ideas from one written or spoken source to another	Comprehension
18	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
19	MC	1	3	Recognize and explain how structural features affect readers' and listeners' understanding and appreciation of text	Comprehension
20	MC	1	2	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Comprehension
21	MC	1	2	Apply learning strategies to comprehend literature and make inferences about literature and produce literary responses	Comprehension
22	MC	1	2	Locate and identify key literary elements in texts and relate those elements to those in other works and to students' own experiences	Comprehension
23	MC	1	3	Form and express responses to a variety of literary, informational, and persuasive materials through reading, listening, viewing, discussing, and writing; use details and evidence as support	Comprehension
24	MC	1	3	Recognize personal point of view in self and others in discussing, interpreting, and evaluating information	Comprehension
Writing					
1	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Phonetic Understanding
2	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Phonetic Understanding
3	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Phonetic Understanding
4	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Phonetic Understanding

A.1: 2010 NYSESLAT Operational Test Information Grades 2–4 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
5	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Mechanics and Structure
6	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Mechanics and Structure
7	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Mechanics and Structure
8	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Mechanics and Structure
9	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Mechanics and Structure
10	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Mechanics and Structure
11	MC	1	3	Form and express responses to a variety of literary, informational, and persuasive materials through reading, listening, viewing, discussing, and writing; use details and evidence as support	Mechanics and Structure
12	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English	Mechanics and Structure
13	CR	2	3	Recognize personal point of view in self and others in discussing, interpreting and evaluating information	Pre-writing
14	CR	2	3	Form and express responses to a variety of literary, informational, and persuasive materials through reading, listening, viewing, discussing, and writing; use details and evidence as support	Pre-writing
15	CR	2	3	Form and express responses to a variety of literary, informational, and persuasive materials through reading, listening, viewing, discussing, and writing; use details and evidence as support	Pre-writing
16	ER	4	3	Form and express responses to a variety of literary, informational, and persuasive materials through reading, listening, viewing, discussing, and writing; use details and evidence as support	Extended Response

A.1: 2010 NYSESLAT Operational Test Information Grades 5–6

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
Speaking					
1	SS	2	1	Make and support inferences about information and ideas with reference to features in oral and written text	Sentence Completion
2	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Sentence Completion
3	SS	2	1	Make and support inferences about information and ideas with reference to features in oral and written text	Sentence Completion
4	SS	2	1	Formulate, ask, and respond to various question forms to obtain, clarify, and extend information and meaning	Sentence Completion
5	SS	2	1	Make and support inferences about information and ideas with reference to features in oral and written text	Sentence Completion
6	SE	4	2	Create stories, poems, sketches, songs, and plays, including those that reflect traditional and popular American culture, using typical features of a given genre; create an effective voice using a variety of writing styles appropriate to different audio	Storytelling
7	SS	2	1	Formulate, ask, and respond to various question forms to obtain, clarify, and extend information and meaning	Picture Description
8	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Picture Description
9	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Picture Description
10	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Picture Description
11	SS	2	1	Convey and organize information, using facts, details, illustrative examples, and a variety of patterns and structures	Picture Description
12	SS	2	4	Listen attentively and take turns speaking when engaged in pair, group, or full-class discussions on personal, social, community, and academic topics	Social Interaction
13	SS	2	3	Read, listen to, view, write about, and discuss texts and performances from a wide range of authors, subjects, and genres	Social Interaction
14	SS	2	5	Recognize and share cross-cultural experiences and ideas and connect with those of others	Social Interaction
15	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Social Interaction

A.1: 2010 NYSESLAT Operational Test Information Grades 5–6 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
16	SS	2	2	Read, listen to, view, write about, and discuss texts and performances from a wide range of authors, subjects, and genres	Social Interaction
Listening					
1	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/sentence comprehension
2	MC	1	1	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Word/sentence comprehension
3	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/sentence comprehension
4	MC	1	4	Negotiate and manage interactions to accomplish social and classroom tasks	Comprehension of conversational language
5	MC	1	1	Select information appropriate to the purpose of the investigation, relate ideas from one written or spoken source to another and exclude nonessential information	Comprehension of conversational language
6	MC	1	5	Demonstrate an understanding of cultural and language patterns and norms in American English, including different regional and social varieties of English	Comprehension of conversational language
7	MC	1	1	Distinguish between fact and opinion and relevant and irrelevant information, and excluded nonessential information in oral and written presentations	Comprehension of conversational language
8	MC	1	1	Select information appropriate to the purpose of the investigation, relate ideas from one written or spoken source to another and exclude nonessential information	Comprehension of conversational language
9	MC	1	4	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Comprehension of conversational language
10	MC	1	2	Read, listen to, view, write about, and discuss texts and performances from a wide range of authors, subjects, and genres	Comprehension of conversational language
11	MC	1	1	Select information appropriate to the purpose of the investigation, relate ideas from one written or spoken source to another and exclude nonessential information	Comprehension of conversational language
12	MC	1	3	Read, listen to, view, write about, and discuss texts and performances from a wide range of authors, subjects, and genres	Comprehension of conversational language

A.1: 2010 NYSESLAT Operational Test Information Grades 5–6 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
13	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language
14	MC	1	5	Demonstrate an understanding of cultural and language patterns and norms in American English, including different regional and social varieties of English	Comprehension of conversational language
15	MC	1	2	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language
16	MC	1	4	Follow oral and written directions to participate in classroom and social activities and provide directions to peers in selected interactions	Task-based
17	MC	1	4	Follow oral and written directions to participate in classroom and social activities and provide directions to peers in selected interactions	Task-based
18	MC	1	4	Follow oral and written directions to participate in classroom and social activities and provide directions to peers in selected interactions	Task-based
19	MC	1	4	Follow oral and written directions to participate in classroom and social activities and provide directions to peers in selected interactions	Task-based
20	MC	1	1	Select information appropriate to the purpose of the investigation, relate ideas from one written or spoken source to another and exclude nonessential information	Task-based
21	MC	1	1	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Task-based
22	MC	1	4	Request and provide information and assistance orally or in writing, for personal, social, and academic purposes	Task-based
23	MC	1	4	Request and provide information and assistance orally or in writing, for personal, social, and academic purposes	Task-based
24	MC	1	1	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Task-based
25	MC	1	4	Request and provide information and assistance orally or in writing, for personal, social, and academic purposes	Task-based
Reading					
1	MC	1	1	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Word/Sentence Reading

A.1: 2010 NYSESLAT Operational Test Information Grades 5–6 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
2	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/Sentence Reading
3	MC	1	1	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Word/Sentence Reading
4	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
5	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
6	MC	1	3	Apply learning strategies to examine and interpret a variety of materials	Comprehension
7	MC	1	1	Select information appropriate to the purpose of the investigation, relate ideas from one written or spoken source to another and exclude nonessential information	Comprehension
8	MC	1	3	Apply learning strategies to examine and interpret a variety of materials	Comprehension
9	MC	1	2	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
10	MC	1	1	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Comprehension
11	MC	1	2	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
12	MC	1	3	Identify and explain the distinguishing features of the major genres, and use those features to aid comprehension, interpretation, and discussion of literature	Comprehension
13	MC	1	2	Make predictions, inferences, and deductions, and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meanings	Comprehension
14	MC	1	2	Make predictions, inferences, and deductions, and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meanings	Comprehension
15	MC	1	2	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
16	MC	1	1	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Comprehension

A.1: 2010 NYSESLAT Operational Test Information Grades 5–6 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
17	MC	1	2	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
18	MC	1	2	Locate and identify selected literary elements and techniques in texts and relate those elements to those in other works and to students' own experiences	Comprehension
19	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
20	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
21	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
22	MC	1	1	Formulate, ask, and respond to various question forms to obtain, clarify, and extend information and meaning	Comprehension
23	MC	1	1	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Comprehension
24	MC	1	2	Make predictions, inferences, and deductions, and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meanings	Comprehension
25	MC	1	2	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
26	MC	1	1	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Comprehension
27	MC	1	2	Locate and identify selected literary elements and techniques in texts and relate those elements to those in other works and to students' own experiences	Comprehension
Writing					
1	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Phonetic Understanding
2	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Phonetic Understanding
3	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Mechanics and Structure

A.1: 2010 NYSESLAT Operational Test Information Grades 5–6 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
4	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
5	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
6	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
7	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
8	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
9	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
10	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
11	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
12	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
13	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure

A.1: 2010 NYSESLAT Operational Test Information Grades 5–6 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
14	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
15	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
16	CR	2	3	Identify and explain the distinguishing features of the major genres, and use those features to aid comprehension, interpretation, and discussion of literature	Pre-writing
17	CR	2	3	Apply learning strategies to examine and interpret a variety of materials	Pre-writing
18	CR	2	3	Apply learning strategies to examine and interpret a variety of materials	Pre-writing
19	ER	4	3	Read, listen to, view, write about, and discuss texts and performances from a wide range of authors, subjects, and genres	Extended Response

A.1: 2010 NYSESLAT Operational Test Information Grades 7–8

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
Speaking					
1	SS	2	1	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Sentence Completion
2	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Sentence Completion
3	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Sentence Completion
4	SS	2	5	Recognize and demonstrate an appreciation of some commonalities and distinctions across cultures and groups (differentiated by gender, ability, generations, etc.) including the students' own	Sentence Completion
5	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Sentence Completion
6	SE	4	2	Create stories, poems, sketches, songs, and plays, including, those that reflect traditional and popular American culture, using typical features of a given genre; create an effective voice, using a variety of writing styles appropriate to different audio	Storytelling
7	SS	2	4	Use appropriate vocabulary, expressions, language, routines, and interactions styles for various audiences and formal and informal social or school situations, noticing how intention is realized through language	Picture Description
8	SS	2	1	Formulate, ask, and respond to various question forms to obtain, clarify, and extend information and meaning	Social Interaction
9	SS	2	3	Read, listen to, view, write about, and discuss texts and performances from a wide range of authors, subjects, and genres	Picture Description
10	SS	2	3	Apply learning strategies to examine and interpret a variety of materials	Picture Description
11	SS	2	5	Recognize and share cross-cultural experiences and ideas and connect with those of others	Picture Description
12	SS	2	4	Use appropriate vocabulary, expressions, language, routines, and interactions styles for various audiences and formal and informal social or school situations, noticing how intention is realized through language	Social Interaction
13	SS	2	4	Request and provide information and assistance orally or in writing, for personal, social, and academic purposes	Picture Description
14	SS	2	3	Read, listen to, view, write about, and discuss texts and performances from a wide range of authors, subjects, and genres	Social Interaction

A.1: 2010 NYSESLAT Operational Test Information Grades 7–8 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
15	SS	2	4	Listen attentively and take turns speaking when engaged in pair, group, or full-class discussions on personal, social, community, and academic topics	Picture Description
16	SS	2	4	Explain actions, choices, and decisions in social and classroom situations	Social Interaction
Listening					
1	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/sentence comprehension
2	MC	1	3	Apply learning strategies to examine and interpret a variety of materials	Word/sentence comprehension
3	MC	1	5	Recognize and share cross-cultural experiences and ideas and connect with those of others	Word/sentence comprehension
4	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language
5	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language
6	MC	1	4	Request and provide information and assistance orally or in writing, for personal, social, and academic purposes	Comprehension of conversational language
7	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension of conversational language
8	MC	1	4	Request and provide information and assistance orally or in writing, for personal, social, and academic purposes	Comprehension of conversational language
9	MC	1	1	Formulate, ask, and respond to various question forms to obtain, clarify, and extend information and meaning	Comprehension of conversational language
10	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension of conversational language
11	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language
12	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension of conversational language
13	MC	1	1	Select information appropriate to the purpose of the investigation, relate ideas from one written or spoken source to another and exclude nonessential information	Comprehension of conversational language
14	MC	1	3	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language

A.1: 2010 NYSESLAT Operational Test Information Grades 7–8 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
15	MC	1	5	Demonstrate an understanding of cultural and language patterns and norms in American English, including different regional and social varieties of English	Comprehension of conversational language
16	MC	1	4	Follow oral and written directions to participate in classroom and social activities and provide directions to peers in selected interactions	Task-based
17	MC	1	4	Follow oral and written directions to participate in classroom and social activities and provide directions to peers in selected interactions	Task-based
18	MC	1	4	Follow oral and written directions to participate in classroom and social activities and provide directions to peers in selected interactions	Task-based
19	MC	1	1	Distinguish between fact and opinion and relevant and irrelevant information, and excluded nonessential information in oral and written presentations	Task-based
20	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Task-based
21	MC	1	1	Convey and organize information, using facts, details, illustrative examples, and a variety of patterns and structures	Task-based
22	MC	1	4	Request and provide information and assistance orally or in writing, for personal, social, and academic purposes	Task-based
23	MC	1	1	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Task-based
24	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Task-based
25	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Task-based
Reading					
1	MC	1	3	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/Sentence Reading
2	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Word/Sentence Reading
3	MC	1	1	Read, gather, view, listen to, organize, discuss, and interpret information related to academic content areas from various sources	Word/Sentence Reading

A.1: 2010 NYSESLAT Operational Test Information Grades 7–8 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
4	MC	1	1	Select information appropriate to the purpose of the investigation, relate ideas from one written or spoken source to another and exclude nonessential information	Comprehension
5	MC	1	3	Make predictions, inferences, and deductions, and discuss the meaning of literary works with some attention to meaning beyond the literal level, to understand and interpret text presented orally, and in written form	Comprehension
6	MC	1	5	Demonstrate an understanding of a broad range of U.S. cultural and political referents through institutions, functions, and processes at the local and national levels, and compare/contrast these with parallels in students' native community	Comprehension
7	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
8	MC	1	3	Apply learning strategies to examine and interpret a variety of materials	Comprehension
9	MC	1	1	Convey and organize information using facts, details, illustrative examples, and a variety of patterns and structures	Comprehension
10	MC	1	1	Convey and organize information using facts, details, illustrative examples, and a variety of patterns and structures	Comprehension
11	MC	1	1	Convey and organize information using facts, details, illustrative examples, and a variety of patterns and structures	Comprehension
12	MC	1	1	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Comprehension
13	MC	1	1	Select information appropriate to the purpose of the investigation, relate ideas from one written or spoken source to another and exclude nonessential information	Comprehension
14	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
15	MC	1	3	Read, listen to, view, write about, and discuss texts and performances from a wide range of authors, subjects, and genres	Comprehension
16	MC	1	1	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Comprehension
17	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
18	MC	1	3	Apply learning strategies to examine and interpret a variety of materials	Comprehension

A.1: 2010 NYSESLAT Operational Test Information Grades 7–8 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
19	MC	1	2	Make predictions, inferences, and deductions, and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meanings	Comprehension
20	MC	1	2	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
21	MC	1	2	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
22	MC	1	2	Make predictions, inferences, and deductions, and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meanings	Comprehension
23	NS	1	2	Make predictions, inferences, and deductions, and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meanings	Comprehension
24	MC	1	2	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
25	MC	1	2	Make predictions, inferences, and deductions, and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meanings	Comprehension
26	MC	1	2	Make predictions, inferences, and deductions, and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meanings	Comprehension
27	MC	1	2	Locate and identify selected literary elements and techniques in texts and relate those elements to those in other works and to students' own experiences	Comprehension
Writing					
1	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Phonetic Understanding
2	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Phonetic Understanding
3	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Mechanics and Structure

A.1: 2010 NYSESLAT Operational Test Information Grades 7–8 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
4	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
5	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
6	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
7	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
8	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
9	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
10	MC	1	1	Apply self-monitoring and self-correcting strategies for accurate language production and oral and written presentation using established criteria for effective presentation of information	Mechanics and Structure
11	MC	1	1	Apply self-monitoring and self-correcting strategies for accurate language production and oral and written presentation using established criteria for effective presentation of information	Mechanics and Structure
12	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
13	MC	1	1	Apply self-monitoring and self-correcting strategies for accurate language production and oral and written presentation using established criteria for effective presentation of information	Mechanics and Structure
14	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of America English appropriate to audience and purpose	Mechanics and Structure
15	MC	1	1	Apply self-monitoring and self-correcting strategies for accurate language production and oral and written presentation using established criteria for effective presentation of information	Mechanics and Structure

A.1: 2010 NYSESLAT Operational Test Information Grades 7–8 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
16	CR	2	1	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Pre-writing
17	CR	2	3	Apply learning strategies to examine and interpret a variety of materials	Pre-writing
18	CR	2	3	Apply learning strategies to examine and interpret a variety of materials	Pre-writing
19	ER	4	3	Read, listen to, view, write about, and discuss texts and performances from a wide range of authors, subjects, and genres	Extended Response

A.1: 2010 NYSESLAT Operational Test Information Grades 9–12

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
Speaking					
1	SS	2	5	Recognize and share cross-cultural experiences and ideas and connect with those of others	Sentence Completion
2	SS	2	4	Explain actions, choices, and decisions in social and academic situations	Sentence Completion
3	SS	2	4	Explain actions, choices, and decisions in social and academic situations	Sentence Completion
4	SS	2	5	Demonstrate an understanding of a broad range of U.S. cultural and political referents through institutions, functions, and processes at the local and national levels, and compare/contrast these with parallels in students' native community	Sentence Completion
5	SS	2	5	Recognize and demonstrate an appreciation of some commonalities and distinctions across cultures and groups (differentiated by gender, ability, generations, etc.)	Sentence Completion
6	SE	4	2	Create stories, poems, sketches, songs, and plays, including, those that reflect traditional and popular American culture, using typical features of a given genre; create an effective voice, using a variety of writing styles appropriate to different audio	Storytelling
7	SS	2	4	Use appropriate vocabulary, expressions, language, routines, and interaction styles for various audiences	Picture Description
8	SS	2	1	Convey and organize information, using facts, details, illustrative examples, and a variety of patterns and structures	Picture Description
9	SS	2	1	Make and support inferences about information and ideas with reference to features in oral and written text	Picture Description
10	SS	2	4	Explain actions, choices, and decisions in social and academic situations	Picture Description
11	SS	2	3	Apply learning strategies to examine, interpret, and evaluate a variety of materials	Picture Description
12	SS	2	4	Request and provide information and assistance orally or in writing, for personal, social, and academic purposes	Social Interaction
13	SS	2	4	Explain actions, choices, and decisions in social and academic situations	Social Interaction
14	SS	2	4	Explain actions, choices, and decisions in social and academic situations	Social Interaction
15	SS	2	4	Describe, read about, participate in, or recommend a favorite activity, book, or song	Social Interaction

A.1: 2010 NYSESLAT Operational Test Information Grades 9–12 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
16	SS	2	4	Listen attentively, take turns speaking, and build on others' ideas when engaged in pair, group, or full-class discussions on personal, social, community, and academic topics	Social Interaction
Listening					
1	MC	1	1	Compare, contrast, categorize, and synthesize information and objects and identify complexities and discrepancies in the information	Word/sentence comprehension
2	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/sentence comprehension
3	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/sentence comprehension
4	MC	1	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information; make inferences	Comprehension of conversational language
5	MC	1	3	Apply learning strategies to examine, interpret, and evaluate a variety of materials	Comprehension of conversational language
6	MC	1	4	Follow oral and written directions to participate in classroom activities, and provide directions to peers in selected interactions	Comprehension of conversational language
7	MC	1	4	Follow oral and written directions to participate in classroom activities, and provide directions to peers in selected interactions	Comprehension of conversational language
8	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension of conversational language
9	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language
10	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language
11	MC	1	4	Follow oral and written directions to participate in classroom activities, and provide directions to peers in selected interactions	Comprehension of conversational language
12	MC	1	1	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Comprehension of conversational language
13	MC	1	3	Apply learning strategies to examine, interpret, and evaluate a variety of materials	Comprehension of conversational language

A.1: 2010 NYSESLAT Operational Test Information Grades 9–12 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
14	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of conversational language
15	MC	1	1	Read, gather, view, listen to, organize, discuss, and interpret information related to academic content areas from various sources	Comprehension of conversational language
16	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based
17	MC	1	4	Follow oral and written directions to participate in classroom activities, and provide directions to peers in selected interactions	Task-based
18	MC	1	1	Compare, contrast, categorize, and synthesize information and objects and identify complexities and discrepancies in the information	Task-based
19	MC	1	4	Follow oral and written directions to participate in classroom activities, and provide directions to peers in selected interactions	Task-based
20	MC	1	3	Recognize, explain, evaluate, and analyze how structural features affect readers' and listeners' understanding and appreciation of text	Task-based
21	MC	1	1	Select information appropriate to the purpose of the investigation with suitable supporting material	Task-based
22	MC	1	4	Negotiate and manage interactions to accomplish social and classroom tasks	Task-based
23	MC	1	4	Follow oral and written directions to participate in classroom activities, and provide directions to peers in selected interactions	Task-based
24	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based
25	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based
Reading					
1	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/Sentence Reading
2	MC	1	1	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Word/Sentence Reading
3	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/Sentence Reading

A.1: 2010 NYSESLAT Operational Test Information Grades 9–12 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
4	MC	1	5	Demonstrate an understanding of a broad range of U.S. cultural and political referents through institutions, functions, and processes at the local and national levels, and compare/contrast these with parallels in students' native community	Comprehension
5	MC	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
6	MC	1	5	Demonstrate an understanding of a broad range of U.S. cultural and political referents through institutions, functions, and processes at the local and national levels, and compare/contrast these with parallels in students' native community	Comprehension
7	MC	1	1	Distinguish between fact and opinions, and opinion and relevant and irrelevant information and exclude nonessential information in oral and written presentations	Comprehension
8	MC	1	2	Identify and explain the distinguishing features of different literary genres, periods, and traditions, and use those features to aid comprehension, interpretation, and discussion of literature	Comprehension
9	MC	1	5	Compare and contrast oral traditions, myths, folktales, and literature from the United States and international regions and cultures, including the students' own, and identify similarities and differences and universal cultural themes, and exploring how	Comprehension
10	MC	1	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information; make inferences	Comprehension
11	MC	1	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information; make inferences	Comprehension
12	MC	1	5	Recognize and share cross-cultural experiences and ideas and connect with those of others	Comprehension
13	MC	1	3	Recognize, explain, evaluate, and analyze how structural features affect readers' and listeners' understanding and appreciation of text	Comprehension

A.1: 2010 NYSESLAT Operational Test Information Grades 9–12 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
14	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
15	MC	1	1	Formulate, ask, and respond to various questions, forms to obtain, clarify and extend information and meaning	Comprehension
16	MC	1	2	Locate and identify a range of significant literary elements and techniques in texts and use those elements to interpret the work comparing and contrasting the work to other works and to students' own experiences	Comprehension
17	MC	1	3	Recognize, explain, evaluate, and analyze how structural features affect readers' and listeners' understanding and appreciation of text	Comprehension
18	MC	1	3	Assess, compare, and evaluate the quality of spoken or written texts and visual presentations, using different criteria related to the organization, subject area, and purpose of text	Comprehension
19	MC	1	2	Make prediction, inferences, and deductions and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meaning	Comprehension
20	MC	1	2	Make prediction, inferences, and deductions and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meaning	Comprehension
21	MC	1	2	Make prediction, inferences, and deductions and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meaning	Comprehension
22	MC	1	2	Make prediction, inferences, and deductions and describe different levels of meaning of literary works presented orally and in written form, including literal and implied meaning	Comprehension
23	MC	1	2	Apply self-monitoring and self-correcting strategies while reading, viewing, discussing, listening to, or producing literary texts and essays	Comprehension
24	MC	1	2	Apply reading and listening strategies to make literary text comprehensible and meaningful	Comprehension

A.1: 2010 NYSESLAT Operational Test Information Grades 9–12 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
25	MC	1	2	Apply learning strategies to comprehend, make inferences about, and analyze literature and to produce literary responses	Comprehension
26	MC	1	2	Identify and explain the distinguishing features of different literary genres, periods, and traditions, and use those features to aid comprehension, interpretation, and discussion of literature	Comprehension
27	MC	1	2	Apply reading and listening strategies to make literary text comprehensible and meaningful	Comprehension
Writing					
1	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Phonetic Understanding
2	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Phonetic Understanding
3	MC	1	4	Follow oral and written directions to participate in classroom activities, and provide directions to peers in selected interactions	Mechanics and Structure
4	MC	1	5	Demonstrate an understanding of cultural and language patterns and norms in American English, including different regional and social varieties of English, and identify and interpret how these patterns and norms are used	Mechanics and Structure
5	MC	1	3	Assess, compare, and evaluate the quality of spoken or written texts and visual presentations, using different criteria related to the organization, subject area, and purpose of text	Mechanics and Structure
6	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Mechanics and Structure
7	MC	1	3	Recognize, explain, evaluate, and analyze how structural features affect readers' and listeners' understanding and appreciation of text	Mechanics and Structure
8	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Mechanics and Structure

A.1: 2010 NYSESLAT Operational Test Information Grades 9–12 (continued)

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Cluster
9	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Mechanics and Structure
10	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Mechanics and Structure
11	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Mechanics and Structure
12	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Mechanics and Structure
13	MC	1	1	Make and support inferences about information and ideas with reference to features in oral and written text	Mechanics and Structure
14	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Mechanics and Structure
15	MC	1	1	Convey information and ideas through spoken and written language using conventions and features of American English appropriate to audience and purpose	Mechanics and Structure
16	CR	2	3	Speak and write, using the conventions and features of American English to effectively influence an audience (e.g., to persuade, negotiate, argue)	Pre-writing
17	CR	2	5	Recognize and demonstrate an appreciation of some commonalities and distinctions across cultures and groups (differentiated by gender, ability, generations, etc.)	Pre-writing
18	CR	2	5	Recognize and demonstrate an appreciation of some commonalities and distinctions across cultures and groups (differentiated by gender, ability, generations, etc.)	Pre-writing
19	ER	4	3	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts, and experiences; justify and explain the rationale for positions using persuasive language, tone, evidence, and well-developed arguments	Extended Response

A. 2: Item Specification**LISTENING: Word/Sentence Comprehension, Grades K–1**

The test taker must choose which of three letters represents the first sound in a word. These items measure phonemic discrimination and are designed especially to test students who are just beginning to learn the sound system in English.

Item Component	Description of Component
Stimulus	--
Stem	closed stem question: “What is the first sound in the word _____?”
Answer options	3 answer options consisting of 3 individual letter graphics
Graphic	--

LISTENING: Word/Sentence Comprehension, all grades

The test taker hears a question and must choose which of three small graphics represents the correct answer to the question. These items measure comprehension of words, phrases, and sentences, and are designed especially to test students at lower levels of proficiency.

Item Component	Description of Component
Stimulus	--
Stem	closed stem question, maximum 15 words: “Which picture shows _____?” ; “Which girl is _____?” ; “Which building is where _____?” ; etc.
Answer options	3 answer options consisting of 3 small individual graphics lettered A–C
Graphic	--

A.2: Item Specification (continued)**LISTENING: Comprehension of Conversational Language, all grades**

The test taker listens to someone speaking (a monologue) in a situation where a person would naturally be speaking without expecting a response from the listener, e.g., a teacher talking to a class, someone making an announcement over a loudspeaker, or someone briefly describing an event. These items measure comprehension of main ideas and details and the ability to make inferences. The language tested should simulate authentic conversational English. The language level should be at the upper level of the grade span.

Item Component	Description of Component
Stimulus	simulated authentic conversation with one person speaking; 20–30 words (Grades K–1), 40–50 words (Grades 2–4), 50–60 words (Grades 5–8), 55–70 words (Grades 9–12); there is one item per stimulus (Grades K–6) or two items per stimulus (Grades 7–12)
Stem	closed stem question, maximum 15 words
Answer options	3 answer options consisting of 3 small individual graphics lettered A–C (Grades K–4) or 4 text answer options of one to eight words (Grades 2–12; both formats are used in Grades 2–4)
Graphic	--

LISTENING: Synthesizing Information (Task-based Listening), Grades 2–12

The test taker is required to perform—through imagined manipulation of the graphics provided—real-world tasks. The test taker must comprehend discrete information and details and then synthesize the information to perform the tasks. The language tested should simulate authentic conversational English. The language level should be at the upper level of the grade span.

Item Component	Description of Component
Advance organizer	1–2 sentence description of an authentic context for the set of items; maximum 25 words; description will appear in test booklet with a large overview graphic depicting the scenario
Stimulus	simulated authentic conversation with one person speaking; 40–50 words (Grades 2–4), 50–60 words (Grades 5–8), 55–70 words (Grades 9–12), with distractor material embedded; there is one item per stimulus

LISTENING: Synthesizing Information (Task-based Listening), Grades 2–12

The test taker is required to perform—through imagined manipulation of the graphics provided— real-world tasks. The test taker must comprehend discrete information and details and then synthesize the information to perform the tasks. The language tested should simulate authentic conversational English. The language level should be at the upper level of the grade span.

Item Component	Description of Component
Stem	closed stem question, maximum 15 words

A.2: Item Specification (continued)**LISTENING: Synthesizing Information (Task-based Listening), Grades 2–12**

The test taker is required to perform—through imagined manipulation of the graphics provided—real-world tasks. The test taker must comprehend discrete information and details and then synthesize the information to perform the tasks. The language tested should simulate authentic conversational English. The language level should be at the upper level of the grade span.

Item Component	Description of Component
Answer options	4 answer options consisting of either 4 small individual graphics lettered A–D or 1 large graphic with 4 options lettered A–D
Graphic	large overview graphic depicting the scenario and containing elements of the items

READING: Word Reading, Grades K–1

The test taker reads a question, looks at a small graphic, and must choose which of three words correctly identifies the graphic. These items measure word recognition skills. This item type is especially designed to test students at lower levels of proficiency.

Item Component	Description of Component
Stimulus	graphic (below)
Stem	closed stem question, maximum 10 words: “What do you see?” and “What is the boy doing?” (Grades K–1); “Which word tells what is in the picture?” and “Which word tells what the girl is holding?” (Grades 2–4); etc.
Answer options	3 answer options consisting of 1 word each
Graphic	graphic depicting correct answer option

A.2: Item Specification (continued)**READING: Sentence Reading, Grades K–1**

The test taker reads a question that includes a descriptive phrase, or several descriptive sentences followed by a question, and must choose which of three graphics correctly represents what is described. These items measure basic phrase and sentence comprehension.

Item Component	Description of Component
Stimulus	(optional) description of context with distractor material embedded; maximum 25 words
Stem	closed stem question, maximum 10 words: “Which one tells you _____?”; “Which one can you _____?”; etc.
Answer options	3 answer options consisting of 3 small individual graphics
Graphic	--

READING: Short Passage with Question, Grades K–1

The test taker reads a brief passage and answers one question about the passage. These items measure comprehension of main ideas and details and the ability to make inferences.

Item Component	Description of Component
Stimulus	fictional, descriptive, informational, or functional passage; maximum 30 words
Stem	closed stem question, maximum 10 words
Answer options	3 answer options consisting of 3 small individual graphics
Graphic	--

A.2: Item Specification (continued)**READING: Word/Sentence Reading, Grades 2–12**

The test taker reads a question that includes a key word or a descriptive phrase, or several descriptive sentences followed by a question, and must choose which of three graphics correctly represents what is named or described. These items measure word-recognition skills and basic phrase and sentence comprehension.

Item Component	Description of Component
Stimulus	(optional) description of context with distractor material embedded; maximum 25 words
Stem	closed stem question, maximum 10 words: “Which picture shows _____?”; “Which person is _____?”; etc.
Answer options	3 answer options consisting of small individual graphics
Graphic	--

READING: Comprehension, Grades 2–12

The test taker reads a passage and answers several questions about the passage. These items measure comprehension of main ideas and details, the ability to make inferences, literary analysis skills, the ability to understand how and why information is organized in a passage, and other typical reading comprehension skills.

Item Component	Description of Component
Stimulus	fictional, descriptive, informational, or functional passage; maximum 200 words (Grades 2–4), 250 words (Grades 5–8), or 350 words (Grades 9–12)
Stem	closed stem question, maximum 20 words
Answer options	4 answer options consisting of 1 word, phrase, or sentence each, maximum 15 words
Graphic	--

A.2: Item Specification (continued)**WRITING CONVENTIONS: Phonemic Understanding, Grades K–1**

The test taker listens to a question, looks at a small graphic, and must choose which of three letters or letter blends correctly relates to the graphic. These items measure understanding of phonemes as they relate to writing. This item type is especially designed to test students at beginning levels of proficiency, who often read and write based on phonics rules.

Item Component	Description of Component
Stimulus	--
Stem	closed stem question: “Which letter(s) begin(s)/end(s) the word _____?”
Answer options	3 answer options consisting of 1 letter each or 1 letter blend each
Graphic	graphic depicting the word that begins or ends with the correct answer option

WRITING CONVENTIONS: Phonemic Understanding, Grades 2–12

The test taker reads a question, looks at a small graphic, and must choose which of three words correctly identifies the graphic. These items measure understanding of phonemes as they relate to writing. This item type is especially designed to test students at lower levels of proficiency, who often read and write based on phonics rules.

Item Component	Description of Component
Stimulus	--
Stem	closed stem question: “Which word goes with the picture?”
Answer options	3 answer options consisting of 1 word each
Graphic	graphic depicting correct answer option

A.2: Item Specification (continued)**WRITING CONVENTIONS: Mechanics and Structure Type 1, Grades 2–12**

The test taker must choose which of the options is a correct sentence. These items measure understanding of sentence-level grammar and other writing conventions in English.

Item Component	Description of Component
Stimulus	--
Stem	closed stem question: “Which sentence is correct?”
Answer options	3 answer options (Grades K–1) or 4 answer options (Grades 2–12) consisting of 1 sentence (or sentence fragment) each, maximum 20 words
Graphic	--

WRITING CONVENTIONS: Mechanics and Structure Type 2, Grades 2–12

The test taker must determine if the underlined portion of the stimulus sentence is correct as is or can be corrected by replacing it with one of three answer options. These items measure understanding of sentence-level grammar and other writing conventions in English.

Item Component	Description of Component
Stimulus	1–2 sentences with part of the sentence underlined; maximum 15 words (Grades 2–8) or 25 words (Grades 9–12)
Stem	closed stem question: “Which answer is correct?”
Answer options	4 answer options consisting of 3 phrases to replace underlined part of stimulus and 1 “Correct as is” option
Graphic	--

WRITING: Sentence Copying, Grades K–1

The test taker must copy a full sentence. (The sentence may be a question.) These items measure understanding of basic mechanical skills.

Item Component	Description of Component
Prompt	a complete sentence five to seven words in length

WRITING: Sentence Copying, Grades K–1

The test taker must copy a full sentence. (The sentence may be a question.) These items measure understanding of basic mechanical skills.

Item Component	Description of Component
Graphic	--

A.2: Item Specification (continued)**WRITING: Letter Writing, Grades K–1**

The test taker must write a spoken letter. These items measure understanding of basic mechanical skills.

Item Component	Description of Component
Prompt	a word that starts with the letter to be written
Graphic	--

WRITING: Word Writing, Grades K–1

The test taker must write a word that is dictated. Word dictations are accompanied by a graphic representation of the dictated word. These items measure understanding of basic mechanical skills.

Item Component	Description of Component
Prompt	a word that can be clearly represented in graphic form
Graphic	graphic depicting the dictated word

WRITING: Sentence Writing, Grades K–1

The test taker must write a sentence that is dictated. (The sentence may be a question.) These items measure understanding of basic mechanical skills.

Item Component	Description of Component
Prompt	a complete sentence four to six words in length
Graphic	--

A.2: Item Specification (continued)**PRE-WRITING: Grades 2–12**

The test taker briefly responds to three short questions related to a subsequent Writing item in order to generate ideas for the Writing item.

Item Component	Description of Component
Prompt	3 questions that relate to the topic of the subsequent Writing item
Graphic	--

WRITING: Extended Response: Description and/or Narrative, Grades K–2

The test taker describes or tells a story about a picture. This item measures the ability to use vocabulary, grammar, features of extended discourse, and other writing conventions correctly and appropriately in order to describe or to tell a story.

Item Component	Description of Component
Prompt	graphic and directions: “Look at the picture. Write about what you see in the picture. Tell a story about this picture.”
Graphic	1 large graphic (drawing for K–1, photograph for 2–6)

WRITING: Extended Response: Exposition, Grades 2–12

The test taker compares and/or discusses activities of a particular type. This item measures the ability to use vocabulary, grammar, features of extended discourse, and other writing conventions correctly and appropriately to write an expository essay.

Item Component	Description of Component
Prompt	graphics and directions: “Write about Write an essay about”
Graphic	2 graphics (photographs) depicting related activities or situations

A.2: Item Specification (continued)**SPEAKING: Repeat/Read Aloud, all grades**

The test taker hears and reads a word, phrase, or sentence and then repeats it. These items measure pronunciation and fluency.

Item Component	Description of Component
Prompt	a word, phrase, or sentence, one to six words in length
Graphic	--

SPEAKING: Sentence Completion, all grades

The test taker hears and reads the beginning of a sentence, and must complete the sentence based on a scenario depicted in an accompanying graphic. These items measure the ability to use vocabulary and grammar correctly and appropriately.

Item Component	Description of Component
Prompt	graphic and beginning of a sentence, maximum 10 words
Graphic	graphic depicting a recognizable scene or activity

SPEAKING: Story-Telling, all grades

The test taker must present a brief narrative based on three graphics that show a sequence of events. This item measures pronunciation, intonation, rhythm, fluency, and the ability to use vocabulary, grammar, and discourse organizational features correctly and appropriately.

Item Component	Description of Component
Prompt	graphics
Graphic	3 sequenced graphics labeled First (1), Next (2), Last (3)

A.2: Item Specification (continued)**SPEAKING: Social Interaction, all grades**

The test taker responds to a question or statement as though in an actual conversation. These items measure appropriateness of response, as well as pronunciation, intonation, rhythm, fluency, and the ability to use vocabulary and grammar correctly.

Item Component	Description of Component
Prompt	a question or statement, maximum 10 words, requiring a rejoinder
Graphic	--

APPENDIX B: ITEM LEVEL STATISTICS BY GRADE SPAN

B.1a: Kindergarten–Grade 1 Listening and Speaking

	Item Num	Item Forma	Max Point	N-count	% at 0	% at 1	% at 2	% at 3	% at 4	% Omit	P-value	Point Biserial
LI	1	MC	1	53,399	0	93	3	3	0	1	0.93	0.31
	2	MC	1	53,399	0	5	91	3	0	1	0.91	0.34
	3	MC	1	53,399	0	5	4	89	0	1	0.89	0.48
	4	MC	1	53,399	0	2	95	2	0	1	0.95	0.42
	5	MC	1	53,399	0	92	4	4	0	1	0.92	0.49
	6	MC	1	53,399	0	1	93	5	0	0	0.93	0.28
	7	MC	1	53,399	0	4	1	95	0	0	0.95	0.32
	8	MC	1	53,399	0	2	4	93	0	0	0.93	0.47
	9	MC	1	53,399	0	90	5	4	0	0	0.90	0.42
	10	MC	1	53,399	0	6	10	83	0	0	0.83	0.41
	11	MC	1	53,399	0	87	9	3	0	0	0.87	0.33
	12	MC	1	53,399	0	6	16	77	0	1	0.77	0.38
	13	MC	1	53,399	0	82	4	14	0	0	0.82	0.38
	14	MC	1	53,399	0	14	83	2	0	1	0.83	0.48
	15	MC	1	53,399	0	10	7	82	0	1	0.82	0.18
	16	MC	1	53,399	0	3	2	94	0	1	0.94	0.44
	17	MC	1	53,399	0	85	2	13	0	0	0.85	0.33
	18	MC	1	53,399	0	79	15	5	0	1	0.79	0.42
	19	MC	1	53,399	0	18	4	77	0	1	0.77	0.42
	20	MC	1	53,399	0	84	5	11	0	1	0.84	0.43
	21	MC	1	53,399	0	5	66	28	0	1	0.66	0.29
	22	MC	1	53,399	0	9	19	71	0	1	0.71	0.50
	23	MC	1	53,399	0	12	73	15	0	1	0.73	0.37
	24	MC	1	53,399	0	14	12	73	0	1	0.73	0.22
SP	1	SS	2	53,399	6	35	59	0	0	0	0.76	0.66
	2	SS	2	53,399	6	33	61	0	0	0	0.77	0.66
	3	SS	2	53,399	6	24	70	0	0	0	0.82	0.70
	4	SS	2	53,399	6	28	66	0	0	0	0.80	0.70
	5	SS	2	53,399	7	40	53	0	0	0	0.73	0.66
	6	SE	4	53,399	3	8	25	35	29	0	0.69	0.78
	7	SS	2	53,399	3	32	64	0	0	0	0.80	0.70
	8	SS	2	53,399	5	27	68	0	0	0	0.82	0.71
	9	SS	2	53,399	3	23	74	0	0	0	0.85	0.69
	10	SS	2	53,399	6	38	56	0	0	0	0.75	0.71
	11	SS	2	53,399	5	29	66	0	0	0	0.80	0.70
	12	SS	2	53,399	6	20	74	0	0	0	0.84	0.71
	13	SS	2	53,399	8	17	75	0	0	0	0.84	0.72
	14	SS	2	53,399	7	22	71	0	0	0	0.82	0.70
	15	SS	2	53,399	10	24	66	0	0	0	0.78	0.72
	16	SS	2	53,399	8	14	79	0	0	0	0.85	0.59

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=218,077) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

B.1b: Kindergarten–Grade 1 Reading and Writing

	Item Num	Item Forma	Max Point	N-count	% at 0	% at 1	% at 2	% at 3	% at 4	% Omit	P-value	Point Biserial
RD	1	MC	1	53,795	0	8	4	88	0	1	0.88	0.51
	2	MC	1	53,795	0	83	5	12	0	0	0.83	0.35
	3	MC	1	53,795	0	5	10	85	0	0	0.85	0.58
	4	MC	1	53,795	0	83	9	8	0	1	0.83	0.47
	5	MC	1	53,795	0	90	5	4	0	1	0.90	0.45
	6	MC	1	53,795	0	9	85	5	0	1	0.85	0.54
	7	MC	1	53,795	0	21	10	68	0	1	0.68	0.61
	8	MC	1	53,795	0	87	7	5	0	1	0.87	0.44
	9	MC	1	53,795	0	73	14	12	0	1	0.73	0.40
	10	MC	1	53,795	0	16	70	12	0	1	0.70	0.59
	11	MC	1	53,795	0	25	10	64	0	1	0.64	0.56
	12	MC	1	53,795	0	14	66	18	0	1	0.66	0.47
	13	MC	1	53,795	0	11	21	67	0	1	0.67	0.49
	14	MC	1	53,795	0	13	18	67	0	2	0.67	0.51
	15	MC	1	53,795	0	10	79	8	0	2	0.79	0.41
WR	1	MC	1	53,795	0	91	4	5	0	0	0.91	0.41
	2	MC	1	53,795	0	4	2	94	0	0	0.94	0.44
	3	MC	1	53,795	0	16	2	82	0	0	0.82	0.54
	4	MC	1	53,795	0	83	11	6	0	0	0.83	0.36
	5	MC	1	53,795	0	6	85	8	0	0	0.85	0.36
	6	MC	1	53,795	0	17	74	8	0	1	0.74	0.54
	7	CR	2	53,795	2	33	65	0	0	0	0.82	0.52
	8	CR	2	53,795	2	31	67	0	0	0	0.83	0.41
	9	CR	2	53,795	7	3	90	0	0	0	0.91	0.39
	10	CR	2	53,795	10	13	78	0	0	0	0.84	0.54
	11	CR	2	53,795	23	23	53	0	0	0	0.65	0.77
	12	CR	2	53,795	15	16	69	0	0	0	0.77	0.75
	13	CR	2	53,795	18	20	61	0	0	0	0.72	0.83
	14	CR	2	53,795	27	27	46	0	0	0	0.60	0.84
	15	CR	2	53,795	28	40	32	0	0	0	0.52	0.79

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=218,077) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

B.2a: Grade 2–Grade 4 Listening and Speaking

	Item Number	Item Format	Max Points	N-count	% at 0	% at 1	% at 2	% at 3	% at 4	% Omit	P-value	Point Biserial
LI	1	MC	1	65,353	0	95	3	2	0	0	0.95	0.46
	2	MC	1	65,353	0	7	86	7	0	0	0.86	0.35
	3	MC	1	65,353	0	6	3	91	0	0	0.91	0.48
	4	MC	1	65,353	0	3	4	93	0	0	0.93	0.44
	5	MC	1	65,353	0	96	2	2	0	0	0.96	0.46
	6	MC	1	65,353	0	90	4	6	0	0	0.90	0.31
	7	MC	1	65,353	0	8	12	80	0	0	0.80	0.39
	8	MC	1	65,353	0	8	61	31	0	0	0.61	0.25
	9	MC	1	65,353	0	71	24	5	0	0	0.71	0.36
	10	MC	1	65,353	0	5	91	4	0	0	0.91	0.48
	11	MC	1	65,353	0	3	2	4	91	0	0.91	0.48
	12	MC	1	65,353	0	4	10	15	70	0	0.70	0.41
	13	MC	1	65,353	0	9	19	64	8	0	0.64	0.34
	14	MC	1	65,353	0	77	10	4	8	0	0.77	0.44
	15	MC	1	65,353	0	4	3	87	6	0	0.87	0.48
	16	MC	1	65,353	0	75	3	4	17	0	0.75	0.41
	17	MC	1	65,353	0	10	5	79	6	0	0.79	0.41
	18	MC	1	65,353	0	6	1	12	80	0	0.80	0.43
	19	MC	1	65,353	0	29	7	53	11	1	0.53	0.27
	20	MC	1	65,353	0	4	90	3	2	0	0.90	0.39
	21	MC	1	65,353	0	1	3	89	6	0	0.89	0.46
	22	MC	1	65,353	0	6	4	86	3	1	0.86	0.40
	23	MC	1	65,353	0	5	87	4	3	0	0.87	0.41
	24	MC	1	65,353	0	85	10	2	2	1	0.85	0.42
SP	1	SS	2	65,353	3	22	75	0	0	0	0.86	0.65
	2	SS	2	65,353	7	28	65	0	0	0	0.79	0.66
	3	SS	2	65,353	3	30	67	0	0	0	0.82	0.61
	4	SS	2	65,353	6	21	73	0	0	0	0.84	0.69
	5	SS	2	65,353	7	29	64	0	0	0	0.78	0.68
	6	SE	4	65,353	2	6	19	37	36	0	0.75	0.78
	7	SS	2	65,353	3	18	79	0	0	0	0.88	0.71
	8	SS	2	65,353	3	20	77	0	0	0	0.87	0.69
	9	SS	2	65,353	4	21	76	0	0	0	0.86	0.73
	10	SS	2	65,353	4	24	72	0	0	0	0.84	0.72
	11	SS	2	65,353	3	19	78	0	0	0	0.87	0.70
	12	SS	2	65,353	5	12	83	0	0	0	0.89	0.72
	13	SS	2	65,353	6	13	81	0	0	0	0.87	0.71
	14	SS	2	65,353	3	18	78	0	0	0	0.88	0.64
	15	SS	2	65,353	6	15	79	0	0	0	0.87	0.68
	16	SS	2	65,353	6	20	74	0	0	0	0.84	0.70

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=218,077) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

B.2b: Grade 2–Grade 4 Reading and Writing

	Item Number	Item Format	Max Points	N-count	% at 0	% at 1	% at 2	% at 3	% at 4	% Omit	P-value	Point Biserial
RD	1	MC	1	62,408	0	99	1	1	0	0	0.99	0.23
	2	MC	1	62,408	0	5	5	90	0	0	0.90	0.44
	3	MC	1	62,408	0	93	3	4	0	0	0.93	0.35
	4	MC	1	62,408	0	3	93	3	0	0	0.93	0.35
	5	MC	1	62,408	0	96	3	2	0	0	0.96	0.41
	6	MC	1	62,408	0	94	3	2	0	0	0.94	0.40
	7	MC	1	62,408	0	5	76	19	0	0	0.76	0.50
	8	MC	1	62,408	0	80	6	8	5	0	0.80	0.36
	9	MC	1	62,408	0	39	44	10	6	0	0.44	0.36
	10	MC	1	62,408	0	4	62	9	25	1	0.62	0.44
	11	MC	1	62,408	0	10	7	10	73	0	0.73	0.46
	12	MC	1	62,408	0	7	81	3	9	0	0.81	0.57
	13	MC	1	62,408	0	13	10	5	71	1	0.71	0.55
	14	MC	1	62,408	0	69	11	10	10	1	0.69	0.38
	15	MC	1	62,408	0	7	72	12	9	1	0.72	0.52
	16	MC	1	62,408	0	18	21	9	51	0	0.51	0.45
	17	MC	1	62,408	0	10	66	10	13	1	0.66	0.57
	18	MC	1	62,408	0	79	5	11	5	1	0.79	0.25
	19	MC	1	62,408	0	17	29	7	46	1	0.46	0.33
	20	MC	1	62,408	0	4	79	5	11	1	0.79	0.44
	21	MC	1	62,408	0	16	9	69	5	1	0.69	0.63
	22	MC	1	62,408	0	56	9	22	12	1	0.56	0.39
	23	MC	1	62,408	0	6	7	20	67	1	0.67	0.47
	24	MC	1	62,408	0	11	6	73	10	1	0.73	0.53
WR	1	MC	1	62,408	0	93	3	4	0	0	0.93	0.37
	2	MC	1	62,408	0	2	97	1	0	0	0.97	0.31
	3	MC	1	62,408	0	10	13	77	0	0	0.77	0.44
	4	MC	1	62,408	0	2	30	68	0	0	0.68	0.34
	5	MC	1	62,408	0	9	81	6	4	0	0.81	0.45
	6	MC	1	62,408	0	5	5	79	10	0	0.79	0.53
	7	MC	1	62,408	0	6	6	19	68	0	0.68	0.48
	8	MC	1	62,408	0	11	20	15	54	0	0.54	0.36
	9	MC	1	62,408	0	78	7	4	10	0	0.78	0.49
	10	MC	1	62,408	0	18	10	6	66	0	0.66	0.51
	11	MC	1	62,408	0	68	13	6	12	0	0.68	0.45
	12	MC	1	62,408	0	48	9	14	28	1	0.48	0.37
	13	CR	2	62,408	2	15	83	0	0	0	0.90	0.51
	14	CR	2	62,408	5	18	77	0	0	0	0.86	0.56
	15	CR	2	62,408	12	34	55	0	0	0	0.72	0.61
	16	ER	4	62,408	7	23	37	26	8	0	0.51	0.71

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=218,077) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

B.3a: Grade 5–Grade 6 Listening and Speaking

	Item Number	Item Format	Max Points	N-count	% at 0	% at 1	% at 2	% at 3	% at 4	% Omit	P-value	Point Biserial
LI	1	MC	1	30,306	0	96	2	2	0	0	0.96	0.39
	2	MC	1	30,306	0	2	6	91	0	0	0.91	0.51
	3	MC	1	30,306	0	19	4	76	0	0	0.76	0.34
	4	MC	1	30,306	0	4	89	3	4	0	0.89	0.61
	5	MC	1	30,306	0	2	4	5	89	0	0.89	0.51
	6	MC	1	30,306	0	2	6	88	4	0	0.88	0.45
	7	MC	1	30,306	0	9	8	4	79	0	0.79	0.44
	8	MC	1	30,306	0	84	5	5	6	0	0.84	0.48
	9	MC	1	30,306	0	79	9	8	3	0	0.79	0.39
	10	MC	1	30,306	0	10	13	6	71	0	0.71	0.39
	11	MC	1	30,306	0	7	5	12	75	0	0.75	0.35
	12	MC	1	30,306	0	5	81	6	7	0	0.81	0.44
	13	MC	1	30,306	0	15	24	53	8	0	0.53	0.31
	14	MC	1	30,306	0	17	62	12	10	0	0.62	0.34
	15	MC	1	30,306	0	13	19	45	22	0	0.45	0.31
	16	MC	1	30,306	0	82	7	6	5	0	0.82	0.50
	17	MC	1	30,306	0	3	4	78	13	1	0.78	0.39
	18	MC	1	30,306	0	18	18	6	58	0	0.58	0.27
	19	MC	1	30,306	0	4	5	86	5	0	0.86	0.58
	20	MC	1	30,306	0	2	16	78	4	1	0.78	0.27
	21	MC	1	30,306	0	26	53	10	11	0	0.53	0.28
	22	MC	1	30,306	0	82	7	8	2	0	0.82	0.56
	23	MC	1	30,306	0	14	46	16	23	0	0.46	0.22
	24	MC	1	30,306	0	12	3	7	77	1	0.77	0.38
	25	MC	1	30,306	0	75	19	2	2	1	0.75	0.44
SP	1	SS	2	30,306	7	25	68	0	0	0	0.80	0.73
	2	SS	2	30,306	7	23	71	0	0	0	0.82	0.73
	3	SS	2	30,306	8	18	74	0	0	0	0.83	0.76
	4	SS	2	30,306	8	34	59	0	0	0	0.76	0.69
	5	SS	2	30,306	8	24	68	0	0	0	0.80	0.72
	6	SE	4	30,306	2	5	13	31	48	0	0.79	0.82
	7	SS	2	30,306	3	22	75	0	0	0	0.86	0.73
	8	SS	2	30,306	4	22	74	0	0	0	0.85	0.72
	9	SS	2	30,306	3	16	82	0	0	0	0.89	0.74
	10	SS	2	30,306	5	26	68	0	0	0	0.81	0.70
	11	SS	2	30,306	6	19	75	0	0	0	0.85	0.77
	12	SS	2	30,306	6	9	85	0	0	0	0.90	0.72
	13	SS	2	30,306	13	17	70	0	0	0	0.78	0.77
	14	SS	2	30,306	8	22	71	0	0	0	0.81	0.68
	15	SS	2	30,306	8	21	71	0	0	0	0.82	0.75
	16	SS	2	30,306	9	19	71	0	0	0	0.81	0.77

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=218,077) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score = 0 or > 900).

B.3b: Grade 5–Grade 6 Reading and Writing

	Item Number	Item Format	Max Points	N-count	% at 0	% at 1	% at 2	% at 3	% at 4	% Omit	P-value	Point Biserial
RD	1	MC	1	28,821	0	4	2	94	0	0	0.94	0.34
	2	MC	1	28,821	0	1	96	3	0	0	0.96	0.34
	3	MC	1	28,821	0	89	5	6	0	0	0.89	0.32
	4	MC	1	28,821	0	2	2	94	2	0	0.94	0.37
	5	MC	1	28,821	0	64	23	8	5	0	0.64	0.44
	6	MC	1	28,821	0	2	13	4	80	0	0.80	0.37
	7	MC	1	28,821	0	3	84	7	6	0	0.84	0.52
	8	MC	1	28,821	0	6	6	6	82	0	0.82	0.56
	9	MC	1	28,821	0	68	22	4	6	0	0.68	0.34
	10	MC	1	28,821	0	8	11	67	13	0	0.67	0.58
	11	MC	1	28,821	0	10	12	7	71	0	0.71	0.48
	12	MC	1	28,821	0	26	61	6	7	0	0.61	0.51
	13	MC	1	28,821	0	9	7	7	77	0	0.77	0.56
	14	MC	1	28,821	0	30	56	7	6	0	0.56	0.48
	15	MC	1	28,821	0	10	6	77	7	0	0.77	0.54
	16	MC	1	28,821	0	74	8	9	9	0	0.74	0.56
	17	MC	1	28,821	0	8	25	9	57	0	0.57	0.47
	18	MC	1	28,821	0	11	10	35	44	0	0.44	0.40
	19	MC	1	28,821	0	22	60	8	9	1	0.60	0.48
	20	MC	1	28,821	0	7	8	10	75	0	0.75	0.50
	21	MC	1	28,821	0	11	74	7	7	0	0.74	0.48
	22	MC	1	28,821	0	52	20	18	10	0	0.52	0.41
	23	MC	1	28,821	0	40	10	21	28	1	0.40	0.35
	24	MC	1	28,821	0	19	12	55	14	1	0.55	0.37
	25	MC	1	28,821	0	16	21	8	54	1	0.54	0.34
	26	MC	1	28,821	0	55	17	12	14	0	0.55	0.45
	27	MC	1	28,821	0	20	47	22	11	1	0.47	0.38
WR	1	MC	1	28,821	0	4	93	3	0	0	0.93	0.39
	2	MC	1	28,821	0	17	79	4	0	0	0.79	0.30
	3	MC	1	28,821	0	6	5	8	81	0	0.81	0.40
	4	MC	1	28,821	0	3	6	87	3	0	0.87	0.48
	5	MC	1	28,821	0	83	4	11	3	0	0.83	0.51
	6	MC	1	28,821	0	15	11	57	18	0	0.57	0.43
	7	MC	1	28,821	0	15	11	65	8	1	0.65	0.31
	8	MC	1	28,821	0	5	6	6	82	0	0.82	0.51
	9	MC	1	28,821	0	11	5	76	7	0	0.76	0.51
	10	MC	1	28,821	0	68	12	2	17	0	0.68	0.41
	11	MC	1	28,821	0	23	9	18	50	0	0.50	0.36
	12	MC	1	28,821	0	4	82	5	9	0	0.82	0.42
	13	MC	1	28,821	0	6	26	60	9	0	0.60	0.39
	14	MC	1	28,821	0	4	4	81	10	0	0.81	0.42
	15	MC	1	28,821	0	5	52	13	30	0	0.52	0.29
	16	CR	2	28,821	8	35	57	0	0	0	0.75	0.42
	17	CR	2	28,821	10	37	53	0	0	0	0.72	0.59
	18	CR	2	28,821	17	37	46	0	0	0	0.65	0.51
	19	ER	4	28,821	8	20	39	26	8	0	0.51	0.68

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=218,077) after cleaning (with valid corresponding subtest raw score).

B.4a: Grade 7–Grade 8 Listening and Speaking

	Item Number	Item Format	Max Points	N-count	% at 0	% at 1	% at 2	% at 3	% at 4	% Omit	P-value	Point Biserial
LI	1	MC	1	26,227	0	12	74	13	0	1	0.74	0.24
	2	MC	1	26,227	0	89	7	4	0	0	0.89	0.39
	3	MC	1	26,227	0	44	24	32	0	0	0.44	0.34
	4	MC	1	26,227	0	9	10	6	74	0	0.74	0.51
	5	MC	1	26,227	0	7	23	65	4	0	0.65	0.49
	6	MC	1	26,227	0	5	89	3	3	0	0.89	0.54
	7	MC	1	26,227	0	7	4	83	6	0	0.83	0.58
	8	MC	1	26,227	0	85	3	9	3	0	0.85	0.55
	9	MC	1	26,227	0	5	79	12	5	0	0.79	0.44
	10	MC	1	26,227	0	7	17	54	22	0	0.54	0.36
	11	MC	1	26,227	0	12	21	22	45	0	0.45	0.34
	12	MC	1	26,227	0	11	17	58	14	1	0.58	0.33
	13	MC	1	26,227	0	6	83	4	7	0	0.83	0.21
	14	MC	1	26,227	0	12	15	61	10	1	0.61	0.41
	15	MC	1	26,227	0	41	5	39	14	0	0.41	0.38
	16	MC	1	26,227	0	11	5	3	81	0	0.81	0.56
	17	MC	1	26,227	0	71	5	7	16	0	0.71	0.49
	18	MC	1	26,227	0	5	86	4	4	0	0.86	0.50
	19	MC	1	26,227	0	82	5	7	5	0	0.82	0.49
	20	MC	1	26,227	0	7	9	69	16	0	0.69	0.51
	21	MC	1	26,227	0	82	8	5	5	0	0.82	0.47
	22	MC	1	26,227	0	6	9	77	8	0	0.77	0.48
	23	MC	1	26,227	0	11	11	67	11	0	0.67	0.49
	24	MC	1	26,227	0	3	12	12	74	0	0.74	0.41
	25	MC	1	26,227	0	3	17	44	35	0	0.44	0.18
SP	1	SS	2	26,227	5	24	71	0	0	0	0.83	0.71
	2	SS	2	26,227	6	27	68	0	0	0	0.81	0.74
	3	SS	2	26,227	5	22	74	0	0	0	0.85	0.73
	4	SS	2	26,227	5	23	72	0	0	0	0.83	0.75
	5	SS	2	26,227	5	25	70	0	0	0	0.82	0.65
	6	SE	4	26,227	3	7	17	31	42	0	0.76	0.83
	7	SS	2	26,227	5	24	71	0	0	0	0.83	0.76
	8	SS	2	26,227	4	21	75	0	0	0	0.86	0.76
	9	SS	2	26,227	7	28	65	0	0	0	0.79	0.75
	10	SS	2	26,227	9	24	67	0	0	0	0.79	0.79
	11	SS	2	26,227	7	31	62	0	0	0	0.77	0.75
	12	SS	2	26,227	9	20	71	0	0	0	0.81	0.80
	13	SS	2	26,227	10	20	70	0	0	0	0.80	0.80
	14	SS	2	26,227	9	22	69	0	0	0	0.80	0.76
	15	SS	2	26,227	12	18	70	0	0	0	0.79	0.78
	16	SS	2	26,227	10	13	77	0	0	0	0.84	0.80

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=218,077) after cleaning (with valid corresponding subtest raw score).

B.4b: Grade 7–Grade 8 Reading and Writing

	Item Number	Item Format	Max Points	N-count	% at 0	% at 1	% at 2	% at 3	% at 4	% Omit	P-value	Point Biserial
RD	1	MC	1	25,039	0	3	95	2	0	0	0.95	0.37
	2	MC	1	25,039	0	3	11	86	0	0	0.86	0.33
	3	MC	1	25,039	0	3	4	94	0	0	0.94	0.36
	4	MC	1	25,039	0	78	7	12	3	0	0.78	0.42
	5	MC	1	25,039	0	18	18	48	16	0	0.48	0.39
	6	MC	1	25,039	0	25	6	10	59	0	0.59	0.40
	7	MC	1	25,039	0	68	21	6	5	0	0.68	0.56
	8	MC	1	25,039	0	9	77	10	4	0	0.77	0.36
	9	MC	1	25,039	0	13	68	6	13	0	0.68	0.57
	10	MC	1	25,039	0	13	8	68	10	0	0.68	0.50
	11	MC	1	25,039	0	17	66	9	7	0	0.66	0.48
	12	MC	1	25,039	0	3	80	8	8	0	0.80	0.46
	13	MC	1	25,039	0	8	7	9	75	0	0.75	0.56
	14	MC	1	25,039	0	26	51	6	16	1	0.51	0.23
	15	MC	1	25,039	0	63	11	17	9	0	0.63	0.47
	16	MC	1	25,039	0	14	7	74	6	0	0.74	0.41
	17	MC	1	25,039	0	19	11	60	9	0	0.60	0.34
	18	MC	1	25,039	0	73	17	5	5	0	0.73	0.55
	19	MC	1	25,039	0	9	68	11	11	0	0.68	0.48
	20	MC	1	25,039	0	21	18	12	49	0	0.49	0.50
	21	MC	1	25,039	0	75	9	8	8	0	0.75	0.54
	22	MC	1	25,039	0	57	19	11	11	0	0.57	0.50
	23	NS	1	25,039	0	10	39	8	42	0	0.00	
	24	MC	1	25,039	0	59	15	14	11	1	0.59	0.44
	25	MC	1	25,039	0	16	67	11	6	1	0.67	0.42
	26	MC	1	25,039	0	24	16	49	11	1	0.49	0.35
	27	MC	1	25,039	0	21	7	16	56	1	0.56	0.52
WR	1	MC	1	25,039	0	13	21	66	0	0	0.66	0.42
	2	MC	1	25,039	0	87	9	4	0	0	0.87	0.35
	3	MC	1	25,039	0	5	3	3	89	0	0.89	0.39
	4	MC	1	25,039	0	4	8	20	68	0	0.68	0.50
	5	MC	1	25,039	0	78	7	12	3	0	0.78	0.53
	6	MC	1	25,039	0	69	21	6	4	0	0.69	0.45
	7	MC	1	25,039	0	30	55	9	5	0	0.55	0.31
	8	MC	1	25,039	0	89	4	5	3	0	0.89	0.42
	9	MC	1	25,039	0	67	5	14	15	0	0.67	0.45
	10	MC	1	25,039	0	5	4	82	10	0	0.82	0.48
	11	MC	1	25,039	0	74	10	10	5	0	0.74	0.48
	12	MC	1	25,039	0	2	16	20	62	0	0.62	0.45
	13	MC	1	25,039	0	29	64	4	3	0	0.64	0.46
	14	MC	1	25,039	0	49	21	9	21	0	0.49	0.25
	15	MC	1	25,039	0	51	7	14	28	0	0.51	0.46
	16	CR	2	25,039	5	18	77	0	0	0	0.86	0.52
	17	CR	2	25,039	4	26	70	0	0	0	0.83	0.52
	18	CR	2	25,039	12	35	53	0	0	0	0.70	0.58
	19	ER	4	25,039	5	15	34	32	14	0	0.59	0.71

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=218,077) after cleaning (with valid corresponding subtest raw score).

B.5a: Grade 9–Grade 12 Listening and Speaking

	Item Number	Item Format	Max Points	N-count	% at 0	% at 1	% at 2	% at 3	% at 4	% Omit	P-value	Point Biserial
LI	1	MC	1	42,792	0	11	84	5	0	1	0.84	0.31
	2	MC	1	42,792	0	81	10	8	0	1	0.81	0.47
	3	MC	1	42,792	0	4	3	92	0	0	0.92	0.41
	4	MC	1	42,792	0	46	26	13	16	0	0.46	0.40
	5	MC	1	42,792	0	16	11	68	5	1	0.68	0.53
	6	MC	1	42,792	0	4	84	4	8	0	0.84	0.31
	7	MC	1	42,792	0	84	7	4	4	0	0.84	0.40
	8	MC	1	42,792	0	13	11	7	68	0	0.68	0.57
	9	MC	1	42,792	0	78	10	7	6	0	0.78	0.50
	10	MC	1	42,792	0	5	79	4	12	0	0.79	0.43
	11	MC	1	42,792	0	10	2	72	16	0	0.72	0.32
	12	MC	1	42,792	0	63	17	13	6	1	0.63	0.43
	13	MC	1	42,792	0	14	21	12	53	1	0.53	0.33
	14	MC	1	42,792	0	6	11	79	3	0	0.79	0.39
	15	MC	1	42,792	0	70	17	4	9	0	0.70	0.46
	16	MC	1	42,792	0	3	4	91	2	0	0.91	0.47
	17	MC	1	42,792	0	87	4	7	2	0	0.87	0.43
	18	MC	1	42,792	0	6	6	8	79	1	0.79	0.42
	19	MC	1	42,792	0	13	63	15	8	0	0.63	0.39
	20	MC	1	42,792	0	5	6	86	3	0	0.86	0.46
	21	MC	1	42,792	0	12	11	14	62	0	0.62	0.42
	22	MC	1	42,792	0	4	87	4	4	0	0.87	0.22
	23	MC	1	42,792	0	3	4	23	70	0	0.70	0.36
	24	MC	1	42,792	0	5	3	16	75	0	0.75	0.39
	25	MC	1	42,792	0	20	60	16	4	0	0.60	0.36
SP	1	SS	2	42,792	7	30	63	0	0	0	0.78	0.73
	2	SS	2	42,792	4	26	70	0	0	0	0.83	0.71
	3	SS	2	42,792	6	20	74	0	0	0	0.84	0.74
	4	SS	2	42,792	12	34	54	0	0	0	0.71	0.71
	5	SS	2	42,792	5	23	72	0	0	0	0.83	0.73
	6	SE	4	42,792	2	8	20	30	39	0	0.74	0.83
	7	SS	2	42,792	5	28	66	0	0	0	0.80	0.74
	8	SS	2	42,792	5	30	66	0	0	0	0.80	0.75
	9	SS	2	42,792	3	27	70	0	0	0	0.83	0.73
	10	SS	2	42,792	4	26	70	0	0	0	0.83	0.70
	11	SS	2	42,792	6	31	63	0	0	0	0.78	0.74
	12	SS	2	42,792	9	16	74	0	0	0	0.83	0.74
	13	SS	2	42,792	8	24	68	0	0	0	0.80	0.76
	14	SS	2	42,792	7	22	72	0	0	0	0.82	0.73
	15	SS	2	42,792	5	24	71	0	0	0	0.83	0.70
	16	SS	2	42,792	10	21	69	0	0	0	0.79	0.70

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=218,077) after cleaning (with valid corresponding subtest raw score).

B.5b: Grade 9–Grade 12 Reading and Writing

	Item Number	Item Format	Max Points	N-count	% at 0	% at 1	% at 2	% at 3	% at 4	% Omit	P-value	Point Biserial
RD	1	MC	1	40,934	0	5	12	83	0	0	0.83	0.37
	2	MC	1	40,934	0	85	13	2	0	0	0.85	0.30
	3	MC	1	40,934	0	71	18	10	0	0	0.71	0.44
	4	MC	1	40,934	0	9	10	12	70	0	0.70	0.47
	5	MC	1	40,934	0	15	10	8	67	0	0.67	0.52
	6	MC	1	40,934	0	65	14	11	10	1	0.65	0.49
	7	MC	1	40,934	0	9	19	67	4	0	0.67	0.45
	8	MC	1	40,934	0	11	34	11	44	0	0.44	0.26
	9	MC	1	40,934	0	14	52	9	25	0	0.52	0.41
	10	MC	1	40,934	0	68	8	11	12	1	0.68	0.43
	11	MC	1	40,934	0	17	6	30	46	1	0.46	0.38
	12	MC	1	40,934	0	66	14	11	8	1	0.66	0.56
	13	MC	1	40,934	0	19	9	23	49	1	0.49	0.41
	14	MC	1	40,934	0	4	85	3	8	1	0.85	0.34
	15	MC	1	40,934	0	4	4	16	76	0	0.76	0.27
	16	MC	1	40,934	0	5	6	14	75	0	0.75	0.40
	17	MC	1	40,934	0	5	3	88	4	0	0.88	0.40
	18	MC	1	40,934	0	26	7	15	51	1	0.51	0.31
	19	MC	1	40,934	0	19	53	14	13	1	0.53	0.27
	20	MC	1	40,934	0	64	13	15	7	1	0.64	0.57
	21	MC	1	40,934	0	65	18	11	5	1	0.65	0.42
	22	MC	1	40,934	0	21	55	8	15	1	0.55	0.40
	23	MC	1	40,934	0	18	10	51	20	1	0.51	0.18
	24	MC	1	40,934	0	17	57	14	12	1	0.57	0.24
	25	MC	1	40,934	0	22	12	52	13	1	0.52	0.53
	26	MC	1	40,934	0	30	34	19	16	1	0.30	0.32
	27	MC	1	40,934	0	21	13	49	16	1	0.49	0.38
WR	1	MC	1	40,934	0	16	6	77	0	0	0.77	0.43
	2	MC	1	40,934	0	5	88	6	0	0	0.88	0.37
	3	MC	1	40,934	0	92	1	3	3	0	0.92	0.32
	4	MC	1	40,934	0	3	88	5	4	0	0.88	0.43
	5	MC	1	40,934	0	6	12	75	7	0	0.75	0.44
	6	MC	1	40,934	0	29	64	3	4	0	0.64	0.41
	7	MC	1	40,934	0	63	10	12	13	1	0.63	0.47
	8	MC	1	40,934	0	1	7	86	5	0	0.86	0.33
	9	MC	1	40,934	0	4	15	63	18	0	0.63	0.39
	10	MC	1	40,934	0	9	8	7	76	0	0.76	0.44
	11	MC	1	40,934	0	7	60	21	12	0	0.60	0.42
	12	MC	1	40,934	0	17	8	64	11	0	0.64	0.32
	13	MC	1	40,934	0	5	9	71	15	0	0.71	0.54
	14	MC	1	40,934	0	21	10	2	66	0	0.66	0.37
	15	MC	1	40,934	0	52	12	16	19	0	0.52	0.46
	16	CR	2	40,934	6	30	64	0	0	0	0.79	0.52
	17	CR	2	40,934	8	34	59	0	0	0	0.75	0.53
	18	CR	2	40,934	16	36	48	0	0	0	0.66	0.58
	19	ER	4	40,934	4	15	30	31	21	0	0.62	0.71

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=218,077) after cleaning (with valid corresponding subtest raw score).

APPENDIX C: RAW SCORE TO SCALE SCORE CONVERSION TABLES

C.1a: Kindergarten–Grade 1 Listening and Speaking

Raw Score	Measure	SE	Scale Score	SE(SS)
0	-7.77	2.01	328	70
1	-6.36	1.02	378	36
2	-5.63	0.73	403	26
3	-5.19	0.61	418	21
4	-4.86	0.53	430	19
5	-4.60	0.49	439	17
6	-4.39	0.45	447	16
7	-4.19	0.42	453	15
8	-4.02	0.40	459	14
9	-3.87	0.39	465	13
10	-3.73	0.37	470	13
11	-3.59	0.36	474	13
12	-3.47	0.35	479	12
13	-3.35	0.34	483	12
14	-3.24	0.33	487	12
15	-3.13	0.32	490	11
16	-3.03	0.32	494	11
17	-2.93	0.31	498	11
18	-2.83	0.31	501	11
19	-2.74	0.31	504	11
20	-2.64	0.30	507	11
21	-2.55	0.30	511	10
22	-2.46	0.30	514	10
23	-2.38	0.29	517	10
24	-2.29	0.29	520	10
25	-2.21	0.29	523	10
26	-2.12	0.29	526	10
27	-2.04	0.29	529	10
28	-1.95	0.29	532	10
29	-1.87	0.29	535	10
30	-1.79	0.29	537	10
31	-1.70	0.29	540	10
32	-1.62	0.29	543	10
33	-1.53	0.29	546	10
34	-1.45	0.29	549	10
35	-1.36	0.30	552	10
36	-1.27	0.30	555	10
37	-1.18	0.30	559	11
38	-1.09	0.30	562	11
39	-1.00	0.31	565	11
40	-0.90	0.31	568	11
41	-0.80	0.32	572	11
42	-0.70	0.32	575	11

C.1a: Kindergarten–Grade 1 Listening and Speaking (continued)

Raw Score	Measure	SE	Scale Score	SE(SS)
43	-0.60	0.33	579	11
44	-0.49	0.34	583	12
45	-0.37	0.34	587	12
46	-0.25	0.35	591	12
47	-0.12	0.36	596	13
48	0.01	0.38	601	13
49	0.16	0.39	606	14
50	0.32	0.41	611	14
51	0.50	0.43	617	15
52	0.70	0.46	624	16
53	0.92	0.50	632	17
54	1.19	0.54	642	19
55	1.53	0.62	653	22
56	1.98	0.74	669	26
57	2.73	1.03	696	36
58	4.16	2.01	746	70

Note: The statistics in Tables C.1–C.5 were based on the equating sample. Invalid cases were removed after cleaning procedures.

C.1b: Kindergarten–Grade 1 Reading and Writing

Raw Score	Measure	SE	Scale Score	SE(SS)
0	-8.20	2.03	313	71
1	-6.72	1.06	365	37
2	-5.91	0.78	393	27
3	-5.40	0.66	411	23
4	-5.01	0.58	425	20
5	-4.70	0.53	435	19
6	-4.44	0.49	445	17
7	-4.22	0.46	452	16
8	-4.02	0.43	459	15
9	-3.84	0.41	466	14
10	-3.68	0.40	471	14
11	-3.53	0.38	477	13
12	-3.38	0.37	482	13
13	-3.25	0.36	486	13
14	-3.12	0.36	491	12
15	-2.99	0.35	495	12
16	-2.87	0.35	499	12
17	-2.75	0.34	504	12
18	-2.63	0.34	508	12
19	-2.52	0.34	512	12
20	-2.40	0.34	516	12
21	-2.28	0.34	520	12
22	-2.16	0.35	524	12
23	-2.04	0.35	528	12
24	-1.92	0.35	533	12
25	-1.80	0.36	537	12
26	-1.67	0.36	542	13
27	-1.53	0.37	546	13
28	-1.40	0.38	551	13
29	-1.25	0.39	556	14
30	-1.09	0.40	562	14
31	-0.92	0.42	568	15
32	-0.74	0.44	574	15
33	-0.53	0.47	581	16
34	-0.30	0.50	590	18
35	-0.02	0.55	599	19
36	0.33	0.63	611	22
37	0.80	0.75	628	26
38	1.57	1.04	655	36
39	3.02	2.02	706	71

C.2a: Grade 2–Grade 4 Listening and Speaking

Raw Score	Measure	SE	Scale Score	SE(SS)
0	-6.11	2.01	386	70
1	-4.70	1.02	436	36
2	-3.96	0.73	461	26
3	-3.52	0.61	477	21
4	-3.19	0.54	488	19
5	-2.93	0.49	497	17
6	-2.71	0.45	505	16
7	-2.52	0.43	512	15
8	-2.34	0.40	518	14
9	-2.19	0.39	523	14
10	-2.04	0.37	528	13
11	-1.91	0.36	533	13
12	-1.78	0.35	538	12
13	-1.66	0.34	542	12
14	-1.55	0.33	546	12
15	-1.44	0.33	550	11
16	-1.34	0.32	553	11
17	-1.23	0.32	557	11
18	-1.14	0.31	560	11
19	-1.04	0.31	564	11
20	-0.95	0.30	567	11
21	-0.86	0.30	570	11
22	-0.77	0.30	573	10
23	-0.68	0.30	576	10
24	-0.59	0.29	579	10
25	-0.51	0.29	582	10
26	-0.42	0.29	585	10
27	-0.34	0.29	588	10
28	-0.25	0.29	591	10
29	-0.17	0.29	594	10
30	-0.08	0.29	597	10
31	0.00	0.29	600	10
32	0.09	0.29	603	10
33	0.17	0.29	606	10
34	0.26	0.30	609	10
35	0.35	0.30	612	10
36	0.44	0.30	615	11
37	0.53	0.30	618	11
38	0.62	0.31	622	11
39	0.71	0.31	625	11
40	0.81	0.31	628	11
41	0.91	0.32	632	11
42	1.02	0.32	636	11
43	1.12	0.33	639	12
44	1.24	0.34	643	12
45	1.35	0.35	647	12

C.2a: Grade 2–Grade 4 Listening and Speaking (continued)

Raw Score	Measure	SE	Scale Score	SE(SS)
46	1.48	0.36	652	12
47	1.61	0.37	656	13
48	1.74	0.38	661	13
49	1.89	0.39	666	14
50	2.05	0.41	672	14
51	2.23	0.43	678	15
52	2.43	0.46	685	16
53	2.66	0.50	693	17
54	2.93	0.54	702	19
55	3.26	0.62	714	22
56	3.71	0.74	730	26
57	4.46	1.02	756	36
58	5.88	2.01	806	70

C.2b: Grade 2–Grade 4 Reading and Writing

Raw Score	Measure	SE	Scale Score	SE(SS)
0	-6.63	2.03	368	71
1	-5.15	1.06	420	37
2	-4.34	0.78	448	27
3	-3.83	0.66	466	23
4	-3.45	0.58	479	20
5	-3.14	0.54	490	19
6	-2.87	0.50	500	17
7	-2.64	0.47	508	16
8	-2.43	0.45	515	16
9	-2.23	0.43	522	15
10	-2.06	0.41	528	14
11	-1.89	0.40	534	14
12	-1.73	0.39	539	14
13	-1.59	0.38	545	13
14	-1.44	0.37	549	13
15	-1.31	0.37	554	13
16	-1.18	0.36	559	13
17	-1.05	0.35	563	12
18	-0.93	0.35	568	12
19	-0.80	0.35	572	12
20	-0.68	0.34	576	12
21	-0.57	0.34	580	12
22	-0.45	0.34	584	12
23	-0.34	0.34	588	12
24	-0.22	0.34	592	12
25	-0.10	0.34	596	12
26	0.01	0.34	600	12
27	0.13	0.34	604	12
28	0.25	0.35	609	12
29	0.37	0.35	613	12
30	0.49	0.35	617	12
31	0.62	0.36	622	12
32	0.75	0.36	626	13
33	0.88	0.37	631	13
34	1.02	0.38	636	13
35	1.17	0.39	641	14
36	1.32	0.40	646	14
37	1.49	0.42	652	15
38	1.67	0.43	659	15
39	1.87	0.46	665	16
40	2.09	0.49	673	17
41	2.35	0.52	682	18
42	2.65	0.57	693	20
43	3.02	0.65	706	23
44	3.52	0.78	723	27
45	4.33	1.06	752	37
46	5.82	2.04	804	71

C.3a: Grade 5–Grade 6 Listening and Speaking

Raw Score	Measure	SE	Scale Score	SE(SS)
0	-5.79	2.02	398	71
1	-4.35	1.03	448	36
2	-3.60	0.75	474	26
3	-3.14	0.62	490	22
4	-2.80	0.55	502	19
5	-2.53	0.50	511	17
6	-2.31	0.46	519	16
7	-2.11	0.43	526	15
8	-1.93	0.41	532	14
9	-1.77	0.39	538	14
10	-1.63	0.37	543	13
11	-1.50	0.36	548	13
12	-1.37	0.35	552	12
13	-1.25	0.34	556	12
14	-1.14	0.33	560	12
15	-1.03	0.32	564	11
16	-0.93	0.32	567	11
17	-0.83	0.31	571	11
18	-0.74	0.31	574	11
19	-0.65	0.30	577	11
20	-0.56	0.30	580	10
21	-0.47	0.29	583	10
22	-0.39	0.29	586	10
23	-0.30	0.29	589	10
24	-0.22	0.29	592	10
25	-0.14	0.28	595	10
26	-0.06	0.28	598	10
27	0.02	0.28	601	10
28	0.10	0.28	604	10
29	0.18	0.28	606	10
30	0.26	0.28	609	10
31	0.34	0.28	612	10
32	0.42	0.28	615	10
33	0.50	0.28	618	10
34	0.58	0.29	620	10
35	0.66	0.29	623	10
36	0.75	0.29	626	10
37	0.83	0.29	629	10
38	0.92	0.29	632	10
39	1.00	0.30	635	10
40	1.09	0.30	638	11
41	1.19	0.31	641	11
42	1.28	0.31	645	11

C.3a: Grade 5–Grade 6 Listening and Speaking (continued)

Raw Score	Measure	SE	Scale Score	SE(SS)
43	1.38	0.32	648	11
44	1.48	0.32	652	11
45	1.59	0.33	656	12
46	1.70	0.34	660	12
47	1.82	0.35	664	12
48	1.95	0.36	668	13
49	2.08	0.37	673	13
50	2.23	0.39	678	14
51	2.39	0.41	684	14
52	2.56	0.43	690	15
53	2.76	0.46	697	16
54	2.98	0.49	704	17
55	3.25	0.54	714	19
56	3.59	0.62	725	22
57	4.04	0.74	741	26
58	4.78	1.02	767	36
59	6.20	2.01	817	70

C.3b: Grade 5–Grade 6 Reading and Writing

Raw Score	Measure	SE	Scale Score	SE(SS)
0	-4.83	2.02	431	71
1	-3.40	1.03	481	36
2	-2.64	0.75	507	26
3	-2.18	0.62	524	22
4	-1.84	0.55	536	19
5	-1.56	0.50	545	18
6	-1.32	0.47	554	16
7	-1.12	0.44	561	15
8	-0.93	0.42	567	15
9	-0.76	0.40	573	14
10	-0.61	0.39	579	14
11	-0.46	0.37	584	13
12	-0.32	0.36	589	13
13	-0.20	0.36	593	12
14	-0.07	0.35	597	12
15	0.05	0.34	602	12
16	0.16	0.33	606	12
17	0.27	0.33	609	12
18	0.38	0.33	613	11
19	0.48	0.32	617	11
20	0.59	0.32	620	11
21	0.69	0.32	624	11
22	0.78	0.31	627	11
23	0.88	0.31	631	11
24	0.98	0.31	634	11
25	1.08	0.31	638	11
26	1.17	0.31	641	11
27	1.27	0.31	644	11
28	1.37	0.31	648	11
29	1.46	0.31	651	11
30	1.56	0.31	655	11
31	1.66	0.31	658	11
32	1.76	0.32	661	11
33	1.86	0.32	665	11
34	1.96	0.32	669	11
35	2.07	0.33	672	11
36	2.18	0.33	676	12
37	2.29	0.34	680	12
38	2.40	0.34	684	12
39	2.52	0.35	688	12
40	2.65	0.36	693	13
41	2.78	0.37	697	13
42	2.92	0.38	702	13
43	3.08	0.40	708	14
44	3.24	0.41	713	14
45	3.42	0.44	720	15
46	3.62	0.46	727	16

C.3b: Grade 5–Grade 6 Reading and Writing (continued)

Raw Score	Measure	SE	Scale Score	SE(SS)
47	3.85	0.50	735	17
48	4.13	0.55	744	19
49	4.47	0.62	756	22
50	4.93	0.75	772	26
51	5.68	1.03	799	36
52	7.11	2.02	849	71

C.4a: Grade 7–Grade 8 Listening and Speaking

Raw Score	Measure	SE	Scale Score	SE(SS)
0	-4.91	2.01	428	70
1	-3.50	1.02	478	36
2	-2.77	0.73	503	26
3	-2.33	0.61	518	21
4	-2.01	0.53	530	19
5	-1.75	0.49	539	17
6	-1.53	0.45	546	16
7	-1.34	0.42	553	15
8	-1.17	0.40	559	14
9	-1.02	0.38	564	13
10	-0.87	0.37	569	13
11	-0.74	0.36	574	12
12	-0.62	0.35	578	12
13	-0.50	0.34	582	12
14	-0.39	0.33	586	12
15	-0.29	0.32	590	11
16	-0.18	0.32	594	11
17	-0.09	0.31	597	11
18	0.01	0.31	600	11
19	0.10	0.30	604	11
20	0.19	0.30	607	10
21	0.28	0.30	610	10
22	0.37	0.29	613	10
23	0.45	0.29	616	10
24	0.54	0.29	619	10
25	0.62	0.29	622	10
26	0.70	0.29	625	10
27	0.78	0.29	627	10
28	0.86	0.29	630	10
29	0.95	0.28	633	10
30	1.03	0.29	636	10
31	1.11	0.29	639	10
32	1.19	0.29	642	10
33	1.27	0.29	645	10
34	1.36	0.29	647	10
35	1.44	0.29	650	10
36	1.53	0.29	653	10
37	1.61	0.30	656	10
38	1.70	0.30	660	10
39	1.79	0.30	663	11
40	1.89	0.31	666	11
41	1.98	0.31	669	11
42	2.08	0.32	673	11
43	2.18	0.32	676	11
44	2.29	0.33	680	12
45	2.40	0.34	684	12

C.4a: Grade 7–Grade 8 Listening and Speaking (continued)

Raw Score	Measure	SE	Scale Score	SE(SS)
46	2.51	0.35	688	12
47	2.64	0.35	692	12
48	2.77	0.37	697	13
49	2.90	0.38	702	13
50	3.05	0.39	707	14
51	3.22	0.41	713	14
52	3.39	0.43	719	15
53	3.59	0.46	726	16
54	3.82	0.50	734	17
55	4.09	0.54	743	19
56	4.43	0.62	755	22
57	4.88	0.74	771	26
58	5.62	1.02	797	36
59	7.04	2.01	847	70

C.4b: Grade 7-Grade 8 Reading and Writing

Raw Score	Measure	SE	Scale Score	SE(SS)
0	-4.98	2.01	426	70
1	-3.56	1.03	475	36
2	-2.81	0.74	502	26
3	-2.36	0.62	517	22
4	-2.02	0.55	529	19
5	-1.75	0.50	539	17
6	-1.53	0.46	547	16
7	-1.33	0.43	554	15
8	-1.15	0.41	560	14
9	-0.99	0.39	565	14
10	-0.84	0.38	571	13
11	-0.70	0.37	576	13
12	-0.57	0.36	580	12
13	-0.44	0.35	585	12
14	-0.32	0.34	589	12
15	-0.21	0.33	593	12
16	-0.10	0.33	597	12
17	0.01	0.32	600	11
18	0.11	0.32	604	11
19	0.21	0.32	607	11
20	0.31	0.31	611	11
21	0.41	0.31	614	11
22	0.51	0.31	618	11
23	0.60	0.31	621	11
24	0.70	0.31	624	11
25	0.79	0.31	628	11
26	0.88	0.31	631	11
27	0.98	0.31	634	11
28	1.07	0.31	638	11
29	1.17	0.31	641	11
30	1.26	0.31	644	11
31	1.36	0.31	648	11
32	1.46	0.31	651	11
33	1.56	0.32	654	11
34	1.66	0.32	658	11
35	1.76	0.33	662	11
36	1.87	0.33	665	12
37	1.98	0.34	669	12
38	2.10	0.34	673	12
39	2.22	0.35	678	12
40	2.35	0.36	682	13
41	2.48	0.37	687	13
42	2.63	0.39	692	14
43	2.78	0.41	697	14
44	2.96	0.43	703	15
45	3.15	0.45	710	16
46	3.37	0.49	718	17
47	3.63	0.54	727	19
48	3.96	0.61	739	21
49	4.41	0.74	754	26
50	5.14	1.02	780	36
51	6.56	2.01	830	70

C.5a: Grade 9–Grade 12 Listening and Speaking

Raw Score	Measure	SE	Scale Score	SE(SS)
0	-4.73	2.01	435	70
1	-3.32	1.02	484	36
2	-2.59	0.73	509	26
3	-2.15	0.61	525	21
4	-1.83	0.53	536	19
5	-1.57	0.48	545	17
6	-1.35	0.45	553	16
7	-1.16	0.42	559	15
8	-0.99	0.40	565	14
9	-0.84	0.38	571	13
10	-0.70	0.37	576	13
11	-0.57	0.36	580	12
12	-0.45	0.35	584	12
13	-0.33	0.34	588	12
14	-0.22	0.33	592	11
15	-0.11	0.32	596	11
16	-0.01	0.32	600	11
17	0.09	0.31	603	11
18	0.18	0.31	606	11
19	0.27	0.30	610	11
20	0.37	0.30	613	10
21	0.45	0.30	616	10
22	0.54	0.29	619	10
23	0.63	0.29	622	10
24	0.71	0.29	625	10
25	0.79	0.29	628	10
26	0.87	0.29	631	10
27	0.96	0.29	633	10
28	1.04	0.28	636	10
29	1.12	0.28	639	10
30	1.20	0.28	642	10
31	1.28	0.28	645	10
32	1.36	0.29	648	10
33	1.44	0.29	651	10
34	1.53	0.29	653	10
35	1.61	0.29	656	10
36	1.69	0.29	659	10
37	1.78	0.29	662	10
38	1.86	0.30	665	10
39	1.95	0.30	668	10
40	2.04	0.30	671	11
41	2.14	0.31	675	11
42	2.23	0.31	678	11
43	2.33	0.32	682	11
44	2.43	0.32	685	11
45	2.54	0.33	689	11
46	2.65	0.34	693	12

C.5a: Grade 9–Grade 12 Listening and Speaking (continued)

Raw Score	Measure	SE	Scale Score	SE(SS)
47	2.76	0.35	697	12
48	2.89	0.36	701	12
49	3.02	0.37	706	13
50	3.16	0.38	711	13
51	3.31	0.40	716	14
52	3.48	0.42	722	15
53	3.67	0.45	728	16
54	3.89	0.48	736	17
55	4.14	0.53	745	19
56	4.47	0.61	756	21
57	4.91	0.73	772	26
58	5.63	1.02	797	36
59	7.05	2.01	847	70

C.5b: Grade 9–Grade 12 Reading and Writing

Raw Score	Measure	SE	Scale Score	SE(SS)
0	-3.89	2.01	464	70
1	-2.48	1.02	513	36
2	-1.75	0.73	539	26
3	-1.30	0.61	554	21
4	-0.98	0.54	566	19
5	-0.72	0.49	575	17
6	-0.49	0.45	583	16
7	-0.30	0.43	589	15
8	-0.13	0.40	595	14
9	0.03	0.39	601	14
10	0.17	0.37	606	13
11	0.31	0.36	611	13
12	0.43	0.35	615	12
13	0.55	0.34	619	12
14	0.67	0.34	623	12
15	0.78	0.33	627	11
16	0.88	0.32	631	11
17	0.99	0.32	635	11
18	1.09	0.31	638	11
19	1.18	0.31	641	11
20	1.28	0.31	645	11
21	1.37	0.31	648	11
22	1.47	0.30	651	11
23	1.56	0.30	655	11
24	1.65	0.30	658	11
25	1.74	0.30	661	11
26	1.83	0.30	664	10
27	1.92	0.30	667	10
28	2.01	0.30	670	11
29	2.10	0.30	673	11
30	2.19	0.30	677	11
31	2.28	0.30	680	11
32	2.38	0.31	683	11
33	2.47	0.31	686	11
34	2.57	0.31	690	11
35	2.66	0.32	693	11
36	2.77	0.32	697	11
37	2.87	0.33	700	11
38	2.98	0.33	704	12
39	3.09	0.34	708	12
40	3.21	0.35	712	12
41	3.33	0.36	717	12
42	3.46	0.37	721	13
43	3.61	0.38	726	13
44	3.76	0.40	732	14
45	3.93	0.42	737	15
46	4.12	0.45	744	16
47	4.33	0.48	752	17
48	4.59	0.53	761	19
49	4.91	0.61	772	21
50	5.35	0.73	787	26
51	6.08	1.02	813	36
52	7.49	2.01	862	70

APPENDIX D: IRT STATISTICS

D.1a: Kindergarten–Grade 1 Listening and Speaking

	Item	N-count	Rasch Difficulty	SE	MNSQ Infit	MNSQ Outfit	Misfit
LI	1	45,063	-3.20	0.02	1.12	1.28	--
	2	45,063	-2.79	0.02	1.12	1.24	--
	3	45,063	-2.52	0.02	0.95	0.85	--
	4	45,063	-3.73	0.03	0.92	0.73	--
	5	45,063	-2.90	0.02	0.92	0.64	--
	6	45,063	-3.15	0.02	1.14	2.03	--
	7	45,063	-3.66	0.03	1.06	1.68	--
	8	45,063	-3.19	0.02	0.91	0.85	--
	9	45,063	-2.54	0.02	1.03	0.97	--
	10	45,063	-1.77	0.01	1.11	1.22	--
	11	45,063	-2.21	0.02	1.17	1.38	--
	12	45,063	-1.24	0.01	1.17	1.29	--
	13	45,063	-1.60	0.01	1.16	1.14	--
	14	45,063	-1.77	0.01	0.99	0.89	--
	15	45,063	-1.73	0.01	1.44	2.45	--
	16	45,063	-3.31	0.02	0.86	0.71	--
	17	45,063	-1.95	0.01	1.22	1.30	--
	18	45,063	-1.51	0.01	1.10	1.20	--
	19	45,063	-1.28	0.01	1.12	1.29	--
	20	45,063	-1.73	0.01	1.03	0.98	--
	21	45,063	-0.49	0.01	1.33	1.56	--
	22	45,063	-0.54	0.01	0.96	0.94	--
	23	45,063	-0.91	0.01	1.16	1.22	--
	24	45,063	-1.00	0.01	1.43	1.88	--
SP	1	45,063	-1.52	0.01	0.88	0.85	--
	2	45,063	-1.38	0.01	0.82	0.81	--
	3	45,063	-1.79	0.01	0.83	0.75	--
	4	45,063	-1.76	0.01	0.83	0.78	--
	5	45,063	-1.35	0.01	0.90	0.89	--
	6	45,063	-1.01	0.01	0.84	0.85	--
	7	45,063	-2.16	0.01	0.79	0.73	--
	8	45,063	-1.96	0.01	0.79	0.72	--
	9	45,063	-2.39	0.01	0.81	0.70	--
	10	45,063	-1.53	0.01	0.80	0.76	--
	11	45,063	-1.94	0.01	0.82	0.77	--
	12	45,063	-1.86	0.01	0.79	0.70	--
	13	45,063	-1.73	0.01	0.80	0.72	--
	14	45,063	-1.84	0.01	0.86	0.82	--
	15	45,063	-1.30	0.01	0.82	0.79	--
	16	45,063	-1.92	0.01	1.09	1.21	--

Note: “√” indicates that the item was flagged as misfit; “--” indicates that the item was not flagged.

D.1b: Kindergarten–Grade 1 Reading and Writing

	Item	N-count	Rasch Difficulty	SE	MNSQ Infit	MNSQ Outfit	Misfit
RD	1	42,701	-3.53	0.02	0.84	0.65	--
	2	42,701	-2.96	0.01	1.20	1.58	--
	3	42,701	-3.09	0.01	0.84	0.64	--
	4	42,701	-2.84	0.01	1.03	0.95	--
	5	42,701	-3.66	0.02	0.95	0.80	--
	6	42,701	-2.97	0.01	0.83	0.65	--
	7	42,701	-1.71	0.01	0.92	0.84	--
	8	42,701	-3.34	0.02	1.02	0.79	--
	9	42,701	-2.03	0.01	1.25	1.47	--
	10	42,701	-1.88	0.01	0.93	0.80	--
	11	42,701	-1.43	0.01	1.02	1.01	--
	12	42,701	-1.59	0.01	1.16	1.17	--
	13	42,701	-1.52	0.01	1.13	1.19	--
	14	42,701	-1.67	0.01	1.11	1.08	--
	15	42,701	-2.55	0.01	1.14	1.25	--
WR	1	42,701	-3.83	0.02	0.97	0.80	--
	2	42,701	-4.50	0.02	0.93	0.60	--
	3	42,701	-2.82	0.01	0.93	0.83	--
	4	42,701	-2.88	0.01	1.17	1.75	--
	5	42,701	-3.07	0.01	1.06	1.78	--
	6	42,701	-2.19	0.01	1.00	0.93	--
	7	42,701	-3.53	0.01	1.14	1.24	--
	8	42,701	-3.48	0.01	1.31	1.49	--
	9	42,701	-3.16	0.01	1.26	4.44	--
	10	42,701	-2.75	0.01	1.17	1.77	--
	11	42,701	-1.50	0.01	0.84	0.81	--
	12	42,701	-2.24	0.01	0.78	0.78	--
	13	42,701	-1.94	0.01	0.59	0.49	--
	14	42,701	-1.19	0.01	0.62	0.59	--
	15	42,701	-0.65	0.01	0.72	0.71	--

D.2a: Grade 2–Grade 4 Listening and Speaking

	Item	N-count	Rasch Difficulty	SE	MNSQ Infit	MNSQ Outfit	Misfit
LI	1	56,489	-1.79	0.02	0.93	0.63	--
	2	56,489	-0.24	0.01	1.24	1.20	--
	3	56,489	-0.86	0.02	0.95	0.65	--
	4	56,489	-1.28	0.02	1.03	0.96	--
	5	56,489	-2.22	0.03	0.91	0.64	--
	6	56,489	-0.74	0.02	1.24	1.97	--
	7	56,489	0.43	0.01	1.16	1.29	--
	8	56,489	1.75	0.01	1.29	1.56	--
	9	56,489	1.22	0.01	1.16	1.23	--
	10	56,489	-1.14	0.02	0.99	0.89	--
	11	56,489	-1.01	0.02	1.00	0.76	--
	12	56,489	1.15	0.01	1.11	1.18	--
	13	56,489	1.61	0.01	1.19	1.37	--
	14	56,489	0.61	0.01	1.11	1.20	--
	15	56,489	-0.32	0.01	1.02	1.13	--
	16	56,489	0.86	0.01	1.13	1.22	--
	17	56,489	0.38	0.01	1.20	1.36	--
	18	56,489	0.40	0.01	1.11	1.19	--
	19	56,489	2.26	0.01	1.24	1.63	--
	20	56,489	-0.81	0.02	1.10	1.33	--
	21	56,489	-0.42	0.02	0.91	0.91	--
	22	56,489	-0.21	0.01	1.15	1.28	--
	23	56,489	-0.41	0.02	1.16	1.13	--
	24	56,489	-0.18	0.01	1.13	1.14	--
SP	1	56,489	-0.43	0.01	0.89	0.88	--
	2	56,489	0.32	0.01	0.87	0.81	--
	3	56,489	-0.44	0.01	0.96	0.93	--
	4	56,489	-0.13	0.01	0.85	0.77	--
	5	56,489	0.34	0.01	0.83	0.79	--
	6	56,489	0.44	0.01	0.78	0.78	--
	7	56,489	-0.70	0.01	0.76	0.65	--
	8	56,489	-0.80	0.01	0.79	0.70	--
	9	56,489	-0.43	0.01	0.74	0.64	--
	10	56,489	-0.29	0.01	0.79	0.69	--
	11	56,489	-0.70	0.01	0.78	0.68	--
	12	56,489	-0.25	0.01	0.75	0.68	--
	13	56,489	-0.15	0.01	0.83	0.71	--
	14	56,489	-0.92	0.01	1.00	0.87	--
	15	56,489	-0.21	0.01	0.87	0.78	--
	16	56,489	-0.03	0.01	0.83	0.73	--

D.2b: Grade 2–Grade 4 Reading and Writing

	Item	N-count	Rasch Difficulty	SE	MNSQ Infit	MNSQ Outfit	Misfit
RD	1	56,168	-4.16	0.04	0.96	0.58	--
	2	56,168	-1.72	0.02	0.95	0.76	--
	3	56,168	-2.20	0.02	0.96	1.00	--
	4	56,168	-2.13	0.02	1.00	0.76	--
	5	56,168	-2.67	0.02	0.87	0.43	--
	6	56,168	-2.64	0.02	0.92	0.48	--
	7	56,168	-0.26	0.01	0.95	0.88	--
	8	56,168	-0.63	0.01	1.09	1.28	--
	9	56,168	1.55	0.01	1.11	1.25	--
	10	56,168	0.60	0.01	1.07	1.11	--
	11	56,168	-0.11	0.01	1.02	1.00	--
	12	56,168	-0.70	0.01	0.84	0.67	--
	13	56,168	0.04	0.01	0.91	0.82	--
	14	56,168	0.15	0.01	1.14	1.18	--
	15	56,168	-0.02	0.01	0.94	0.87	--
	16	56,168	1.26	0.01	1.02	1.08	--
	17	56,168	0.47	0.01	0.85	0.79	--
	18	56,168	-0.76	0.01	1.34	1.86	--
	19	56,168	1.46	0.01	1.17	1.34	--
	20	56,168	-0.54	0.01	1.03	1.04	--
	21	56,168	0.12	0.01	0.79	0.67	--
	22	56,168	0.87	0.01	1.12	1.20	--
	23	56,168	0.21	0.01	1.00	1.03	--
	24	56,168	-0.23	0.01	0.90	0.85	--
WR	1	56,168	-1.95	0.02	0.96	0.77	--
	2	56,168	-3.07	0.03	0.93	0.62	--
	3	56,168	-0.27	0.01	0.98	0.95	--
	4	56,168	0.21	0.01	1.19	1.28	--
	5	56,168	-0.67	0.01	0.98	0.96	--
	6	56,168	-0.47	0.01	0.91	0.80	--
	7	56,168	0.34	0.01	0.98	0.97	--
	8	56,168	1.04	0.01	1.16	1.24	--
	9	56,168	-0.46	0.01	0.95	0.85	--
	10	56,168	0.20	0.01	1.00	0.98	--
	11	56,168	0.33	0.01	0.98	0.98	--
	12	56,168	1.35	0.01	1.13	1.25	--
	13	56,168	-1.79	0.01	0.93	0.91	--
	14	56,168	-1.05	0.01	0.93	0.87	--
	15	56,168	-0.07	0.01	0.97	0.96	--
	16	56,168	1.11	0.01	0.89	0.89	--

D.3a: Grade 5–Grade 6 Listening and Speaking

	Item	N-count	Rasch Difficulty	SE	MNSQ Infit	MNSQ Outfit	Misfit
LI	1	26,761	-2.78	0.05	0.94	0.55	--
	2	26,761	-1.43	0.03	1.08	0.82	--
	3	26,761	0.73	0.02	1.28	1.41	--
	4	26,761	-0.78	0.02	0.83	0.56	--
	5	26,761	-0.79	0.02	0.97	1.00	--
	6	26,761	-0.64	0.02	1.09	1.19	--
	7	26,761	0.39	0.02	1.15	1.24	--
	8	26,761	-0.14	0.02	1.07	1.10	--
	9	26,761	0.35	0.02	1.23	1.35	--
	10	26,761	1.00	0.02	1.20	1.27	--
	11	26,761	0.86	0.02	1.24	1.26	--
	12	26,761	0.20	0.02	1.15	1.13	--
	13	26,761	2.18	0.01	1.15	1.44	--
	14	26,761	1.68	0.01	1.22	1.38	--
	15	26,761	2.68	0.01	1.09	1.39	--
	16	26,761	0.22	0.02	1.00	0.97	--
	17	26,761	0.42	0.02	1.21	1.40	--
	18	26,761	1.85	0.01	1.28	1.65	--
	19	26,761	-0.37	0.02	0.89	0.71	--
	20	26,761	0.37	0.02	1.47	1.88	--
	21	26,761	2.47	0.01	1.19	1.61	--
	22	26,761	0.10	0.02	0.95	0.91	--
	23	26,761	2.51	0.01	1.29	1.84	--
	24	26,761	0.47	0.02	1.26	1.43	--
	25	26,761	0.67	0.02	1.16	1.22	--
SP	1	26,761	0.18	0.01	0.76	0.71	--
	2	26,761	0.15	0.01	0.76	0.72	--
	3	26,761	0.22	0.01	0.71	0.60	--
	4	26,761	0.38	0.01	0.86	0.91	--
	5	26,761	0.33	0.01	0.79	0.74	--
	6	26,761	0.18	0.01	0.71	0.72	--
	7	26,761	-0.72	0.02	0.72	0.62	--
	8	26,761	-0.43	0.02	0.77	0.70	--
	9	26,761	-1.02	0.02	0.68	0.49	--
	10	26,761	-0.02	0.01	0.80	0.74	--
	11	26,761	-0.11	0.01	0.66	0.58	--
	12	26,761	-0.43	0.02	0.75	0.59	--
	13	26,761	0.75	0.01	0.69	0.58	--
	14	26,761	0.25	0.01	0.90	0.81	--
	15	26,761	0.20	0.01	0.71	0.62	--
	16	26,761	0.38	0.01	0.68	0.57	--

D.3b: Grade 5–Grade 6 Reading and Writing

	Item	N-count	Rasch Difficulty	SE	MNSQ Infit	MNSQ Outfit	Misfit
RD	1	25,854	-1.07	0.03	0.97	0.84	--
	2	25,854	-1.55	0.03	0.92	0.72	--
	3	25,854	-0.37	0.02	1.03	1.06	--
	4	25,854	-1.33	0.03	1.04	1.03	--
	5	25,854	1.64	0.01	1.03	1.04	--
	6	25,854	0.53	0.02	1.05	1.20	--
	7	25,854	0.16	0.02	0.92	0.79	--
	8	25,854	0.28	0.02	0.88	0.75	--
	9	25,854	1.36	0.01	1.15	1.24	--
	10	25,854	1.41	0.01	0.84	0.78	--
	11	25,854	1.21	0.02	0.97	0.92	--
	12	25,854	1.73	0.01	0.94	0.91	--
	13	25,854	0.75	0.02	0.86	0.75	--
	14	25,854	2.03	0.01	0.96	0.96	--
	15	25,854	0.78	0.02	0.89	0.75	--
	16	25,854	0.99	0.02	0.87	0.78	--
	17	25,854	1.97	0.01	0.98	0.97	--
	18	25,854	2.65	0.01	1.03	1.11	--
	19	25,854	1.80	0.01	0.97	0.95	--
	20	25,854	0.93	0.02	0.94	0.86	--
	21	25,854	0.97	0.02	0.97	0.89	--
	22	25,854	2.22	0.01	1.05	1.07	--
	23	25,854	2.86	0.01	1.07	1.19	--
	24	25,854	2.04	0.01	1.10	1.15	--
	25	25,854	2.00	0.01	1.11	1.14	--
	26	25,854	2.02	0.01	1.00	1.01	--
	27	25,854	2.48	0.01	1.07	1.14	--
WR	1	25,854	-0.89	0.03	0.94	0.74	--
	2	25,854	0.56	0.02	1.20	1.31	--
	3	25,854	0.49	0.02	1.02	1.06	--
	4	25,854	-0.09	0.02	0.88	0.79	--
	5	25,854	0.27	0.02	0.95	0.84	--
	6	25,854	2.24	0.01	1.02	1.07	--
	7	25,854	1.42	0.01	1.14	1.23	--
	8	25,854	0.39	0.02	0.89	0.85	--
	9	25,854	0.90	0.02	0.91	0.89	--
	10	25,854	1.36	0.01	1.03	1.04	--
	11	25,854	2.57	0.01	1.12	1.20	--
	12	25,854	0.35	0.02	0.98	1.04	--
	13	25,854	2.03	0.01	1.07	1.10	--
	14	25,854	0.50	0.02	1.00	1.04	--
	15	25,854	2.19	0.01	1.19	1.30	--
	16	25,854	0.73	0.01	1.20	1.25	--
	17	25,854	1.01	0.01	0.92	0.90	--
	18	25,854	1.52	0.01	1.12	1.15	--
	19	25,854	2.26	0.01	0.87	0.87	--

D.4a: Grade 7–Grade 8 Listening and Speaking

	Item	N-count	Rasch Difficulty	SE	MNSQ Infit	MNSQ Outfit	Misfit
LI	1	23,840	1.41	0.02	1.54	1.76	--
	2	23,840	-0.44	0.03	1.11	1.03	--
	3	23,840	3.32	0.02	1.15	1.68	--
	4	23,840	1.42	0.02	1.08	1.06	--
	5	23,840	1.97	0.02	1.08	1.12	--
	6	23,840	-0.10	0.02	0.92	0.74	--
	7	23,840	0.62	0.02	0.94	0.78	--
	8	23,840	0.23	0.02	0.95	0.79	--
	9	23,840	1.02	0.02	1.19	1.24	--
	10	23,840	2.59	0.02	1.24	1.40	--
	11	23,840	3.23	0.02	1.20	1.55	--
	12	23,840	2.44	0.02	1.30	1.54	--
	13	23,840	0.55	0.02	1.52	2.34	--
	14	23,840	2.42	0.02	1.14	1.32	--
	15	23,840	3.42	0.02	1.04	1.42	--
	16	23,840	0.75	0.02	0.99	0.93	--
	17	23,840	1.50	0.02	1.12	1.16	--
	18	23,840	0.20	0.02	1.04	0.93	--
	19	23,840	0.59	0.02	1.09	1.07	--
	20	23,840	1.79	0.02	1.04	1.05	--
	21	23,840	0.70	0.02	1.13	1.13	--
	22	23,840	1.11	0.02	1.13	1.05	--
	23	23,840	1.85	0.02	1.10	1.20	--
	24	23,840	1.43	0.02	1.25	1.37	--
	25	23,840	3.31	0.02	1.44	2.15	--
SP	1	23,840	0.36	0.02	0.82	0.76	--
	2	23,840	0.45	0.01	0.82	0.77	--
	3	23,840	0.29	0.02	0.75	0.64	--
	4	23,840	0.45	0.01	0.73	0.62	--
	5	23,840	0.47	0.01	0.95	0.95	--
	6	23,840	0.98	0.01	0.69	0.70	--
	7	23,840	0.33	0.02	0.70	0.63	--
	8	23,840	-0.10	0.02	0.72	0.57	--
	9	23,840	0.77	0.01	0.73	0.68	--
	10	23,840	0.93	0.01	0.66	0.57	--
	11	23,840	0.88	0.01	0.74	0.71	--
	12	23,840	0.87	0.01	0.66	0.57	--
	13	23,840	0.82	0.01	0.74	0.62	--
	14	23,840	0.92	0.01	0.75	0.64	--
	15	23,840	0.94	0.01	0.82	0.67	--
	16	23,840	0.79	0.01	0.62	0.47	--

D.4b: Grade 7–Grade 8 Reading and Writing

	Item	N-count	Rasch Difficulty	SE	MNSQ Infit	MNSQ Outfit	Misfit
RD	1	23,024	-1.68	0.03	0.92	0.66	--
	2	23,024	-0.19	0.02	0.90	0.99	--
	3	23,024	-1.38	0.03	0.94	0.78	--
	4	23,024	0.39	0.02	1.03	0.97	--
	5	23,024	2.11	0.01	1.08	1.11	--
	6	23,024	1.46	0.02	1.09	1.13	--
	7	23,024	1.05	0.02	0.88	0.83	--
	8	23,024	0.40	0.02	1.07	1.23	--
	9	23,024	1.04	0.02	0.87	0.80	--
	10	23,024	1.13	0.02	0.94	0.88	--
	11	23,024	1.19	0.02	0.96	0.93	--
	12	23,024	0.17	0.02	0.99	0.87	--
	13	23,024	0.45	0.02	0.92	0.80	--
	14	23,024	1.87	0.01	1.26	1.37	--
	15	23,024	1.31	0.02	0.98	1.00	--
	16	23,024	0.63	0.02	1.05	0.99	--
	17	23,024	1.48	0.02	1.16	1.19	--
	18	23,024	0.69	0.02	0.90	0.78	--
	19	23,024	1.03	0.02	0.98	0.93	--
	20	23,024	2.08	0.01	0.93	0.93	--
	21	23,024	0.60	0.02	0.89	0.75	--
	22	23,024	1.62	0.02	0.95	0.94	--
	23	23,024	1.51	0.02	1.03	1.03	--
	24	23,024	1.00	0.02	1.02	0.98	--
	25	23,024	2.08	0.01	1.12	1.19	--
	26	23,024	1.69	0.02	0.93	0.93	--
WR	1	23,024	1.18	0.02	1.06	1.08	--
	2	23,024	-0.51	0.02	1.06	1.17	--
	3	23,024	-0.77	0.02	0.96	0.88	--
	4	23,024	1.09	0.02	0.94	0.92	--
	5	23,024	0.40	0.02	0.90	0.80	--
	6	23,024	0.98	0.02	1.01	1.02	--
	7	23,024	1.71	0.02	1.16	1.21	--
	8	23,024	-0.64	0.02	0.94	0.81	--
	9	23,024	1.18	0.02	1.01	1.00	--
	10	23,024	0.08	0.02	0.93	0.88	--
	11	23,024	0.67	0.02	0.96	0.94	--
	12	23,024	1.49	0.02	1.02	1.04	--
	13	23,024	1.26	0.02	1.00	1.02	--
	14	23,024	1.82	0.01	1.28	1.38	--
	15	23,024	1.98	0.01	1.01	1.02	--
	16	23,024	-0.28	0.01	0.94	0.97	--
	17	23,024	-0.21	0.01	0.95	0.96	--
	18	23,024	0.82	0.01	0.96	0.97	--
	19	23,024	1.41	0.01	0.86	0.86	--

D.5a: Grade 9–Grade 12 Listening and Speaking

	Item	N-count	Rasch Difficulty	SE	MNSQ Infit	MNSQ Outfit	Misfit
LI	1	37,718	0.72	0.02	1.26	1.55	--
	2	37,718	0.97	0.02	1.09	1.01	--
	3	37,718	-0.29	0.02	0.98	0.94	--
	4	37,718	3.39	0.01	1.12	1.41	--
	5	37,718	2.03	0.01	0.99	0.96	--
	6	37,718	0.76	0.02	1.26	1.85	--
	7	37,718	0.93	0.02	1.04	1.23	--
	8	37,718	2.00	0.01	0.94	0.88	--
	9	37,718	1.31	0.01	1.04	1.02	--
	10	37,718	1.20	0.01	1.13	1.29	--
	11	37,718	1.77	0.01	1.34	1.49	--
	12	37,718	2.50	0.01	1.15	1.25	--
	13	37,718	2.90	0.01	1.30	1.55	--
	14	37,718	1.16	0.01	1.19	1.34	--
	15	37,718	1.84	0.01	1.12	1.21	--
	16	37,718	-0.11	0.02	0.95	0.68	--
	17	37,718	0.42	0.02	1.04	1.12	--
	18	37,718	1.19	0.01	1.15	1.26	--
	19	37,718	2.34	0.01	1.23	1.36	--
	20	37,718	0.25	0.02	1.17	1.17	--
	21	37,718	2.36	0.01	1.20	1.33	--
	22	37,718	0.36	0.02	1.33	2.03	--
	23	37,718	1.82	0.01	1.28	1.51	--
	24	37,718	1.50	0.01	1.21	1.39	--
	25	37,718	2.46	0.01	1.28	1.46	--
SP	1	37,718	1.18	0.01	0.77	0.71	--
	2	37,718	0.62	0.01	0.79	0.70	--
	3	37,718	0.69	0.01	0.77	0.66	--
	4	37,718	1.74	0.01	0.85	0.81	--
	5	37,718	0.76	0.01	0.75	0.68	--
	6	37,718	1.35	0.01	0.70	0.70	--
	7	37,718	0.87	0.01	0.74	0.66	--
	8	37,718	0.81	0.01	0.71	0.64	--
	9	37,718	0.42	0.01	0.74	0.64	--
	10	37,718	0.52	0.01	0.79	0.70	--
	11	37,718	1.05	0.01	0.73	0.69	--
	12	37,718	1.07	0.01	0.76	0.62	--
	13	37,718	1.04	0.01	0.75	0.66	--
	14	37,718	0.92	0.01	0.77	0.64	--
	15	37,718	0.63	0.01	0.86	0.77	--
	16	37,718	1.28	0.01	0.89	0.89	--

D.5b: Grade 9–Grade 12 Reading and Writing

	Item	N-count	Rasch Difficulty	SE	MNSQ Infit	MNSQ Outfit	Misfit
RD	1	37,005	0.82	0.01	0.97	0.90	--
	2	37,005	0.52	0.02	1.12	1.18	--
	3	37,005	1.66	0.01	0.97	0.94	--
	4	37,005	1.72	0.01	0.93	0.89	--
	5	37,005	1.87	0.01	0.90	0.84	--
	6	37,005	1.96	0.01	0.93	0.89	--
	7	37,005	1.86	0.01	0.97	0.95	--
	8	37,005	3.07	0.01	1.18	1.26	--
	9	37,005	2.60	0.01	1.02	1.03	--
	10	37,005	1.67	0.01	1.02	1.00	--
	11	37,005	2.97	0.01	1.05	1.09	--
	12	37,005	1.95	0.01	0.85	0.78	--
	13	37,005	2.97	0.01	1.03	1.04	--
	14	37,005	0.67	0.02	0.99	1.02	--
	15	37,005	1.25	0.01	1.10	1.20	--
	16	37,005	1.40	0.01	1.00	1.00	--
	17	37,005	0.38	0.02	0.93	0.74	--
	18	37,005	2.69	0.01	1.14	1.18	--
	19	37,005	2.60	0.01	1.17	1.23	--
	20	37,005	2.02	0.01	0.84	0.78	--
	21	37,005	1.99	0.01	1.01	0.96	--
	22	37,005	2.51	0.01	1.04	1.05	--
	23	37,005	2.67	0.01	1.27	1.37	--
	24	37,005	2.39	0.01	1.20	1.28	--
	25	37,005	2.68	0.01	0.89	0.87	--
	26	37,005	3.79	0.01	1.07	1.18	--
	27	37,005	2.80	0.01	1.04	1.06	--
WR	1	37,005	1.30	0.01	0.95	0.92	--
	2	37,005	0.13	0.02	1.12	1.02	--
	3	37,005	-0.17	0.02	0.95	0.87	--
	4	37,005	0.48	0.02	0.83	0.65	--
	5	37,005	1.43	0.01	0.95	0.91	--
	6	37,005	2.06	0.01	1.02	1.05	--
	7	37,005	2.10	0.01	0.95	0.93	--
	8	37,005	0.54	0.02	1.00	1.01	--
	9	37,005	2.19	0.01	1.03	1.05	--
	10	37,005	1.36	0.01	0.95	0.91	--
	11	37,005	2.28	0.01	1.01	1.03	--
	12	37,005	2.02	0.01	1.13	1.14	--
	13	37,005	1.72	0.01	0.86	0.79	--
	14	37,005	1.95	0.01	1.06	1.12	--
	15	37,005	2.71	0.01	0.97	0.98	--
	16	37,005	1.06	0.01	0.95	0.93	--
	17	37,005	1.33	0.01	0.95	0.95	--
	18	37,005	1.95	0.01	0.95	0.94	--
	19	37,005	2.08	0.01	0.87	0.87	--

APPENDIX E: ITEM INFORMATION AT DIFFERENT CUT POINTS BY GRADES

E.1: Kindergarten–Grade 1 Grade K Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	-3.247	0.22	0.07	0.02
	2	-2.787	0.25	0.10	0.03
	3	-3.696	0.18	0.05	0.01
	4	-3.252	0.22	0.07	0.02
	5	-3.734	0.18	0.04	0.01
	6	-4.584	0.10	0.02	0.01
	7	-3.577	0.19	0.05	0.02
	8	-1.927	0.23	0.18	0.07
	9	-2.537	0.25	0.12	0.04
	10	-3.871	0.17	0.04	0.01
	11	-1.174	0.16	0.24	0.13
	12	-2.359	0.25	0.13	0.05
	13	-2.547	0.25	0.12	0.04
	14	-1.709	0.21	0.20	0.09
	15	-3.929	0.16	0.04	0.01
	16	-3.307	0.22	0.06	0.02
	17	-2.377	0.25	0.13	0.05
	18	-2.519	0.25	0.12	0.04
	19	-2.441	0.25	0.13	0.05
	20	-1.731	0.21	0.19	0.09
	21	-0.647	0.11	0.25	0.18
	22	0.043	0.06	0.22	0.24
	23	-1.284	0.17	0.23	0.12
	24	-0.364	0.09	0.24	0.21
SP	1	-2.505	0.36	0.39	0.23
	2	-3.198	0.42	0.30	0.14
	3	-2.086	0.34	0.44	0.26
	4	-2.760	0.39	0.36	0.20
	5	-2.883	0.40	0.35	0.18
	6	-2.607	0.63	0.63	0.54
	7	-3.252	0.33	0.32	0.22
	8	-3.572	0.27	0.28	0.25
	9	-3.268	0.37	0.32	0.18
	10	-3.563	0.28	0.29	0.24
	11	-3.211	0.30	0.30	0.25
	12	-2.126	0.53	0.35	0.11
	13	-2.741	0.49	0.32	0.12
	14	-2.844	0.59	0.22	0.07
	15	-2.568	0.47	0.35	0.14
	16	-1.863	0.41	0.46	0.18

Grade K Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	-3.716	0.11	0.03	0.02
	2	-2.956	0.18	0.06	0.03
	3	-3.406	0.14	0.04	0.02
	4	-2.229	0.24	0.11	0.07
	5	-2.492	0.22	0.09	0.05
	6	-2.972	0.18	0.06	0.03
	7	-3.333	0.15	0.04	0.02
	8	-1.894	0.25	0.14	0.09
	9	-2.419	0.23	0.10	0.06
	10	-2.147	0.24	0.12	0.07
	11	-1.555	0.25	0.17	0.11
	12	-2.446	0.23	0.09	0.06
	13	-1.522	0.25	0.18	0.12
	14	-1.516	0.25	0.18	0.12
	15	-2.551	0.22	0.09	0.05
WR	1	-4.850	0.04	0.01	0.01
	2	-3.908	0.10	0.03	0.01
	3	-3.158	0.16	0.05	0.03
	4	-3.428	0.14	0.04	0.02
	5	-3.826	0.10	0.03	0.02
	6	-4.416	0.06	0.02	0.01
	7	-5.751	0.27	0.14	0.09
	8	-5.402	0.27	0.18	0.12
	9	-2.647	0.07	0.01	0.00
	10	-1.211	0.09	0.01	0.00
	11	-2.572	0.49	0.11	0.06
	12	-2.223	0.61	0.17	0.08
	13	-2.007	0.48	0.39	0.26
	14	-2.073	0.45	0.39	0.28
	15	-1.637	0.33	0.40	0.37

Grade 1 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	-3.247	0.20	0.06	0.01
	2	-2.787	0.24	0.08	0.01
	3	-3.696	0.16	0.04	0.01
	4	-3.252	0.20	0.06	0.01
	5	-3.734	0.15	0.04	0.01
	6	-4.584	0.08	0.02	0.00
	7	-3.577	0.17	0.04	0.01
	8	-1.927	0.24	0.15	0.03
	9	-2.537	0.25	0.10	0.02
	10	-3.871	0.14	0.03	0.00
	11	-1.174	0.19	0.22	0.06
	12	-2.359	0.25	0.12	0.02
	13	-2.547	0.25	0.10	0.02
	14	-1.709	0.23	0.18	0.04
	15	-3.929	0.14	0.03	0.00
	16	-3.307	0.20	0.05	0.01
	17	-2.377	0.25	0.11	0.02
	18	-2.519	0.25	0.10	0.02
	19	-2.441	0.25	0.11	0.02
	20	-1.731	0.23	0.17	0.04
	21	-0.647	0.14	0.25	0.09
	22	0.043	0.08	0.23	0.15
	23	-1.284	0.20	0.21	0.05
	24	-0.364	0.11	0.25	0.11
SP	1	-2.505	0.39	0.37	0.11
	2	-3.198	0.43	0.27	0.06
	3	-2.086	0.38	0.41	0.12
	4	-2.760	0.41	0.34	0.09
	5	-2.883	0.41	0.32	0.08
	6	-2.607	0.68	0.61	0.42
	7	-3.252	0.33	0.31	0.11
	8	-3.572	0.26	0.28	0.15
	9	-3.268	0.37	0.30	0.08
	10	-3.563	0.27	0.29	0.14
	11	-3.211	0.30	0.30	0.14
	12	-2.126	0.59	0.29	0.04
	13	-2.741	0.51	0.27	0.05
	14	-2.844	0.58	0.18	0.02
	15	-2.568	0.50	0.30	0.05
	16	-1.863	0.48	0.40	0.07

Grade 1 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	-3.716	0.06	0.02	0.01
	2	-2.956	0.12	0.03	0.01
	3	-3.406	0.08	0.02	0.01
	4	-2.229	0.18	0.07	0.02
	5	-2.492	0.16	0.05	0.02
	6	-2.972	0.11	0.03	0.01
	7	-3.333	0.09	0.02	0.01
	8	-1.894	0.21	0.09	0.03
	9	-2.419	0.17	0.06	0.02
	10	-2.147	0.19	0.07	0.02
	11	-1.555	0.24	0.11	0.04
	12	-2.446	0.16	0.06	0.02
	13	-1.522	0.24	0.12	0.04
	14	-1.516	0.24	0.12	0.04
	15	-2.551	0.15	0.05	0.02
WR	1	-4.850	0.02	0.01	0.00
	2	-3.908	0.05	0.01	0.00
	3	-3.158	0.10	0.03	0.01
	4	-3.428	0.08	0.02	0.01
	5	-3.826	0.06	0.02	0.00
	6	-4.416	0.03	0.01	0.00
	7	-5.751	0.22	0.09	0.03
	8	-5.402	0.25	0.12	0.04
	9	-2.647	0.02	0.00	0.00
	10	-1.211	0.03	0.00	0.00
	11	-2.572	0.26	0.06	0.02
	12	-2.223	0.37	0.08	0.02
	13	-2.007	0.50	0.26	0.10
	14	-2.073	0.47	0.28	0.11
	15	-1.637	0.39	0.37	0.22

Note: The statistics in Tables E.1–E.5 were based on the equating sample. Invalid cases were removed after cleaning procedures.

E.2: Grade 2–Grade 4
Grade 2 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	-2.103	0.23	0.09	0.02
	2	0.347	0.11	0.25	0.16
	3	-0.861	0.22	0.20	0.06
	4	-0.780	0.22	0.21	0.07
	5	0.874	0.07	0.21	0.21
	6	-1.596	0.25	0.13	0.03
	7	-1.909	0.24	0.11	0.02
	8	0.737	0.08	0.23	0.20
	9	1.219	0.06	0.18	0.23
	10	1.765	0.03	0.13	0.25
	11	1.051	0.06	0.20	0.22
	12	0.245	0.12	0.25	0.15
	13	-0.656	0.21	0.22	0.08
	14	1.551	0.04	0.15	0.25
	15	0.609	0.09	0.23	0.18
	16	0.741	0.08	0.23	0.20
	17	0.382	0.11	0.24	0.16
	18	1.056	0.06	0.20	0.22
	19	1.332	0.05	0.17	0.24
	20	-0.810	0.22	0.21	0.07
	21	0.156	0.13	0.25	0.14
	22	-1.433	0.25	0.15	0.04
	23	-0.903	0.23	0.20	0.06
	24	-0.512	0.19	0.23	0.09
SP	1	-1.946	0.35	0.35	0.19
	2	-1.149	0.33	0.44	0.23
	3	-1.493	0.39	0.42	0.17
	4	-0.977	0.32	0.46	0.24
	5	-1.639	0.32	0.34	0.25
	6	-1.428	0.67	0.66	0.51
	7	-2.468	0.31	0.31	0.18
	8	-1.202	0.33	0.42	0.24
	9	-1.633	0.33	0.36	0.23
	10	-1.530	0.33	0.36	0.24
	11	-1.426	0.31	0.36	0.26
	12	-2.225	0.37	0.34	0.16
	13	-0.391	0.31	0.60	0.19
	14	-2.121	0.37	0.35	0.16
	15	-0.484	0.33	0.58	0.18
	16	-1.385	0.43	0.43	0.14

Grade 2 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	-2.544	0.13	0.03	0.01
	2	-1.720	0.20	0.07	0.03
	3	-1.615	0.21	0.08	0.03
	4	-2.633	0.12	0.03	0.01
	5	-2.082	0.17	0.05	0.02
	6	-2.639	0.12	0.03	0.01
	7	-1.862	0.19	0.06	0.02
	8	-1.993	0.18	0.06	0.02
	9	0.206	0.20	0.23	0.13
	10	-0.717	0.25	0.15	0.07
	11	0.874	0.13	0.25	0.20
	12	-1.278	0.24	0.10	0.04
	13	0.449	0.17	0.24	0.16
	14	0.326	0.18	0.24	0.14
	15	1.949	0.06	0.18	0.25
	16	1.263	0.10	0.23	0.23
	17	0.472	0.17	0.25	0.16
	18	1.000	0.12	0.25	0.21
	19	0.126	0.20	0.23	0.13
	20	-0.233	0.23	0.20	0.10
	21	1.370	0.09	0.23	0.23
	22	0.579	0.16	0.25	0.17
	23	-0.652	0.25	0.16	0.07
	24	1.570	0.08	0.21	0.24
WR	1	-2.010	0.18	0.06	0.02
	2	-2.504	0.13	0.04	0.01
	3	-0.268	0.23	0.20	0.09
	4	-1.031	0.25	0.12	0.05
	5	0.720	0.15	0.25	0.18
	6	-0.187	0.23	0.20	0.10
	7	0.339	0.18	0.24	0.15
	8	0.187	0.20	0.23	0.13
	9	0.308	0.19	0.24	0.14
	10	-0.829	0.25	0.14	0.06
	11	0.332	0.18	0.24	0.15
	12	0.438	0.17	0.24	0.16
	13	-2.910	0.35	0.14	0.05
	14	-0.815	0.51	0.39	0.17
	15	-1.090	0.36	0.37	0.27
	16	-1.518	0.46	0.52	0.54

Grade 3 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	-2.103	0.21	0.07	0.01
	2	0.347	0.14	0.25	0.11
	3	-0.861	0.24	0.18	0.04
	4	-0.780	0.24	0.18	0.04
	5	0.874	0.10	0.23	0.16
	6	-1.596	0.24	0.11	0.02
	7	-1.909	0.22	0.09	0.02
	8	0.737	0.11	0.24	0.15
	9	1.219	0.07	0.21	0.20
	10	1.765	0.05	0.16	0.24
	11	1.051	0.08	0.22	0.18
	12	0.245	0.15	0.25	0.11
	13	-0.656	0.23	0.20	0.05
	14	1.551	0.05	0.18	0.22
	15	0.609	0.12	0.25	0.14
	16	0.741	0.11	0.24	0.15
	17	0.382	0.14	0.25	0.12
	18	1.056	0.08	0.22	0.18
	19	1.332	0.07	0.20	0.21
	20	-0.810	0.24	0.18	0.04
	21	0.156	0.16	0.25	0.10
	22	-1.433	0.25	0.12	0.02
	23	-0.903	0.24	0.17	0.04
	24	-0.512	0.22	0.21	0.06
SP	1	-1.946	0.37	0.34	0.13
	2	-1.149	0.38	0.42	0.16
	3	-1.493	0.42	0.38	0.12
	4	-0.977	0.37	0.44	0.17
	5	-1.639	0.33	0.34	0.20
	6	-1.428	0.74	0.62	0.46
	7	-2.468	0.31	0.31	0.13
	8	-1.202	0.36	0.41	0.17
	9	-1.633	0.35	0.35	0.18
	10	-1.530	0.35	0.36	0.19
	11	-1.426	0.33	0.35	0.21
	12	-2.225	0.37	0.32	0.11
	13	-0.391	0.39	0.55	0.12
	14	-2.121	0.37	0.33	0.11
	15	-0.484	0.42	0.53	0.11
	16	-1.385	0.48	0.38	0.09

Grade 3 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	-2.544	0.11	0.03	0.01
	2	-1.720	0.18	0.06	0.02
	3	-1.615	0.19	0.07	0.02
	4	-2.633	0.10	0.03	0.01
	5	-2.082	0.15	0.05	0.01
	6	-2.639	0.10	0.03	0.01
	7	-1.862	0.17	0.06	0.01
	8	-1.993	0.16	0.05	0.01
	9	0.206	0.22	0.22	0.09
	10	-0.717	0.25	0.14	0.04
	11	0.874	0.15	0.25	0.15
	12	-1.278	0.22	0.09	0.03
	13	0.449	0.20	0.24	0.11
	14	0.326	0.21	0.23	0.10
	15	1.949	0.07	0.19	0.24
	16	1.263	0.12	0.24	0.19
	17	0.472	0.19	0.24	0.12
	18	1.000	0.14	0.25	0.16
	19	0.126	0.22	0.22	0.09
	20	-0.233	0.24	0.19	0.07
	21	1.370	0.11	0.24	0.20
	22	0.579	0.18	0.24	0.12
	23	-0.652	0.25	0.15	0.05
	24	1.570	0.09	0.22	0.22
WR	1	-2.010	0.15	0.05	0.01
	2	-2.504	0.11	0.03	0.01
	3	-0.268	0.24	0.18	0.06
	4	-1.031	0.24	0.11	0.03
	5	0.720	0.17	0.25	0.14
	6	-0.187	0.24	0.19	0.07
	7	0.339	0.20	0.23	0.10
	8	0.187	0.22	0.22	0.09
	9	0.308	0.21	0.23	0.10
	10	-0.829	0.25	0.13	0.04
	11	0.332	0.21	0.23	0.10
	12	0.438	0.20	0.24	0.11
	13	-2.910	0.32	0.13	0.04
	14	-0.815	0.54	0.36	0.11
	15	-1.090	0.37	0.36	0.21
	16	1.19	0.48	0.52	0.55

Grade 4 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	-2.103	0.18	0.06	0.01
	2	0.347	0.17	0.25	0.10
	3	-0.861	0.25	0.15	0.03
	4	-0.780	0.25	0.16	0.04
	5	0.874	0.12	0.25	0.14
	6	-1.596	0.23	0.09	0.02
	7	-1.909	0.20	0.07	0.01
	8	0.737	0.13	0.25	0.13
	9	1.219	0.09	0.23	0.18
	10	1.765	0.06	0.18	0.22
	11	1.051	0.11	0.24	0.16
	12	0.245	0.18	0.24	0.09
	13	-0.656	0.24	0.17	0.04
	14	1.551	0.07	0.20	0.21
	15	0.609	0.14	0.25	0.12
	16	0.741	0.13	0.25	0.13
	17	0.382	0.17	0.25	0.10
	18	1.056	0.10	0.24	0.16
	19	1.332	0.08	0.22	0.19
	20	-0.810	0.25	0.16	0.04
	21	0.156	0.19	0.24	0.08
	22	-1.433	0.24	0.10	0.02
	23	-0.903	0.25	0.15	0.03
	24	-0.512	0.24	0.18	0.05
SP	1	-1.946	0.37	0.32	0.12
	2	-1.149	0.41	0.39	0.14
	3	-1.493	0.44	0.34	0.10
	4	-0.977	0.41	0.41	0.15
	5	-1.639	0.33	0.33	0.18
	6	-1.428	0.77	0.60	0.43
	7	-2.468	0.31	0.29	0.11
	8	-1.202	0.39	0.39	0.15
	9	-1.633	0.36	0.34	0.15
	10	-1.530	0.36	0.35	0.16
	11	-1.426	0.35	0.35	0.18
	12	-2.225	0.37	0.29	0.09
	13	-0.391	0.48	0.48	0.10
	14	-2.121	0.38	0.30	0.10
	15	-0.484	0.50	0.46	0.09
	16	-1.385	0.50	0.33	0.08

Grade 4 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	-2.544	0.08	0.02	0.01
	2	-1.720	0.15	0.04	0.01
	3	-1.615	0.16	0.05	0.01
	4	-2.633	0.08	0.02	0.01
	5	-2.082	0.12	0.03	0.01
	6	-2.639	0.08	0.02	0.01
	7	-1.862	0.14	0.04	0.01
	8	-1.993	0.12	0.03	0.01
	9	0.206	0.24	0.19	0.07
	10	-0.717	0.24	0.10	0.03
	11	0.874	0.19	0.24	0.12
	12	-1.278	0.19	0.06	0.02
	13	0.449	0.22	0.21	0.09
	14	0.326	0.23	0.20	0.08
	15	1.949	0.09	0.23	0.22
	16	1.263	0.15	0.25	0.16
	17	0.472	0.22	0.21	0.09
	18	1.000	0.18	0.24	0.14
	19	0.126	0.24	0.18	0.07
	20	-0.233	0.25	0.14	0.05
	21	1.370	0.14	0.25	0.17
	22	0.579	0.21	0.22	0.10
	23	-0.652	0.24	0.11	0.03
	24	1.570	0.12	0.25	0.19
WR	1	-2.010	0.12	0.03	0.01
	2	-2.504	0.08	0.02	0.01
	3	-0.268	0.25	0.14	0.05
	4	-1.031	0.21	0.08	0.02
	5	0.720	0.20	0.23	0.11
	6	-0.187	0.25	0.15	0.05
	7	0.339	0.23	0.20	0.08
	8	0.187	0.24	0.18	0.07
	9	0.308	0.23	0.20	0.08
	10	-0.829	0.23	0.09	0.03
	11	0.332	0.23	0.20	0.08
	12	0.438	0.22	0.21	0.09
	13	-2.910	0.27	0.09	0.03
	14	-0.815	0.55	0.26	0.08
	15	-1.090	0.38	0.33	0.17
	16	-1.518	0.49	0.53	0.55

E.3: Grade 5–Grade 6
Grade 5 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	-1.098	0.23	0.10	0.02
	2	1.266	0.12	0.24	0.17
	3	1.290	0.12	0.24	0.17
	4	-0.458	0.25	0.16	0.04
	5	-0.131	0.24	0.19	0.06
	6	0.668	0.18	0.25	0.11
	7	1.219	0.12	0.24	0.16
	8	-0.479	0.25	0.16	0.04
	9	0.097	0.22	0.21	0.07
	10	0.576	0.18	0.24	0.11
	11	0.862	0.16	0.25	0.13
	12	1.022	0.14	0.25	0.15
	13	1.503	0.10	0.23	0.19
	14	-0.692	0.25	0.14	0.04
	15	1.836	0.08	0.20	0.22
	16	0.234	0.21	0.22	0.08
	17	2.318	0.05	0.16	0.25
	18	2.221	0.06	0.17	0.24
	19	0.369	0.20	0.23	0.09
	20	0.370	0.20	0.23	0.09
	21	1.500	0.10	0.23	0.19
	22	0.692	0.17	0.25	0.12
	23	2.790	0.03	0.12	0.25
	24	0.751	0.17	0.25	0.12
	25	2.303	0.05	0.16	0.25
SP	1	-1.116	0.39	0.36	0.17
	2	-0.550	0.44	0.43	0.15
	3	-0.202	0.39	0.48	0.20
	4	-0.614	0.38	0.41	0.21
	5	-0.279	0.34	0.43	0.25
	6	-0.619	0.69	0.71	0.55
	7	-1.170	0.40	0.36	0.15
	8	-0.453	0.39	0.44	0.19
	9	-0.269	0.37	0.46	0.22
	10	-0.983	0.38	0.37	0.18
	11	-1.504	0.40	0.33	0.12
	12	0.596	0.49	0.58	0.10
	13	-0.117	0.37	0.49	0.21
	14	-0.499	0.61	0.35	0.07
	15	0.258	0.46	0.56	0.13
	16	0.602	0.32	0.62	0.21

Grade 5 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	-1.581	0.09	0.03	0.01
	2	-0.341	0.20	0.10	0.03
	3	0.232	0.24	0.15	0.05
	4	-1.330	0.11	0.04	0.01
	5	1.643	0.19	0.25	0.16
	6	0.533	0.25	0.18	0.07
	7	0.163	0.24	0.14	0.05
	8	0.276	0.24	0.15	0.05
	9	2.831	0.09	0.19	0.25
	10	0.269	0.24	0.15	0.05
	11	1.264	0.22	0.24	0.12
	12	0.751	0.25	0.20	0.08
	13	2.446	0.12	0.22	0.23
	14	0.294	0.24	0.15	0.05
	15	1.689	0.19	0.25	0.16
	16	1.166	0.23	0.23	0.11
	17	1.131	0.23	0.23	0.11
	18	0.522	0.25	0.17	0.07
	19	0.533	0.25	0.18	0.07
	20	1.306	0.22	0.24	0.12
	21	2.065	0.15	0.24	0.20
	22	1.632	0.19	0.25	0.15
	23	0.933	0.24	0.21	0.09
	24	1.638	0.19	0.25	0.16
	25	0.660	0.25	0.19	0.08
	26	1.229	0.23	0.23	0.12
	27	1.310	0.22	0.24	0.12
WR	1	0.535	0.25	0.18	0.07
	2	-0.520	0.19	0.08	0.03
	3	-0.083	0.22	0.12	0.04
	4	0.240	0.24	0.15	0.05
	5	1.941	0.16	0.25	0.18
	6	0.015	0.23	0.13	0.04
	7	2.222	0.14	0.24	0.21
	8	-0.748	0.17	0.07	0.02
	9	0.227	0.24	0.15	0.05
	10	0.298	0.24	0.15	0.06
	11	2.567	0.11	0.21	0.23
	12	-0.614	0.18	0.08	0.02
	13	1.332	0.22	0.24	0.13
	14	2.005	0.16	0.25	0.19
	15	2.117	0.15	0.24	0.20
	16	-1.310	0.36	0.16	0.05
	17	-0.226	0.53	0.29	0.09
	18	0.296	0.41	0.42	0.25
	19	-0.443	0.58	0.56	0.55

Grade 6 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	-1.098	0.21	0.08	0.02
	2	1.266	0.14	0.25	0.15
	3	1.290	0.14	0.25	0.15
	4	-0.458	0.25	0.14	0.04
	5	-0.131	0.25	0.17	0.05
	6	0.668	0.20	0.23	0.10
	7	1.219	0.15	0.25	0.15
	8	-0.479	0.25	0.13	0.04
	9	0.097	0.24	0.19	0.06
	10	0.576	0.21	0.23	0.09
	11	0.862	0.18	0.24	0.11
	12	1.022	0.17	0.25	0.13
	13	1.503	0.12	0.24	0.17
	14	-0.692	0.24	0.12	0.03
	15	1.836	0.09	0.23	0.20
	16	0.234	0.23	0.20	0.07
	17	2.318	0.06	0.18	0.24
	18	2.221	0.07	0.19	0.23
	19	0.369	0.22	0.21	0.08
	20	0.370	0.22	0.21	0.08
	21	1.500	0.12	0.24	0.17
	22	0.692	0.20	0.24	0.10
	23	2.790	0.04	0.14	0.25
	24	0.751	0.19	0.24	0.10
	25	2.303	0.06	0.19	0.24
SP	1	-1.116	0.40	0.34	0.14
	2	-0.550	0.47	0.39	0.13
	3	-0.202	0.43	0.44	0.17
	4	-0.614	0.40	0.39	0.18
	5	-0.279	0.37	0.41	0.22
	6	-0.619	0.74	0.69	0.52
	7	-1.170	0.41	0.33	0.13
	8	-0.453	0.43	0.41	0.17
	9	-0.269	0.40	0.43	0.19
	10	-0.983	0.40	0.35	0.16
	11	-1.504	0.40	0.29	0.11
	12	0.596	0.60	0.47	0.08
	13	-0.117	0.42	0.46	0.18
	14	-0.499	0.63	0.28	0.06
	15	0.258	0.54	0.47	0.10
	16	0.602	0.40	0.58	0.17

Grade 6 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	-1.581	0.09	0.03	0.01
	2	-0.341	0.19	0.08	0.03
	3	0.232	0.24	0.13	0.04
	4	-1.330	0.10	0.03	0.01
	5	1.643	0.20	0.24	0.14
	6	0.533	0.25	0.16	0.06
	7	0.163	0.23	0.12	0.04
	8	0.276	0.24	0.13	0.05
	9	2.831	0.09	0.21	0.24
	10	0.269	0.24	0.13	0.05
	11	1.264	0.23	0.22	0.11
	12	0.751	0.25	0.18	0.07
	13	2.446	0.13	0.24	0.21
	14	0.294	0.24	0.13	0.05
	15	1.689	0.20	0.25	0.14
	16	1.166	0.24	0.21	0.10
	17	1.131	0.24	0.21	0.10
	18	0.522	0.25	0.15	0.06
	19	0.533	0.25	0.16	0.06
	20	1.306	0.23	0.22	0.11
	21	2.065	0.16	0.25	0.18
	22	1.632	0.20	0.24	0.14
	23	0.933	0.25	0.19	0.08
	24	1.638	0.20	0.24	0.14
	25	0.660	0.25	0.17	0.07
	26	1.229	0.23	0.22	0.10
	27	1.310	0.23	0.23	0.11
WR	1	0.535	0.25	0.16	0.06
	2	-0.520	0.18	0.07	0.02
	3	-0.083	0.22	0.10	0.03
	4	0.240	0.24	0.13	0.05
	5	1.941	0.17	0.25	0.17
	6	0.015	0.22	0.11	0.04
	7	2.222	0.15	0.25	0.19
	8	-0.748	0.16	0.06	0.02
	9	0.227	0.24	0.13	0.04
	10	0.298	0.24	0.13	0.05
	11	2.567	0.11	0.23	0.22
	12	-0.614	0.17	0.07	0.02
	13	1.332	0.23	0.23	0.11
	14	2.005	0.17	0.25	0.17
	15	2.117	0.16	0.25	0.19
	16	-1.310	0.34	0.13	0.04
	17	-0.226	0.52	0.25	0.08
	18	0.296	0.42	0.40	0.23
	19	-0.443	0.59	0.56	0.54

E.4: Grade 7–Grade 8
Grade 7 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	-1.003	0.20	0.07	0.02
	2	-0.439	0.24	0.11	0.03
	3	0.804	0.21	0.22	0.09
	4	1.136	0.18	0.24	0.11
	5	2.777	0.05	0.17	0.25
	6	2.943	0.04	0.16	0.25
	7	3.180	0.04	0.13	0.25
	8	1.822	0.11	0.24	0.17
	9	0.882	0.20	0.23	0.09
	10	2.587	0.06	0.19	0.24
	11	2.825	0.05	0.17	0.25
	12	2.772	0.05	0.17	0.25
	13	3.771	0.02	0.09	0.22
	14	2.416	0.07	0.21	0.23
	15	1.694	0.12	0.25	0.16
	16	1.558	0.14	0.25	0.15
	17	0.982	0.19	0.23	0.10
	18	2.260	0.08	0.22	0.21
	19	2.145	0.09	0.23	0.20
	20	1.788	0.12	0.25	0.17
	21	1.199	0.17	0.24	0.12
	22	0.138	0.25	0.16	0.05
	23	2.211	0.08	0.22	0.21
	24	3.675	0.02	0.09	0.23
	25	1.583	0.13	0.25	0.15
SP	1	-0.243	0.38	0.39	0.20
	2	0.690	0.30	0.50	0.27
	3	0.649	0.32	0.50	0.25
	4	-0.403	0.35	0.36	0.22
	5	0.877	0.28	0.52	0.28
	6	-0.042	0.68	0.79	0.56
	7	-0.535	0.35	0.35	0.21
	8	-0.241	0.39	0.39	0.19
	9	-0.254	0.38	0.39	0.20
	10	-0.295	0.37	0.38	0.20
	11	-0.888	0.38	0.33	0.15
	12	0.532	0.34	0.49	0.24
	13	1.016	0.27	0.55	0.28
	14	0.781	0.34	0.55	0.21
	15	1.890	0.17	0.63	0.36
	16	0.947	0.31	0.57	0.23

Grade 7 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	-0.203	0.19	0.07	0.03
	2	-0.846	0.13	0.04	0.02
	3	-1.212	0.10	0.03	0.01
	4	1.440	0.23	0.22	0.13
	5	0.675	0.25	0.15	0.07
	6	-0.392	0.17	0.06	0.03
	7	1.398	0.23	0.21	0.13
	8	1.003	0.25	0.18	0.09
	9	1.042	0.25	0.18	0.10
	10	1.131	0.25	0.19	0.10
	11	1.192	0.24	0.19	0.11
	12	0.170	0.22	0.10	0.05
	13	0.448	0.24	0.12	0.06
	14	2.097	0.18	0.25	0.19
	15	0.976	0.25	0.17	0.09
	16	1.365	0.24	0.21	0.12
	17	1.380	0.24	0.21	0.12
	18	1.653	0.22	0.23	0.15
	19	1.764	0.21	0.24	0.16
	20	1.846	0.20	0.24	0.17
	21	2.350	0.15	0.25	0.21
	22	1.623	0.22	0.23	0.15
	23	2.483	0.14	0.25	0.22
	24	0.902	0.25	0.17	0.09
	25	1.810	0.20	0.24	0.16
	26	1.043	0.25	0.18	0.10
	27	1.839	0.20	0.24	0.17
WR	1	-0.518	0.16	0.06	0.02
	2	-0.507	0.16	0.06	0.02
	3	2.378	0.15	0.25	0.22
	4	1.474	0.23	0.22	0.13
	5	0.057	0.21	0.09	0.04
	6	-0.009	0.21	0.09	0.04
	7	1.072	0.25	0.18	0.10
	8	0.221	0.22	0.11	0.05
	9	1.805	0.20	0.24	0.16
	10	1.015	0.25	0.18	0.09
	11	1.320	0.24	0.21	0.12
	12	-0.031	0.20	0.09	0.04
	13	1.168	0.25	0.19	0.11
	14	1.094	0.25	0.19	0.10
	15	1.474	0.23	0.22	0.13
	16	0.544	0.57	0.34	0.16
	17	-0.877	0.32	0.10	0.04
	18	0.095	0.51	0.30	0.15
	19	-0.471	0.63	0.61	0.58

Grade 8 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	-1.003	0.18	0.05	0.02
	2	-0.439	0.23	0.09	0.03
	3	0.804	0.23	0.20	0.09
	4	1.136	0.20	0.22	0.11
	5	2.777	0.07	0.20	0.25
	6	2.943	0.06	0.18	0.25
	7	3.180	0.05	0.16	0.25
	8	1.822	0.14	0.25	0.17
	9	0.882	0.22	0.20	0.09
	10	2.587	0.08	0.21	0.24
	11	2.825	0.06	0.19	0.25
	12	2.772	0.07	0.20	0.25
	13	3.771	0.03	0.11	0.22
	14	2.416	0.09	0.23	0.23
	15	1.694	0.15	0.25	0.16
	16	1.558	0.16	0.25	0.15
	17	0.982	0.21	0.21	0.10
	18	2.260	0.10	0.24	0.21
	19	2.145	0.11	0.24	0.20
	20	1.788	0.14	0.25	0.17
	21	1.199	0.20	0.23	0.12
	22	0.138	0.25	0.13	0.05
	23	2.211	0.10	0.24	0.21
	24	3.675	0.03	0.11	0.23
	25	1.583	0.16	0.25	0.15
SP	1	-0.243	0.40	0.36	0.20
	2	0.690	0.36	0.48	0.27
	3	0.649	0.38	0.48	0.25
	4	-0.403	0.36	0.35	0.22
	5	0.877	0.34	0.51	0.28
	6	-0.042	0.76	0.76	0.56
	7	-0.535	0.36	0.34	0.21
	8	-0.241	0.41	0.36	0.19
	9	-0.254	0.40	0.36	0.20
	10	-0.295	0.39	0.36	0.20
	11	-0.888	0.39	0.30	0.15
	12	0.532	0.39	0.46	0.24
	13	1.016	0.33	0.54	0.28
	14	0.781	0.41	0.50	0.21
	15	1.890	0.22	0.67	0.36
	16	0.947	0.38	0.53	0.23

Grade 8 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	-0.203	0.18	0.07	0.03
	2	-0.846	0.12	0.04	0.02
	3	-1.212	0.09	0.03	0.01
	4	1.440	0.24	0.21	0.13
	5	0.675	0.24	0.13	0.07
	6	-0.392	0.16	0.06	0.03
	7	1.398	0.24	0.20	0.13
	8	1.003	0.25	0.16	0.09
	9	1.042	0.25	0.17	0.10
	10	1.131	0.25	0.18	0.10
	11	1.192	0.25	0.18	0.11
	12	0.170	0.21	0.09	0.05
	13	0.448	0.23	0.11	0.06
	14	2.097	0.19	0.25	0.19
	15	0.976	0.25	0.16	0.09
	16	1.365	0.24	0.20	0.12
	17	1.380	0.24	0.20	0.12
	18	1.653	0.22	0.22	0.15
	19	1.764	0.22	0.23	0.16
	20	1.846	0.21	0.24	0.17
	21	2.350	0.16	0.25	0.21
	22	1.623	0.23	0.22	0.15
	23	2.483	0.15	0.25	0.22
	24	0.902	0.25	0.15	0.09
	25	1.810	0.21	0.23	0.16
	26	1.043	0.25	0.17	0.10
	27	1.839	0.21	0.23	0.17
WR	1	-0.518	0.15	0.05	0.02
	2	-0.507	0.15	0.05	0.02
	3	2.378	0.16	0.25	0.22
	4	1.474	0.24	0.21	0.13
	5	0.057	0.20	0.08	0.04
	6	-0.009	0.20	0.08	0.04
	7	1.072	0.25	0.17	0.10
	8	0.221	0.22	0.10	0.05
	9	1.805	0.21	0.23	0.16
	10	1.015	0.25	0.17	0.09
	11	1.320	0.24	0.19	0.12
	12	-0.031	0.20	0.08	0.04
	13	1.168	0.25	0.18	0.11
	14	1.094	0.25	0.17	0.10
	15	1.474	0.24	0.21	0.13
	16	0.544	0.57	0.31	0.16
	17	-0.877	0.30	0.09	0.04
	18	0.095	0.50	0.28	0.15
	19	-0.471	0.63	0.61	0.58

E.5: Grade 9–Grade 12
Grade 9 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	0.291	0.24	0.11	0.04
	2	0.967	0.25	0.17	0.08
	3	0.412	0.24	0.12	0.05
	4	1.264	0.23	0.20	0.10
	5	2.051	0.16	0.25	0.17
	6	0.762	0.25	0.15	0.07
	7	0.925	0.25	0.17	0.08
	8	1.580	0.21	0.23	0.13
	9	1.057	0.24	0.18	0.09
	10	2.749	0.10	0.23	0.23
	11	2.923	0.09	0.22	0.24
	12	1.726	0.20	0.23	0.14
	13	2.895	0.09	0.22	0.24
	14	2.487	0.12	0.25	0.21
	15	2.886	0.09	0.22	0.24
	16	2.493	0.12	0.25	0.21
	17	2.328	0.14	0.25	0.20
	18	2.300	0.14	0.25	0.20
	19	1.722	0.20	0.23	0.14
	20	2.149	0.15	0.25	0.18
	21	2.358	0.14	0.25	0.20
	22	0.828	0.25	0.16	0.07
	23	0.242	0.24	0.11	0.04
	24	2.349	0.14	0.25	0.20
	25	2.060	0.16	0.25	0.17
SP	1	-0.739	0.30	0.30	0.21
	2	-0.821	0.32	0.29	0.19
	3	0.086	0.52	0.29	0.12
	4	0.109	0.41	0.35	0.21
	5	0.320	0.51	0.34	0.15
	6	-0.251	0.80	0.65	0.47
	7	-0.807	0.35	0.28	0.16
	8	-0.242	0.37	0.33	0.21
	9	-1.001	0.33	0.28	0.16
	10	-1.010	0.33	0.28	0.16
	11	0.000	0.40	0.35	0.21
	12	0.190	0.45	0.35	0.18
	13	0.199	0.45	0.35	0.18
	14	0.435	0.49	0.37	0.17
	15	0.879	0.45	0.45	0.23
	16	-0.264	0.45	0.29	0.14

Grade 9 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	0.361	0.20	0.07	0.04
	2	0.517	0.22	0.08	0.04
	3	0.616	0.22	0.09	0.05
	4	-0.348	0.14	0.04	0.02
	5	2.099	0.21	0.22	0.15
	6	2.736	0.15	0.25	0.21
	7	0.576	0.22	0.08	0.04
	8	2.697	0.16	0.25	0.20
	9	1.770	0.24	0.19	0.12
	10	3.073	0.12	0.25	0.23
	11	2.973	0.13	0.25	0.23
	12	1.954	0.22	0.20	0.14
	13	1.177	0.25	0.13	0.07
	14	1.546	0.25	0.17	0.10
	15	1.619	0.24	0.17	0.11
	16	1.814	0.23	0.19	0.12
	17	2.164	0.21	0.22	0.15
	18	2.022	0.22	0.21	0.14
	19	2.807	0.15	0.25	0.21
	20	2.316	0.19	0.23	0.17
	21	2.694	0.16	0.25	0.20
	22	2.665	0.16	0.25	0.20
	23	3.562	0.08	0.22	0.25
	24	3.102	0.12	0.25	0.23
	25	2.850	0.14	0.25	0.22
	26	1.962	0.22	0.20	0.14
	27	3.421	0.09	0.23	0.25
WR	1	0.227	0.19	0.06	0.03
	2	-0.656	0.11	0.03	0.01
	3	1.749	0.24	0.19	0.12
	4	0.476	0.21	0.08	0.04
	5	2.174	0.21	0.22	0.16
	6	2.782	0.15	0.25	0.21
	7	0.240	0.19	0.06	0.03
	8	0.951	0.24	0.11	0.06
	9	2.185	0.21	0.22	0.16
	10	0.966	0.24	0.11	0.06
	11	0.669	0.23	0.09	0.05
	12	2.983	0.13	0.25	0.23
	13	2.737	0.15	0.25	0.21
	14	2.118	0.21	0.22	0.15
	15	1.818	0.23	0.19	0.12
	16	0.521	0.55	0.16	0.08
	17	0.699	0.58	0.21	0.10
	18	0.710	0.46	0.33	0.21
	19	0.525	0.77	0.73	0.61

Grade 10 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	0.291	0.23	0.09	0.04
	2	0.967	0.25	0.14	0.07
	3	0.412	0.24	0.10	0.04
	4	1.264	0.24	0.17	0.09
	5	2.051	0.17	0.24	0.16
	6	0.762	0.25	0.12	0.06
	7	0.925	0.25	0.14	0.07
	8	1.580	0.21	0.20	0.11
	9	1.057	0.25	0.15	0.07
	10	2.749	0.11	0.25	0.22
	11	2.923	0.09	0.24	0.23
	12	1.726	0.20	0.21	0.13
	13	2.895	0.10	0.24	0.23
	14	2.487	0.13	0.25	0.20
	15	2.886	0.10	0.24	0.23
	16	2.493	0.13	0.25	0.20
	17	2.328	0.15	0.25	0.18
	18	2.300	0.15	0.25	0.18
	19	1.722	0.20	0.21	0.13
	20	2.149	0.16	0.24	0.17
	21	2.358	0.14	0.25	0.19
	22	0.828	0.25	0.13	0.06
	23	0.242	0.23	0.08	0.04
	24	2.349	0.14	0.25	0.18
	25	2.060	0.17	0.24	0.16
SP	1	-0.739	0.30	0.28	0.19
	2	-0.821	0.32	0.27	0.17
	3	0.086	0.52	0.23	0.10
	4	0.109	0.42	0.32	0.18
	5	0.320	0.52	0.28	0.13
	6	-0.251	0.80	0.60	0.43
	7	-0.807	0.35	0.25	0.14
	8	-0.242	0.37	0.30	0.19
	9	-1.001	0.33	0.25	0.14
	10	-1.010	0.33	0.25	0.14
	11	0.000	0.40	0.31	0.19
	12	0.190	0.46	0.30	0.16
	13	0.199	0.46	0.30	0.16
	14	0.435	0.50	0.31	0.15
	15	0.879	0.46	0.39	0.20
	16	-0.264	0.45	0.24	0.12

Grade 10 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	0.361	0.20	0.06	0.03
	2	0.517	0.21	0.07	0.04
	3	0.616	0.22	0.08	0.04
	4	-0.348	0.13	0.03	0.02
	5	2.099	0.22	0.21	0.13
	6	2.736	0.16	0.25	0.19
	7	0.576	0.21	0.08	0.04
	8	2.697	0.17	0.25	0.19
	9	1.770	0.24	0.18	0.11
	10	3.073	0.13	0.25	0.22
	11	2.973	0.14	0.25	0.22
	12	1.954	0.23	0.19	0.12
	13	1.177	0.25	0.12	0.07
	14	1.546	0.25	0.16	0.09
	15	1.619	0.25	0.16	0.09
	16	1.814	0.24	0.18	0.11
	17	2.164	0.21	0.21	0.14
	18	2.022	0.23	0.20	0.13
	19	2.807	0.16	0.25	0.20
	20	2.316	0.20	0.22	0.15
	21	2.694	0.17	0.25	0.19
	22	2.665	0.17	0.24	0.19
	23	3.562	0.09	0.23	0.25
	24	3.102	0.13	0.25	0.22
	25	2.850	0.15	0.25	0.20
	26	1.962	0.23	0.20	0.12
	27	3.421	0.10	0.24	0.24
WR	1	0.227	0.18	0.06	0.03
	2	-0.656	0.10	0.03	0.01
	3	1.749	0.24	0.18	0.10
	4	0.476	0.21	0.07	0.03
	5	2.174	0.21	0.21	0.14
	6	2.782	0.16	0.25	0.20
	7	0.240	0.18	0.06	0.03
	8	0.951	0.24	0.10	0.05
	9	2.185	0.21	0.21	0.14
	10	0.966	0.24	0.10	0.05
	11	0.669	0.22	0.08	0.04
	12	2.983	0.14	0.25	0.22
	13	2.737	0.16	0.25	0.19
	14	2.118	0.22	0.21	0.14
	15	1.818	0.24	0.18	0.11
	16	0.521	0.53	0.14	0.06
	17	0.699	0.56	0.19	0.09
	18	0.710	0.47	0.31	0.19
	19	0.525	0.78	0.72	0.57

Grade 11 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	0.291	0.22	0.07	0.03
	2	0.967	0.25	0.12	0.06
	3	0.412	0.23	0.08	0.04
	4	1.264	0.25	0.15	0.08
	5	2.051	0.20	0.22	0.14
	6	0.762	0.25	0.10	0.05
	7	0.925	0.25	0.12	0.06
	8	1.580	0.23	0.18	0.10
	9	1.057	0.25	0.13	0.06
	10	2.749	0.13	0.25	0.20
	11	2.923	0.11	0.25	0.22
	12	1.726	0.22	0.19	0.11
	13	2.895	0.12	0.25	0.22
	14	2.487	0.15	0.25	0.18
	15	2.886	0.12	0.25	0.22
	16	2.493	0.15	0.25	0.18
	17	2.328	0.17	0.24	0.16
	18	2.300	0.17	0.24	0.16
	19	1.722	0.22	0.19	0.11
	20	2.149	0.19	0.23	0.15
	21	2.358	0.17	0.24	0.17
	22	0.828	0.25	0.11	0.05
	23	0.242	0.21	0.07	0.03
	24	2.349	0.17	0.24	0.17
	25	2.060	0.19	0.22	0.14
SP	1	-0.739	0.31	0.26	0.17
	2	-0.821	0.32	0.25	0.15
	3	0.086	0.51	0.19	0.09
	4	0.109	0.42	0.28	0.16
	5	0.320	0.52	0.23	0.11
	6	-0.251	0.78	0.57	0.39
	7	-0.807	0.34	0.23	0.13
	8	-0.242	0.37	0.28	0.17
	9	-1.001	0.33	0.22	0.12
	10	-1.010	0.33	0.22	0.12
	11	0.000	0.40	0.28	0.16
	12	0.190	0.46	0.26	0.14
	13	0.199	0.46	0.27	0.14
	14	0.435	0.51	0.26	0.13
	15	0.879	0.50	0.34	0.17
	16	-0.264	0.44	0.20	0.10

Grade 11 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	0.361	0.19	0.05	0.02
	2	0.517	0.20	0.06	0.03
	3	0.616	0.21	0.06	0.03
	4	-0.348	0.12	0.03	0.01
	5	2.099	0.23	0.19	0.10
	6	2.736	0.17	0.24	0.16
	7	0.576	0.21	0.06	0.03
	8	2.697	0.18	0.23	0.16
	9	1.770	0.24	0.16	0.08
	10	3.073	0.14	0.25	0.19
	11	2.973	0.15	0.25	0.18
	12	1.954	0.24	0.17	0.09
	13	1.177	0.24	0.10	0.05
	14	1.546	0.25	0.13	0.07
	15	1.619	0.25	0.14	0.07
	16	1.814	0.24	0.16	0.08
	17	2.164	0.22	0.19	0.11
	18	2.022	0.23	0.18	0.10
	19	2.807	0.16	0.24	0.17
	20	2.316	0.21	0.21	0.12
	21	2.694	0.18	0.23	0.16
	22	2.665	0.18	0.23	0.15
	23	3.562	0.10	0.24	0.23
	24	3.102	0.14	0.25	0.20
	25	2.850	0.16	0.24	0.17
	26	1.962	0.24	0.17	0.09
	27	3.421	0.11	0.25	0.22
WR	1	0.227	0.17	0.05	0.02
	2	-0.656	0.10	0.02	0.01
	3	1.749	0.25	0.15	0.08
	4	0.476	0.20	0.06	0.02
	5	2.174	0.22	0.19	0.11
	6	2.782	0.17	0.24	0.17
	7	0.240	0.18	0.05	0.02
	8	0.951	0.23	0.09	0.04
	9	2.185	0.22	0.19	0.11
	10	0.966	0.23	0.09	0.04
	11	0.669	0.21	0.07	0.03
	12	2.983	0.15	0.25	0.18
	13	2.737	0.17	0.24	0.16
	14	2.118	0.23	0.19	0.11
	15	1.818	0.24	0.16	0.08
	16	0.521	0.50	0.11	0.05
	17	0.699	0.55	0.15	0.06
	18	0.710	0.47	0.27	0.14
	19	0.525	0.79	0.68	0.49

Grade 12 Listening and Speaking

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LI	1	0.291	0.20	0.06	0.03
	2	0.967	0.25	0.10	0.06
	3	0.412	0.21	0.06	0.04
	4	1.264	0.25	0.13	0.08
	5	2.051	0.21	0.20	0.14
	6	0.762	0.24	0.09	0.05
	7	0.925	0.25	0.10	0.06
	8	1.580	0.24	0.16	0.10
	9	1.057	0.25	0.11	0.06
	10	2.749	0.14	0.25	0.20
	11	2.923	0.13	0.25	0.22
	12	1.726	0.23	0.17	0.11
	13	2.895	0.13	0.25	0.22
	14	2.487	0.17	0.23	0.18
	15	2.886	0.13	0.25	0.22
	16	2.493	0.17	0.23	0.18
	17	2.328	0.18	0.22	0.16
	18	2.300	0.19	0.22	0.16
	19	1.722	0.23	0.17	0.11
	20	2.149	0.20	0.21	0.15
	21	2.358	0.18	0.22	0.17
	22	0.828	0.24	0.09	0.05
	23	0.242	0.20	0.06	0.03
	24	2.349	0.18	0.22	0.17
	25	2.060	0.21	0.20	0.14
SP	1	-0.739	0.31	0.24	0.17
	2	-0.821	0.32	0.22	0.15
	3	0.086	0.49	0.15	0.09
	4	0.109	0.42	0.25	0.16
	5	0.320	0.51	0.19	0.11
	6	-0.251	0.77	0.52	0.39
	7	-0.807	0.34	0.20	0.13
	8	-0.242	0.37	0.25	0.17
	9	-1.001	0.32	0.19	0.12
	10	-1.010	0.32	0.19	0.12
	11	0.000	0.40	0.25	0.16
	12	0.190	0.46	0.22	0.14
	13	0.199	0.46	0.22	0.14
	14	0.435	0.51	0.22	0.13
	15	0.879	0.51	0.29	0.17
	16	-0.264	0.43	0.17	0.10

Grade 12 Reading and Writing

	Item	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
RD	1	0.361	0.18	0.05	0.02
	2	0.517	0.19	0.05	0.03
	3	0.616	0.20	0.06	0.03
	4	-0.348	0.11	0.02	0.01
	5	2.099	0.23	0.17	0.10
	6	2.736	0.18	0.23	0.16
	7	0.576	0.20	0.06	0.03
	8	2.697	0.18	0.23	0.16
	9	1.770	0.25	0.14	0.08
	10	3.073	0.15	0.25	0.19
	11	2.973	0.16	0.24	0.18
	12	1.954	0.24	0.16	0.09
	13	1.177	0.24	0.09	0.05
	14	1.546	0.25	0.12	0.07
	15	1.619	0.25	0.13	0.07
	16	1.814	0.25	0.15	0.08
	17	2.164	0.23	0.18	0.11
	18	2.022	0.24	0.17	0.10
	19	2.807	0.17	0.23	0.17
	20	2.316	0.22	0.20	0.12
	21	2.694	0.18	0.23	0.16
	22	2.665	0.19	0.22	0.15
	23	3.562	0.10	0.25	0.23
	24	3.102	0.15	0.25	0.20
	25	2.850	0.17	0.24	0.17
	26	1.962	0.24	0.16	0.09
	27	3.421	0.12	0.25	0.22
WR	1	0.227	0.17	0.04	0.02
	2	-0.656	0.09	0.02	0.01
	3	1.749	0.25	0.14	0.08
	4	0.476	0.19	0.05	0.02
	5	2.174	0.23	0.18	0.11
	6	2.782	0.18	0.23	0.17
	7	0.240	0.17	0.04	0.02
	8	0.951	0.23	0.08	0.04
	9	2.185	0.23	0.18	0.11
	10	0.966	0.23	0.08	0.04
	11	0.669	0.21	0.06	0.03
	12	2.983	0.16	0.24	0.18
	13	2.737	0.18	0.23	0.16
	14	2.118	0.23	0.18	0.11
	15	1.818	0.25	0.15	0.08
	16	0.521	0.48	0.10	0.05
	17	0.699	0.53	0.14	0.06
	18	0.710	0.47	0.25	0.14
	19	0.525	0.80	0.66	0.49

APPENDIX F: CONFIRMATORY FACTOR ANALYSIS FOR THE 2010 NYSESLAT

New York State regulations, under Federal Title III of the *No Child Left Behind* (NCLB) Act of 2001, require annual assessment of limited English proficient students using the NYSESLAT as a state-approved assessment tool. This test has been developed in accordance with the Standards for Educational and Psychological Testing (American Educational Research Association, 1999) and the New York State testing requirements, within the principles and consistency of Universal Design and applicable federal and state testing requirements. To meet these requirements, the test was developed for five grade spans (K–1, 2–4, 5–6, 7–8, and 9–12) in four modalities (i.e., Speaking, Listening, Reading, and Writing), to assess the English language proficiency of ELL students from Kindergarten through Grade 12.

The maximum number of points by modality and grade span for the 2010 administration is depicted in the following table:

Table 52 - Maximum Number of Points by Modality and Grade Span for the 2010 NYSESLAT

Maximum Number of Points in NYSESLAT Subtests							
Grade Span	Speaking	Listening	Reading	Writing			Total Number of Points
				Writing Conventions	Pre-Writing	Writing Prompt	
	CR	MC	MC	MC	SR	CR/ER	
K–1	34	24	15	6	0	16 CR (Dev. Writing) + 2 ER	97
2–4	34	24	24	12	6	4	104
5–6	34	25	27	15	6	4	111
7–8	34	25	27	15	6	4	111
9–12	34	25	27	15	6	4	111

F.1: Reporting of the NYSESLAT Scores

The NYSESLAT scores are reported on two different scales: a scale that is a combination of Speaking and Listening modalities and another that is a combination of the Reading and Writing modalities. In 2010, the New York State Department of Education (NYSDE) expressed an interest in examining the dimensionality of the data, with an eye toward the reporting of scores on a single English Language Performance Achievement (ELPA) score.

In order to test the assumption of unidimensionality and compare it with the currently used two-dimensional model, a confirmatory factor analysis was used to examine the factor loadings based on the different combinations of the observed variables.

F.2: Method and Process

Since the primary interest of this study was to compare the current reporting structure with an alternate preconceived unidimensional structure, a confirmatory factor model was the obvious choice for analysis from the various other factor models available.

This study, therefore, tested the following two structures:

- A one-factor model with the latent variable being ELPA (English Language Proficiency Assessment)
- A two-factor model that tested the current structure of Speaking and Listening as one factor (SPKLIS), and Reading, Writing Conventions (WRCON), and Writing Constructed-Response items (WRCON, Pre-writing + Writing Prompt) as the other factor (WRTRD)

The results of these confirmatory factor analyses are based on the 2010 data set, and on the full population responses.

Since this study was not intended to assess the appropriateness of individual item loading on the different modality factors, all items were collapsed within their respective modality to produce five scores for each individual student (i.e., for Speaking, Listening, Reading, Writing Conventions (WRCON), and Writing Constructed-Response Items (WRCON)).

The factor loadings for the one- and the two-factor models were therefore based on subtests of Speaking, Listening, Reading, WRCON, and WRCON.

For the NYSESLAT, since the items were collapsed to produce the subtests discussed above, there were five modality measures that extend from a low of 6 scoring categories for the Writing Conventions to a high of 34 categories for the Speaking subtest across all grade spans. The subtests can therefore be looked upon as polytomously scored items.

Furthermore, as stated by Jöreskog and Sörbom (1995a), for ordered categorical data that are not normally distributed but assumed to reflect normally distributed underlying variables, a better way to estimate the population correlations is by using polychoric correlation coefficients. With this in mind, polychoric correlations were used with the ML estimation procedure in this study.

Although there are various methods to evaluate the confirmatory factor model, a more global approach was used to assess the relative performance of the different models. The most commonly used indices were referenced in this study, i.e., the goodness-of-fit index (GFI), the adjusted-goodness-of-fit index (AGFI), the root-mean-squared residual (RMR) and the root-mean-squared error of approximation (RMSEA), and the χ^2 index. However, a consensus as to which of these methods is best does not exist, and no special weight is given to any of the indices outlined above in this study.

According to Swygert, et al. (2001), in the case of the overall fit of the model, the χ^2 index should not be seen as a test of significance for the null of no difference in the fit of the model to the data, but should be used as an indication of how far the data deviate from the model. High values

indicate large deviance and small values indicate small deviance. Furthermore, large sample sizes will generally produce larger Chi-square values.

The GFI and AGFI indices have a range of 0 to 1, where values closer to 1 indicate a better fit of the model to the data. On the other hand, the RMR is an index produced by the average of the fitted residuals. Hu and Bentler (1999) suggest $RMR < 0.06$ as acceptable. Similarly, Hu and Bentler (1999) suggest < 0.05 as acceptable for the RMSEA, which is a measure of the goodness of fit with known distribution properties. On the other hand, Browne and Cudeck (1993) and MacCallum, Browne, and Sugawara(1996) suggest $RMSEA \leq 0.05$ as close model fit, values ranging from 0.06 to 0.08 as adequate fit, those between 0.09 and 0.10 as mediocre fit, and values greater than 0.10 as poor fit.

The path diagrams for the one- and the two-factor models are shown in Figures 25 and 26:

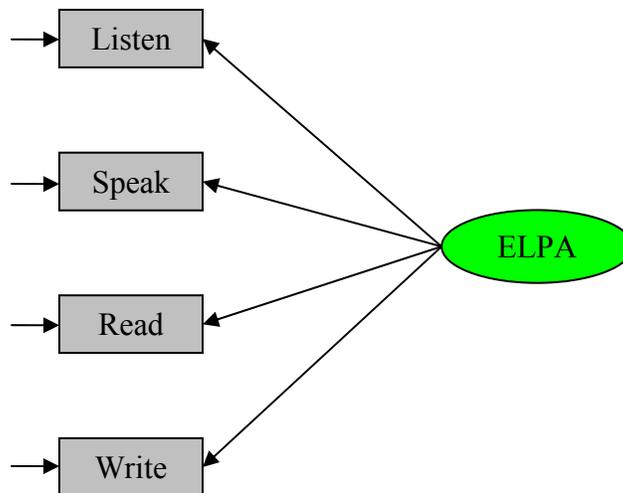


Figure 25 - Path Diagram for the One-Factor Model

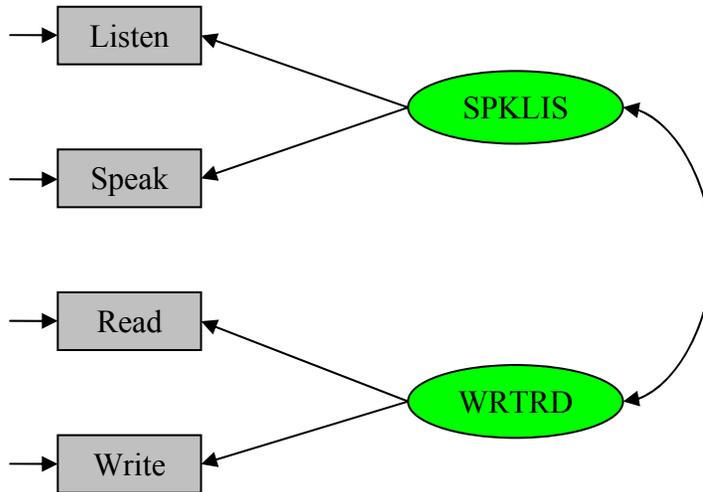


Figure 26 - Path Diagram for the Two-Factor Model

The structural equation modeling together with the polychoric correlation coefficients and ML estimation procedures were implemented by the use of Jöreskog’s and Sörbom’s Lisrel, version 8.80, computer program. The latent variable is standardized.

F.3: Results and Comments

Grade Span K–1

The summary statistics are provided in Table 52. Factor loadings for the one- and the two- factor models are provided in Table 53 with the fit indices for the two models displayed in Table 54.

Table 53 - Summary Statistics

Variable	N-count	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
Listen	52,751	20.42	3.62	-1.51	2.47	1.00	24.00
Speak	52,751	27.03	7.13	-1.47	1.92	1.00	34.00
Read	52,751	11.63	3.27	-0.78	-0.35	0.00	15.00
Write	52,751	18.51	5.19	-0.93	-0.06	1.00	24.00

Table 54 - Factor Loading Estimates for the One- and Two-Factor Models

Modalities	One-Factor Model	Two-Factor Model	
	Λ	Λ_1	Λ_2
Listen	0.68	0.82	–
Speak	0.57	0.70	–
Read	0.86	–	0.87
Write	0.89	–	0.90
		Correlation	0.77

Table 55 - Global Fit Indices for the One- and Two-Factor Models

Fit Indices	One-Factor Model	Two-Factor Model
Chi-square	7617.130	512.309
GFI	0.933	0.995
AGFI	0.663	0.976
RMR	0.068	0.028
RMSEA	0.269	0.070

Note: χ^2 with 2 *df*, $p=0.00$ for the one-factor model and χ^2 with 1 *df*, $p=0.00$ for the two-factor model.

For the one-factor model, the values of GFI and AGFI are 0.933 and 0.663, respectively, and the RMR = 0.068 and the RMSEA = 0.269 are not sufficiently small, indicating that the model fit leaves much to be desired. The Chi-square value is also sufficiently large and significant with 5 degrees of freedom, suggesting a revision of the model for a better fit.

In keeping with the current dimensionality of the test, the two-factor model with the loading of Speaking and Listening on one factor (SPKLIS), and the Writing Conventions and Writing Constructed-Response and Reading on the second factor (WRTRD) was estimated. The results indicate a great improvement in fit of the model to the data. The rather weak loading of the Listening modality (0.68) is considerably improved with the two-factor model to 0.82. Although the p -value of the Chi-square is still in the range of significance at the 0.05 level of rejection ($\chi^2 = 512.309$, $p = 0.00$), the Chi-square itself is reduced tremendously. The RMR and RMSEA values have decreased a great deal to 0.028 and 0.070 respectively, and the GFI as well as the AGFI is approaching the peak of 1.000.

Grade Span 2–4

The summary statistics are provided in Table 54. Factor loadings for the one- and the two-factor models are provided in Table 55 with the fit indices for the two models displayed in Table 56.

Table 56 - Summary Statistics

Variable	N-count	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
Listen	62,102	20.009	3.689	-1.693	3.240	1.000	24.000
Speak	62,102	29.172	5.820	-1.998	4.437	1.000	34.000
Read	62,102	17.814	4.469	-0.801	-0.043	0.000	24.000
Write	62,102	15.797	4.084	-0.792	0.042	1.000	22.000

Table 57 - Factor Loading Estimates for the One- and Two-Factor Models

Modalities	One-Factor Model	Two-Factor Model	
	Λ	Λ_1	Λ_2
Listen	0.67	0.84	–
Speak	0.61	0.69	–
Read	0.85	–	0.80
Write	0.86	–	0.89
		Correlation	0.79

Table 58 - Global Fit Indices for the One- and Two-Factor Models

Fit Indices	One-Factor Model	Two-Factor Model
Chi-square	4309.044	3263.943
GFI	0.966	0.974
AGFI	0.832	0.872
RMR	0.045	0.085
RMSEA	0.186	0.162

Note: χ^2 with 2 *df*, $p=0.00$ for the one-factor model and χ^2 with 1 *df*, $p=0.00$ for the two-factor model.

The fit of the model to the data is very poor with the one-factor model. The Chi-square value is extremely large ($\chi^2 = 4309.044$ with 5 *df*, $p = 0.00$). The value of GFI is 0.966 with a poor value for the AGFI, i.e., 0.832. The RMR is acceptable at 0.046, while the RMSEA is large at 0.186. A summary of the fit indices suggest that an alternate model is justifiable.

With the two-factor model, the Chi-square reduces to 3263.943 (with 4 *df*, $p = 0.00$) and values for GFI and AGFI increased to 0.974 and 0.872 respectively. The increase in the GFI and AGFI to an the acceptable value suggest that the two-factor model is a substantial improvement over the one-factor model and is an acceptable model. However, the RMR and RMSEA decreased slightly. Although the two-factor model shows improvement over the one-factor model, the values of RMR = 0.085 and RMSEA = 0.162 indicate less than desirable fit.

Grade Span 5–6

The summary statistics are provided in Table 58. Factor loadings for the one- and the two-factor models are provided in Table 59 with the fit indices for the two models displayed in Table 60.

Table 59 - Summary Statistics

Variable	N-count	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
Listen	28,540	19.153	4.292	-1.433	2.036	1.000	25.000
Speak	28,540	28.806	6.790	-1.868	3.162	1.000	34.000
Read	28,540	18.529	5.487	-0.556	-0.520	0.000	27.000
Write	28,540	17.300	4.618	-0.801	0.099	1.000	25.000

Table 60 - Factor Loading Estimates for the One- and Two-Factor Models

Modalities	One-Factor Model		Two-Factor Model	
	Λ	Λ_1	Λ_2	
Listen	0.69	0.81	–	
Speak	0.70	0.77	–	
Read	0.81	–	0.73	
Write	0.86	–	0.90	
		Correlation	0.83	

Table 61 - Global Fit Indices for the One- and Two-Factor Models

Fit Indices	One-Factor Model	Two-Factor Model
Chi-square	1884.852	1820.544
GFI	0.968	0.969
AGFI	0.840	0.845
RMR	0.038	0.094
RMSEA	0.182	0.178

Note: χ^2 with 2 *df*, $p=0.00$ for the one-factor model and χ^2 with 1 *df*, $p=0.00$ for the two-factor model.

There was not any improvement in the 5–6 grade span for the two-factor model when compared to the one-factor model ($\chi^2 = 1820.544$ with 4 *df*, $p = 0.00$ for the two-factor model). The RMR and the RMSEA also show an increase. Similarly, there is only a small improvement for the GFI and the AGFI, with the GFI showing a value of 0.969.

Grade Span 7–8

The summary statistics are provided in Table 61. Factor loadings for the one- and the two-factor models are provided in Table 62 with the fit indices for the two models displayed in Table 63.

Table 62 - Summary Statistics

Variable	N-count	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
Listen	24,811	17.957	4.837	-0.929	0.218	1.000	25.000
Speak	24,811	28.294	7.241	-1.596	1.896	1.000	34.000
Read	24,811	17.655	5.478	-0.493	-0.762	0.000	26.000
Write	24,811	17.703	4.835	-0.834	-0.028	2.000	25.000

Table 63 - Factor Loading Estimates for the One- and Two-Factor Models

Modalities	One-Factor Model	Two-Factor Model	
	Λ	Λ_1	Λ_2
Listen	0.82	0.86	–
Speak	0.75	0.78	–
Read	0.79	–	0.78
Write	0.89	–	0.92
		Correlation	0.93

Table 64 - Global Fit Indices for the One- and Two-Factor Models

Fit Indices	One-Factor Model	Two-Factor Model
Chi-square	1510.399	998.103
GFI	0.970	0.980
AGFI	0.852	0.901
RMR	0.029	0.026
RMSEA	0.174	0.142

Note: χ^2 with 2 *df*, $p=0.00$ for the one-factor model and χ^2 with 1 *df*, $p=0.00$ for the two-factor model.

There was slight improvement in the 7–8 grade span for the two-factor model when compared to the one-factor model. The Chi-square sees an increase from the one-factor value of 1510.399 with 5 *df* to 998.103 with 4 degrees of freedom. Similarly, GFI, AGFI, RMR, and RMSEA all indicate slightly improved fit over their counterparts in the one-factor model.

Grade Span 9–12

The summary statistics are provided in Table 64. Factor loadings for the one- and the two-factor models are provided in Table 65 with the fit indices for the two models displayed in Table 66.

Table 65 - Summary Statistics

Variable	N-count	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
Listen	40,708	18.938	4.686	-0.918	0.215	1.000	25.000
Speak	40,708	27.876	7.096	-1.385	1.370	1.000	34.000
Read	40,708	16.881	5.250	-0.270	-0.741	0.000	27.000
Write	40,708	17.801	4.902	-0.662	-0.287	1.000	25.000

Table 66 - Factor Loading Estimates for the One- and Two-Factor Models

Modalities	One-Factor Model		Two-Factor Model	
	Λ	Λ_1	Λ_2	
Listen	0.79	0.83	–	
Speak	0.72	0.75	–	
Read	0.75	–	0.75	
Write	0.87	–	0.89	
		Correlation		0.92

Table 67 - Global Fit Indices for the One- and Two-Factor Models

Fit Indices	One-Factor Model	Two-Factor Model
Chi-square	2082.209	1448.785
GFI	0.975	0.983
AGFI	0.875	0.913
SRMR	0.030	0.024
RMSEA	0.160	0.133

Note: χ^2 with 2 *df*, $p=0.00$ for the one-factor model and χ^2 with 1 *df*, $p=0.00$ for the two-factor model.

Test results for the 9–12 grade span showed slight improvement with the two-factor model over the one-factor model. The factor loadings increase with the use of the two-factor model in comparison to their loadings on the one-factor model. The Chi-square decreases ($\chi^2 = 2082.209$ with 5 *df* for the one-factor model and $\chi^2 = 1448.785$ with 4 *df* for the two-factor model). The other fit indices all indicated slightly poorer fit over the one-factor model.

F.4: Conclusion

Overall, the 2010 NYSESLAT dimensionality tests, with the use of Confirmatory Factor Analysis, produced results that substantiate the use of the two-factor model in reporting IRT scores (i.e., Speaking and Listening as one score and Writing Conventions and Writing Constructed-Response Items and Reading as the other), especially with lower grade spans. With higher grade spans, even though the one-factor model produced slightly better model fit than the two-factor model, the difference was not as noticeable as the difference observed with the lower grade span.

If one were to accept the recommendations of Browne et al. (1993) and MacCallum et al. (1996) for the evaluation of the RMSEA as a fit index, and Hu and Bentler (1999) for the RMR index, in conjunction with acceptable values of other fit indices, then the two-factor model for the K–1 grade span has the most acceptable fit of the model to the data. Besides having acceptable RMR and RMSEA under the various criteria discussed above, the GFI and the AGFI were very close to 1.0.

In the 2–4 grade span, the unidimensionality of the NYSESLAT scoring cannot be justifiable because of lack of substantial improvement of the two-factor model fit to the data over the unidimensional model.

With respect to the analysis of the 5–6, 7–8, and the 9–12 grade span tests, the factor loadings based on the one-factor model remain much the same when they are loaded on two-factors. Since the two-factor model does not show any improvement over the fit of the one-factor model, it

could be left alone to preserve the status quo of the NYSESLAT, at least for the next few years. Extra care could also be taken towards content and substantive issues in selecting items for the future.

APPENDIX G: SCALE SCORE SUMMARY BY SUBGROUPS**G.1a: Kindergarten Listening and Speaking**

Grade	Test	Group	N-count	Mean	Median	SD	IQR
K	LS	Female	12,578	596.19	596	50.85	59
K	LS	Male	13,289	594.68	591	49.24	59
K	LS	American Indian or Alaska Native	64	601.84	593.5	39.71	58.5
K	LS	Black or African American	1,225	592.26	591	48.96	52
K	LS	Asian	5,376	591.21	591	51.65	65
K	LS	Native Hawaiian/Other Pacific Islander	19	616.68	606	48.04	59
K	LS	Hispanic or Latino	15,705	597.93	596	46.54	56
K	LS	White	3,279	590.71	596	61.71	80
K	LS	Multiracial (not Hispanic origin)	188	604.38	601	45.10	47
K	LS	NYC	15,745	595.94	596	50.87	59
K	LS	Big 4 Cities	1,092	587.74	591	46.42	58
K	LS	High Need Urban/Suburban	3,431	586.67	587	49.72	62
K	LS	High Need Rural	164	591.00	587	45.43	55
K	LS	Average Need	3,034	601.84	601	48.36	60
K	LS	Low Need	1,358	599.69	601	43.39	52
K	LS	Charter Schools	277	611.09	611	46.14	49
K	LS	0 Years LEP	15,181	594.59	596	52.25	62
K	LS	1 Years LEP	10,675	596.50	596	46.63	56
K	LS	Arabic	587	590.34	591	54.91	62
K	LS	Bengali	682	598.93	596	50.42	56
K	LS	Chinese	1,250	589.23	587	50.40	58
K	LS	English	1,569	604.83	601	44.60	45
K	LS	Spanish	14,826	597.22	596	46.60	56
K	LS	Other Language	7,013	590.43	591	56.64	69
K	LS	Students with Disabilities	4,606	573.68	575	48.62	63

G.1b: Kindergarten Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
K	RW	Female	12,663	551.53	546	47.81	54
K	RW	Male	13,506	544.23	542	47.74	56
K	RW	American Indian or Alaska Native	65	558.31	546	53.70	50
K	RW	Black or African American	1,225	537.74	533	45.92	54
K	RW	Asian	5,585	565.68	562	52.74	71
K	RW	Native Hawaiian/Other Pacific Islander	19	567.74	551	53.61	83
K	RW	Hispanic or Latino	15,925	542.98	537	42.72	56
K	RW	White	3,149	543.28	537	55.78	70
K	RW	Multiracial (not Hispanic origin)	189	555.25	551	48.09	54
K	RW	NYC	15,903	551.96	546	49.73	65
K	RW	Big 4 Cities	1,116	528.93	524	39.45	47
K	RW	High Need Urban/Suburban	3,542	536.20	533	44.33	54
K	RW	High Need Rural	169	529.32	533	39.36	43
K	RW	Average Need	3,034	546.17	542	41.58	48
K	RW	Low Need	1,369	547.13	542	46.38	58
K	RW	Charter Schools	273	558.78	556	43.64	46
K	RW	0 Years LEP	15,386	551.38	546	49.60	65
K	RW	1 Years LEP	10,769	542.56	537	44.93	56
K	RW	Arabic	604	539.83	537	45.61	60
K	RW	Bengali	696	570.06	568	54.95	71
K	RW	Chinese	1,291	569.43	562	49.84	62
K	RW	English	1,567	545.76	542	42.88	52
K	RW	Spanish	15,046	543.12	542	43.15	56
K	RW	Other Language	7,021	552.40	546	55.07	65
K	RW	Students with Disabilities	4,382	527.07	524	44.09	52

G.2a: Grade 1 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
1	LS	Female	13,075	635.26	632	55.78	68
1	LS	Male	14,307	632.20	632	56.98	68
1	LS	American Indian or Alaska Native	44	643.89	647.5	66.89	97.5
1	LS	Black or African American	1,277	628.88	624	59.01	73
1	LS	Asian	5,278	635.97	632	59.66	68
1	LS	Native Hawaiian/Other Pacific Islander	16	645.50	647.5	58.16	100
1	LS	Hispanic or Latino	16,982	635.63	632	52.30	68
1	LS	White	3,633	623.35	617	66.39	90
1	LS	Multiracial (not Hispanic origin)	126	643.99	642	53.64	63
1	LS	NYC	16,740	635.00	632	55.93	68
1	LS	Big 4 Cities	1,233	620.93	617	56.68	66
1	LS	High Need Urban/Suburban	4,055	618.71	617	57.82	70
1	LS	High Need Rural	190	645.56	642	51.44	58
1	LS	Average Need	2,985	642.10	642	55.72	63
1	LS	Low Need	1,306	646.54	642	52.63	58
1	LS	Charter Schools	242	647.17	642	44.33	52
1	LS	0 Years LEP	14,618	632.92	632	57.73	73
1	LS	1 Years LEP	4,379	620.25	617	59.09	66
1	LS	2 Years LEP	8,423	641.46	642	51.31	63
1	LS	Arabic	646	626.97	628	63.16	78
1	LS	Bengali	716	638.66	642	56.21	63
1	LS	Chinese	1,123	642.09	642	56.68	58
1	LS	English	1,249	645.29	642	53.59	58
1	LS	Spanish	16,057	635.08	632	52.19	68
1	LS	Other Language	7,681	627.15	624	63.76	82
1	LS	Students with Disabilities	5,541	606.70	606	53.40	70

G.2b: Grade 1 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
1	RW	Female	13,311	622.75	628	59.13	74
1	RW	Male	14,533	610.73	611	62.46	87
1	RW	American Indian or Alaska Native	45	610.31	628	66.28	93
1	RW	Black or African American	1,308	601.87	599	61.06	93
1	RW	Asian	5,446	635.65	655	60.19	107
1	RW	Native Hawaiian/Other Pacific Islander	16	631.00	655	61.79	103
1	RW	Hispanic or Latino	17,233	612.92	611	60.27	87
1	RW	White	3,638	610.67	611	60.84	81
1	RW	Multiracial (not Hispanic origin)	129	624.32	628	59.76	65
1	RW	NYC	17,116	617.71	611	61.80	81
1	RW	Big 4 Cities	1,259	594.27	590	63.54	82
1	RW	High Need Urban/Suburban	4,091	608.53	611	59.06	87
1	RW	High Need Rural	194	618.01	611	56.81	74
1	RW	Average Need	2,996	623.30	628	59.56	74
1	RW	Low Need	1,311	631.08	628	55.22	65
1	RW	Charter Schools	242	626.64	628	51.51	65
1	RW	0 Years LEP	14,992	615.56	611	63.14	87
1	RW	1 Years LEP	4,449	605.41	611	61.21	93
1	RW	2 Years LEP	8,440	623.73	628	56.54	74
1	RW	Arabic	675	601.86	599	67.81	109
1	RW	Bengali	724	638.78	655	59.80	107
1	RW	Chinese	1,150	644.97	655	53.91	95
1	RW	English	1,251	624.83	628	58.40	74
1	RW	Spanish	16,310	612.30	611	60.40	87
1	RW	Other Language	7,824	618.57	611	61.89	81
1	RW	Students with Disabilities	5,513	591.03	590	57.34	77

G.3a: Grade 2 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
2	LS	Female	11,037	666.09	666	45.01	54
2	LS	Male	12,476	663.21	666	47.26	57
2	LS	American Indian or Alaska Native	53	668.34	678	39.52	46
2	LS	Black or African American	1,259	661.34	666	48.11	57
2	LS	Asian	4,144	665.47	672	50.68	54
2	LS	Native Hawaiian/Other Pacific Islander	16	669.31	689	42.13	49
2	LS	Hispanic or Latino	14,929	667.82	666	41.89	50
2	LS	White	3,064	649.66	643	54.90	67
2	LS	Multiracial (not Hispanic origin)	59	662.76	661	51.60	49
2	LS	NYC	14,874	663.79	666	45.03	54
2	LS	Big 4 Cities	1,236	656.29	661	49.27	57
2	LS	High Need Urban/Suburban	3,445	656.67	661	48.03	57
2	LS	High Need Rural	134	670.25	672	42.86	59
2	LS	Average Need	2,304	675.56	678	46.78	50
2	LS	Low Need	1,040	681.40	678	44.93	50
2	LS	Charter Schools	202	682.63	678	37.29	41
2	LS	0 Years LEP	12,771	664.36	666	45.51	54
2	LS	1 Years LEP	2,440	642.95	647	54.55	72
2	LS	2 Years LEP	2,827	660.06	656	45.71	57
2	LS	3 Years LEP	5,544	676.43	672	39.92	50
2	LS	Arabic	588	660.07	666	52.84	61
2	LS	Bengali	601	668.38	672	46.79	54
2	LS	Chinese	827	671.35	672	48.22	59
2	LS	English	958	676.71	672	43.91	50
2	LS	Spanish	14,201	667.22	666	41.89	50
2	LS	Other Language	6,460	655.55	656	52.97	68
2	LS	Students with Disabilities	5,778	644.58	643	42.90	54

G.3 b: Grade 2 Reading and Writing

Grade	Test	Group	N- count	Mean	Median	SD	IQR
2	RW	Female	10,606	625.50	626	39.47	52
2	RW	Male	11,502	620.79	622	40.98	54
2	RW	American Indian or Alaska Native	47	632.81	631	33.57	46
2	RW	Black or African American	1,155	614.35	613	39.09	53
2	RW	Asian	3,911	635.25	636	42.82	56
2	RW	Native Hawaiian/Other Pacific Islander	16	630.81	646	44.74	74
2	RW	Hispanic or Latino	14,176	618.82	617	37.73	54
2	RW	White	2,753	632.44	631	44.16	65
2	RW	Multiracial (not Hispanic origin)	56	621.30	619.5	47.46	64
2	RW	NYC	13,949	622.33	622	39.83	50
2	RW	Big 4 Cities	1,100	612.28	613	39.52	57
2	RW	High Need Urban/Suburban	3,223	620.51	622	38.04	54
2	RW	High Need Rural	124	629.22	626	42.16	59
2	RW	Average Need	2,229	628.39	631	42.16	59
2	RW	Low Need	1,016	642.67	641	44.21	60
2	RW	Charter Schools	199	630.16	626	32.83	37
2	RW	0 Years LEP	12,013	620.36	622	38.96	54
2	RW	1 Years LEP	2,051	615.41	613	44.18	62
2	RW	2 Years LEP	2,697	628.37	626	42.84	63
2	RW	3 Years LEP	5,416	629.18	631	39.35	48
2	RW	Arabic	528	617.87	622	40.90	49
2	RW	Bengali	581	631.64	636	41.44	59
2	RW	Chinese	802	641.71	641	43.52	52
2	RW	English	916	628.70	631	41.03	59
2	RW	Spanish	13,460	618.34	617	37.53	49
2	RW	Other Language	5,937	630.07	631	43.76	59
2	RW	Students with Disabilities	5,075	609.50	604	40.90	56

G.4a: Grade 3 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
3	LS	Female	10,198	685.15	685	50.46	58
3	LS	Male	11,628	679.75	685	53.77	62
3	LS	American Indian or Alaska Native	32	696.19	693	45.28	53
3	LS	Black or African American	1,165	672.45	678	52.93	59
3	LS	Asian	3,668	679.11	685	56.13	64.5
3	LS	Native Hawaiian/Other Pacific Islander	8	707.75	697.5	46.38	23
3	LS	Hispanic or Latino	14,029	688.40	685	47.09	53
3	LS	White	2,875	661.68	661	62.59	80
3	LS	Multiracial (not Hispanic origin)	29	673.66	672	65.60	82
3	LS	NYC	14,022	683.19	685	51.17	58
3	LS	Big 4 Cities	1,234	673.95	678	51.48	59
3	LS	High Need Urban/Suburban	3,206	673.66	678	55.67	59
3	LS	High Need Rural	132	681.03	685	43.89	41
3	LS	Average Need	2,057	689.20	693	55.34	53
3	LS	Low Need	854	694.73	693	49.91	64
3	LS	Charter Schools	130	695.87	693	37.54	42
3	LS	0 Years LEP	12,497	684.54	685	51.52	58
3	LS	1 Years LEP	1,829	647.45	652	62.33	79
3	LS	2 Years LEP	1,759	674.90	678	50.81	59
3	LS	3 Years LEP	1,389	673.75	672	50.65	66
3	LS	4 Years LEP	4,373	695.61	693	43.43	42
3	LS	Arabic	561	668.50	678	57.05	74
3	LS	Bengali	578	687.16	693	51.37	53
3	LS	Chinese	608	679.84	685	57.93	67
3	LS	English	794	694.80	693	44.86	48
3	LS	Spanish	13,386	687.84	685	46.98	53
3	LS	Other Language	5,962	668.89	672	60.34	74
3	LS	Students with Disabilities	5,790	663.17	666	49.57	61

G.4b: Grade 3 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
3	RW	Female	9,968	648.92	652	41.21	47
3	RW	Male	11,018	642.82	646	43.87	60
3	RW	American Indian or Alaska Native	32	656.34	652	41.46	36.5
3	RW	Black or African American	1,098	634.62	636	44.22	65
3	RW	Asian	3,521	652.86	659	45.66	56
3	RW	Native Hawaiian/Other Pacific Islander	8	666.50	664	27.03	46.5
3	RW	Hispanic or Latino	13,659	644.88	646	40.61	51
3	RW	White	2,623	646.21	652	46.18	56
3	RW	Multiracial (not Hispanic origin)	27	639.70	641	47.60	78
3	RW	NYC	13,487	645.58	646	41.90	56
3	RW	Big 4 Cities	1,162	634.95	636	45.84	65
3	RW	High Need Urban/Suburban	3,056	644.66	646	42.33	51
3	RW	High Need Rural	130	641.98	641	39.09	48
3	RW	Average Need	1,989	650.62	652	44.88	56
3	RW	Low Need	846	657.91	659	43.85	46
3	RW	Charter Schools	127	651.95	652	35.12	42
3	RW	0 Years LEP	12,090	644.81	646	42.05	56
3	RW	1 Years LEP	1,576	629.95	631	49.03	73
3	RW	2 Years LEP	1,693	645.55	646	45.49	56
3	RW	3 Years LEP	1,330	646.22	646	43.70	56
3	RW	4 Years LEP	4,309	654.01	652	38.64	51
3	RW	Arabic	506	634.50	641	45.07	65
3	RW	Bengali	560	656.97	659	41.21	51
3	RW	Chinese	581	655.89	659	47.96	56
3	RW	English	775	653.07	652	41.54	56
3	RW	Spanish	13,027	644.47	646	40.58	56
3	RW	Other Language	5,591	646.39	652	46.48	56
3	RW	Students with Disabilities	5,361	627.34	626	41.77	63

G.5a: Grade 4 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
4	LS	Female	8,861	694.88	693	55.78	64
4	LS	Male	10,896	690.24	693	59.38	69
4	LS	American Indian or Alaska Native	20	724.90	722	46.90	71
4	LS	Black or African American	1,134	684.47	685	57.79	62
4	LS	Asian	3,240	683.95	693	64.52	83
4	LS	Native Hawaiian/Other Pacific Islander	9	679.56	678	54.12	71
4	LS	Hispanic or Latino	12,703	700.67	702	51.47	58
4	LS	White	2,634	666.94	666	66.42	89
4	LS	Multiracial (not Hispanic origin)	23	675.35	678	74.41	78
4	LS	NYC	12,909	693.95	693	56.88	69
4	LS	Big 4 Cities	1,140	681.00	685	56.64	62
4	LS	High Need Urban/Suburban	3,014	680.86	685	62.69	83
4	LS	High Need Rural	142	694.06	693	52.44	74
4	LS	Average Need	1,665	700.58	702	56.11	58
4	LS	Low Need	710	704.09	702	54.09	58
4	LS	Charter Schools	99	707.89	714	46.23	52
4	LS	0 Years LEP	11,384	695.82	702	57.19	64
4	LS	1 Years LEP	1,831	656.37	661	64.19	81
4	LS	2 Years LEP	1,211	683.54	685	55.04	62
4	LS	3 Years LEP	977	668.92	672	63.89	86
4	LS	4 Years LEP	1,154	694.33	693	51.43	69
4	LS	5 Years LEP	3,249	708.75	702	45.98	52
4	LS	Arabic	506	678.78	685	65.82	75
4	LS	Bengali	423	695.38	693	61.94	74
4	LS	Chinese	486	684.35	693	65.41	83
4	LS	English	685	707.71	702	49.06	52
4	LS	Spanish	12,217	700.27	702	51.60	58
4	LS	Other Language	5,512	673.89	678	65.21	86
4	LS	Students with Disabilities	5,771	678.71	678	55.01	67

G.5b: Grade 4 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
4	RW	Female	8,703	663.94	665	45.55	57
4	RW	Male	10,526	655.85	659	49.13	56
4	RW	American Indian or Alaska Native	20	668.50	659	40.29	69.5
4	RW	Black or African American	1,083	649.24	652	47.38	65
4	RW	Asian	3,114	662.79	665	51.83	62
4	RW	Native Hawaiian/Other Pacific Islander	9	663.33	665	38.81	52
4	RW	Hispanic or Latino	12,503	661.48	665	44.46	57
4	RW	White	2,499	650.99	659	55.09	65
4	RW	Multiracial (not Hispanic origin)	19	652.53	665	42.27	69
4	RW	NYC	12,600	659.67	665	46.81	62
4	RW	Big 4 Cities	1,078	647.29	652	46.31	56
4	RW	High Need Urban/Suburban	2,899	656.27	665	49.83	67
4	RW	High Need Rural	141	658.30	659	47.93	51
4	RW	Average Need	1,636	665.61	665	49.19	57
4	RW	Low Need	704	672.94	673	49.46	60
4	RW	Charter Schools	99	668.67	665	40.74	47
4	RW	0 Years LEP	11,113	660.44	665	46.26	57
4	RW	1 Years LEP	1,697	637.28	636	54.73	77
4	RW	2 Years LEP	1,181	656.30	659	50.46	71
4	RW	3 Years LEP	927	647.63	652	55.89	69
4	RW	4 Years LEP	1,129	666.46	665	42.94	52
4	RW	5 Years LEP	3,227	669.59	665	42.13	47
4	RW	Arabic	460	647.34	652	49.76	69
4	RW	Bengali	411	667.59	673	46.18	52
4	RW	Chinese	471	662.29	665	55.68	67
4	RW	English	675	667.93	665	45.75	52
4	RW	Spanish	12,016	661.36	665	44.52	57
4	RW	Other Language	5,263	654.10	659	53.28	71
4	RW	Students with Disabilities	5,541	641.71	646	44.51	60

G.6a: Grade 5 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
5	LS	Female	7,144	678.55	684	45.78	48
5	LS	Male	9,191	676.08	684	49.58	52
5	LS	American Indian or Alaska Native	21	681.24	684	48.27	69
5	LS	Black or African American	1,025	672.00	678	46.17	56
5	LS	Asian	2,514	668.34	678	54.16	69
5	LS	Native Hawaiian/Other Pacific Islander	10	675.70	673	28.06	30
5	LS	Hispanic or Latino	10,449	684.05	690	43.02	40
5	LS	White	2,278	658.59	660	54.59	71
5	LS	Multiracial (not Hispanic origin)	21	654.52	668	60.71	95
5	LS	NYC	10,765	678.08	684	46.63	48
5	LS	Big 4 Cities	973	668.85	673	46.16	56
5	LS	High Need Urban/Suburban	2,448	668.82	678	53.23	64.5
5	LS	High Need Rural	112	678.36	684	45.45	46
5	LS	Average Need	1,231	683.67	690	50.12	54
5	LS	Low Need	506	687.36	690	41.66	46
5	LS	Charter Schools	171	699.68	697	39.01	47
5	LS	0 Years LEP	9,572	679.34	684	47.57	44
5	LS	1 Years LEP	1,691	649.65	652	50.48	66
5	LS	2 Years LEP	1,120	675.17	678	43.95	54
5	LS	3 Years LEP	606	673.45	673	39.28	52
5	LS	4 Years LEP	751	665.01	678	61.97	66
5	LS	5 Years LEP	767	678.85	678	41.89	52
5	LS	6 Years LEP or More	1,858	696.51	697	34.74	36
5	LS	Arabic	428	666.69	678	49.89	71.5
5	LS	Bengali	311	672.98	678	46.44	59
5	LS	Chinese	378	659.45	664	56.38	84
5	LS	English	510	691.81	690	43.56	46
5	LS	Spanish	10,018	683.62	690	43.14	40
5	LS	Other Language	4,736	664.18	668	53.79	72
5	LS	Students with Disabilities	5,333	671.22	678	44.13	49

G.6b: Grade 5 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
5	RW	Female	6,910	684.13	684	39.28	50
5	RW	Male	8,624	677.44	680	40.85	51
5	RW	American Indian or Alaska Native	20	678.50	672	36.27	53.5
5	RW	Black or African American	971	669.56	669	41.32	59
5	RW	Asian	2,266	685.67	688	44.27	58
5	RW	Native Hawaiian/Other Pacific Islander	10	702.00	723.5	42.11	66
5	RW	Hispanic or Latino	10,130	680.83	684	36.57	50
5	RW	White	2,102	678.58	680	49.49	65
5	RW	Multiracial (not Hispanic origin)	20	661.05	661	39.18	51
5	RW	NYC	10,185	682.09	684	39.10	50
5	RW	Big 4 Cities	912	664.88	665	37.65	55
5	RW	High Need Urban/Suburban	2,345	674.04	680	43.93	54
5	RW	High Need Rural	109	672.66	676	36.45	49
5	RW	Average Need	1,197	683.75	684	40.88	50
5	RW	Low Need	492	695.16	695	41.15	51
5	RW	Charter Schools	169	692.70	693	32.48	44
5	RW	0 Years LEP	9,096	681.86	684	39.14	50
5	RW	1 Years LEP	1,505	665.71	665	43.47	63
5	RW	2 Years LEP	1,075	679.43	680	43.46	60
5	RW	3 Years LEP	582	680.57	684	39.30	53
5	RW	4 Years LEP	721	667.92	676	49.86	64
5	RW	5 Years LEP	756	687.53	688	37.81	49.5
5	RW	6 Years LEP or More	1,826	687.17	688	34.09	43
5	RW	Arabic	386	669.69	669	44.59	59
5	RW	Bengali	294	682.55	684	42.59	58
5	RW	Chinese	330	681.74	684	46.15	65
5	RW	English	500	686.72	688	36.89	48
5	RW	Spanish	9,706	680.51	684	36.59	50
5	RW	Other Language	4,360	679.96	684	46.82	62
5	RW	Students with Disabilities	5,045	669.44	669	39.68	49

G.7a: Grade 6 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
6	LS	Female	6,327	676.84	684	49.10	56
6	LS	Male	7,555	672.92	684	52.90	59
6	LS	American Indian or Alaska Native	27	675.07	673	44.33	62
6	LS	Black or African American	938	669.93	673	49.43	66
6	LS	Asian	2,279	658.39	664	58.51	85
6	LS	Native Hawaiian/Other Pacific Islander	10	676.40	671	22.66	26
6	LS	Hispanic or Latino	8,595	683.41	690	45.85	50
6	LS	White	1,984	660.20	664	55.74	71
6	LS	Multiracial (not Hispanic origin)	30	646.23	654	41.28	53
6	LS	NYC	9,072	674.04	684	50.47	59
6	LS	Big 4 Cities	888	671.33	678	52.38	66
6	LS	High Need Urban/Suburban	1,885	670.28	684	57.31	69
6	LS	High Need Rural	82	692.33	690	52.76	54
6	LS	Average Need	1,032	682.35	690	48.72	58
6	LS	Low Need	401	684.82	690	48.21	54
6	LS	Charter Schools	136	693.39	690	32.46	41
6	LS	0 Years LEP	8,062	675.33	684	51.01	56
6	LS	1 Years LEP	1,701	646.23	648	52.98	78
6	LS	2 Years LEP	1,111	674.41	678	46.76	59
6	LS	3 Years LEP	571	679.17	684	40.87	52
6	LS	4 Years LEP	460	669.43	684	63.39	82
6	LS	5 Years LEP	470	686.12	690	42.33	46
6	LS	6 Years LEP or More	1,530	698.32	697	39.67	47
6	LS	Arabic	389	662.35	673	52.64	68
6	LS	Bengali	283	671.49	678	49.42	66
6	LS	Chinese	300	643.77	648	59.17	96.5
6	LS	English	386	687.69	690	42.15	50
6	LS	Spanish	8,291	683.19	690	45.94	54
6	LS	Other Language	4,276	660.27	664	56.63	84
6	LS	Students with Disabilities	4,686	674.32	678	46.03	56

G.7b: Grade 6 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
6	RW	Female	6,154	685.52	688	40.94	52
6	RW	Male	7,090	676.25	676	42.48	60
6	RW	American Indian or Alaska Native	26	679.04	686	35.89	53
6	RW	Black or African American	893	675.06	676	40.96	58
6	RW	Asian	2,057	678.77	680	48.41	72
6	RW	Native Hawaiian/Other Pacific Islander	10	706.00	705	41.60	33
6	RW	Hispanic or Latino	8,404	680.88	684	38.36	53
6	RW	White	1,801	686.05	688	48.45	65
6	RW	Multiracial (not Hispanic origin)	29	665.79	665	51.37	53
6	RW	NYC	8,572	680.35	680	41.58	57
6	RW	Big 4 Cities	840	674.07	676	42.05	58
6	RW	High Need Urban/Suburban	1,837	678.26	684	45.02	60
6	RW	High Need Rural	78	690.74	695	35.85	44
6	RW	Average Need	1,013	685.13	688	39.88	52
6	RW	Low Need	388	698.25	695	42.83	48
6	RW	Charter Schools	135	693.96	697	26.31	28
6	RW	0 Years LEP	7,705	679.82	680	41.41	57
6	RW	1 Years LEP	1,490	666.57	665	42.58	63
6	RW	2 Years LEP	1,077	682.65	680	45.21	62
6	RW	3 Years LEP	562	689.68	688	39.06	55
6	RW	4 Years LEP	463	672.94	684	54.17	82
6	RW	5 Years LEP	462	688.70	693	38.41	44
6	RW	6 Years LEP or More	1,502	692.73	693	34.82	41
6	RW	Arabic	340	668.47	672	43.23	63
6	RW	Bengali	271	679.83	684	44.70	69
6	RW	Chinese	254	678.94	676	49.74	79
6	RW	English	382	686.86	688	39.18	52
6	RW	Spanish	8,103	680.56	684	38.43	53
6	RW	Other Language	3,933	681.09	684	48.10	65
6	RW	Students with Disabilities	4,424	675.41	676	40.55	54

G.8a: Grade 7 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
7	LS	Female	5,923	699.65	707	51.60	65
7	LS	Male	7,279	694.65	702	53.39	71
7	LS	American Indian or Alaska Native	20	696.80	707	41.90	32
7	LS	Black or African American	927	696.25	702	47.20	60
7	LS	Asian	2,470	678.43	684	58.44	86
7	LS	Native Hawaiian/Other Pacific Islander	6	717.17	719.5	28.25	32
7	LS	Hispanic or Latino	8,078	704.58	713	48.62	50
7	LS	White	1,656	689.49	697	56.00	73
7	LS	Multiracial (not Hispanic origin)	20	660.75	651.5	49.90	75
7	LS	NYC	9,092	696.53	707	52.64	68
7	LS	Big 4 Cities	808	693.09	702	50.60	63
7	LS	High Need Urban/Suburban	1,674	692.08	702	55.67	74
7	LS	High Need Rural	69	706.54	713	56.64	50
7	LS	Average Need	861	701.52	713	51.32	61
7	LS	Low Need	405	712.01	713	46.61	55
7	LS	Charter Schools	90	726.38	726	36.44	41
7	LS	0 Years LEP	7,834	698.40	707	53.32	65
7	LS	1 Years LEP	1,718	666.74	666	51.46	72
7	LS	2 Years LEP	1,205	695.26	697	46.08	60
7	LS	3 Years LEP	426	706.48	707	38.24	50
7	LS	4 Years LEP	444	697.03	713	59.82	63
7	LS	5 Years LEP	291	721.90	719	33.48	41
7	LS	6 Years LEP or More	1,287	719.67	726	41.65	46
7	LS	Arabic	386	690.63	697	52.54	70
7	LS	Bengali	259	693.28	697	53.85	70
7	LS	Chinese	393	670.90	673	58.68	88
7	LS	English	310	718.35	719	44.70	46
7	LS	Spanish	7,853	704.04	713	48.75	54
7	LS	Other Language	4,029	684.45	692	56.35	81
7	LS	Students with Disabilities	3,891	702.06	707	44.90	54

G.8b: Grade 7 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
7	RW	Female	5,777	670.69	673	40.78	53
7	RW	Male	6,801	663.23	665	40.98	58
7	RW	American Indian or Alaska Native	20	676.50	673.5	39.04	64
7	RW	Black or African American	903	664.49	665	39.79	54
7	RW	Asian	2,267	663.18	662	46.03	69
7	RW	Native Hawaiian/Other Pacific Islander	6	695.33	710	38.60	67
7	RW	Hispanic or Latino	7,807	668.07	669	38.02	51
7	RW	White	1,535	667.71	673	47.09	65
7	RW	Multiracial (not Hispanic origin)	16	656.44	658	40.61	57
7	RW	NYC	8,601	666.28	669	40.21	54
7	RW	Big 4 Cities	752	660.96	665	40.49	61
7	RW	High Need Urban/Suburban	1,648	661.29	665	43.15	61
7	RW	High Need Rural	65	675.14	673	35.00	59
7	RW	Average Need	832	673.97	678	40.66	52
7	RW	Low Need	399	686.94	687	44.78	52
7	RW	Charter Schools	89	686.44	687	35.21	38
7	RW	0 Years LEP	7,472	666.48	669	40.02	54
7	RW	1 Years LEP	1,503	650.61	648	41.63	60
7	RW	2 Years LEP	1,178	667.66	669	43.13	59
7	RW	3 Years LEP	432	676.03	678	39.99	46
7	RW	4 Years LEP	436	665.81	673	47.95	65
7	RW	5 Years LEP	290	677.89	682	35.23	49
7	RW	6 Years LEP or More	1,272	679.85	682	37.03	45
7	RW	Arabic	343	661.71	662	44.12	64
7	RW	Bengali	247	666.96	669	43.51	63
7	RW	Chinese	354	661.31	658	47.69	68
7	RW	English	305	679.49	682	41.55	52
7	RW	Spanish	7,581	667.76	669	38.05	51
7	RW	Other Language	3,777	664.22	665	45.30	66
7	RW	Students with Disabilities	3,712	661.67	662	37.68	49

G.9a: Grade 8 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
8	LS	Female	5,816	698.71	707	52.56	68
8	LS	Male	7,155	698.27	707	54.55	68
8	LS	American Indian or Alaska Native	14	682.29	694.5	58.16	80
8	LS	Black or African American	956	700.12	707	50.42	65
8	LS	Asian	2,613	682.78	684	57.06	84
8	LS	Native Hawaiian/Other Pacific Islander	4	700.25	707	37.35	59.5
8	LS	Hispanic or Latino	7,782	704.77	713	51.23	58
8	LS	White	1,533	694.30	697	55.05	71
8	LS	Multiracial (not Hispanic origin)	34	662.29	656.5	49.09	91
8	LS	NYC	9,136	698.99	707	53.28	68
8	LS	Big 4 Cities	836	684.65	692	53.60	81
8	LS	High Need Urban/Suburban	1,537	696.13	707	55.88	78
8	LS	High Need Rural	80	707.51	719	51.15	63
8	LS	Average Need	823	702.10	707	54.11	77
8	LS	Low Need	386	706.15	713	50.56	63
8	LS	Charter Schools	72	729.78	730	35.93	38.5
8	LS	0 Years LEP	7,606	701.19	707	54.34	68
8	LS	1 Years LEP	1,859	670.39	669	52.92	77
8	LS	2 Years LEP	1,296	694.13	697	46.79	63
8	LS	3 Years LEP	471	704.04	707	43.21	58
8	LS	4 Years LEP	393	699.53	707	56.34	68
8	LS	5 Years LEP	264	727.67	726	36.47	30
8	LS	6 Years LEP or More	1,074	722.27	726	43.14	53
8	LS	Arabic	404	687.77	692	57.56	82.5
8	LS	Bengali	278	699.72	702	51.29	74
8	LS	Chinese	427	678.17	676	55.83	74
8	LS	English	302	711.13	713	50.22	55
8	LS	Spanish	7,575	704.55	713	51.19	58
8	LS	Other Language	4,011	689.08	692	55.94	76
8	LS	Students with Disabilities	3,412	707.28	713	45.58	54

G.9b: Grade 8 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
8	RW	Female	5,630	674.54	678	41.20	59
8	RW	Male	6,792	668.11	669	41.82	59
8	RW	American Indian or Alaska Native	12	659.92	665.5	34.26	45
8	RW	Black or African American	927	668.27	669	41.14	59
8	RW	Asian	2,416	668.80	669	44.68	63
8	RW	Native Hawaiian/Other Pacific Islander	4	697.75	704.5	55.17	85.5
8	RW	Hispanic or Latino	7,574	671.45	673	39.71	53
8	RW	White	1,423	676.23	687	44.76	62
8	RW	Multiracial (not Hispanic origin)	30	637.63	626	50.54	72
8	RW	NYC	8,739	671.03	673	40.36	56
8	RW	Big 4 Cities	780	656.94	658	42.28	63
8	RW	High Need Urban/Suburban	1,493	668.56	673	45.02	65
8	RW	High Need Rural	77	677.86	678	40.22	62
8	RW	Average Need	787	677.49	678	44.22	62
8	RW	Low Need	373	690.20	692	42.78	56
8	RW	Charter Schools	72	693.53	692	29.55	41
8	RW	0 Years LEP	7,285	671.64	673	40.23	53
8	RW	1 Years LEP	1,702	656.69	651	44.08	63
8	RW	2 Years LEP	1,268	669.22	669	42.18	59
8	RW	3 Years LEP	467	678.48	682	43.70	55
8	RW	4 Years LEP	392	670.92	678	50.82	80.5
8	RW	5 Years LEP	261	688.98	692	31.69	41
8	RW	6 Years LEP or More	1,038	684.22	687	36.85	48
8	RW	Arabic	355	661.36	662	45.40	64
8	RW	Bengali	273	673.71	678	44.42	65
8	RW	Chinese	392	669.85	667	45.52	65
8	RW	English	291	680.19	682	41.91	59
8	RW	Spanish	7,377	671.21	673	39.70	53
8	RW	Other Language	3,756	670.74	673	44.22	65
8	RW	Students with Disabilities	3,271	668.44	673	38.28	53

G.10a: Grade 9 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
9	LS	Female	6,490	694.94	697	57.22	83
9	LS	Male	7,950	695.66	697	58.00	83
9	LS	American Indian or Alaska Native	30	693.07	683.5	74.73	92
9	LS	Black or African American	1,299	693.28	693	51.97	72
9	LS	Asian	3,239	680.10	678	55.04	74
9	LS	Native Hawaiian/Other Pacific Islander	6	685.33	676.5	50.02	109
9	LS	Hispanic or Latino	8,458	700.81	706	58.75	89
9	LS	White	1,316	702.36	701	54.71	68
9	LS	Multiracial (not Hispanic origin)	67	681.00	678	46.98	61
9	LS	NYC	9,106	696.94	701	57.08	80
9	LS	Big 4 Cities	952	695.64	689	64.67	85
9	LS	High Need Urban/Suburban	1,707	690.28	693	56.69	80
9	LS	High Need Rural	79	701.01	701	54.35	68
9	LS	Average Need	1,050	696.79	701	58.10	80
9	LS	Low Need	539	698.98	701	54.38	77
9	LS	Charter Schools	48	735.33	745	30.64	40
9	LS	0 Years LEP	8,203	697.14	701	59.43	83
9	LS	1 Years LEP	2,621	667.72	665	51.44	70
9	LS	2 Years LEP	1,450	694.22	693	48.22	60
9	LS	3 Years LEP	618	708.45	708.5	45.25	54
9	LS	4 Years LEP	309	723.16	722	50.32	52
9	LS	5 Years LEP	234	733.36	736	39.29	50
9	LS	6 Years LEP or More	985	729.33	736	49.18	63
9	LS	Arabic	438	688.21	682	56.51	77
9	LS	Bengali	380	682.36	678	53.60	68
9	LS	Chinese	493	677.05	675	52.65	64
9	LS	English	391	705.84	711	54.38	80
9	LS	Spanish	8,194	700.71	706	58.86	89
9	LS	Other Language	4,560	688.47	685	55.28	77
9	LS	Students with Disabilities	2,424	715.50	716	47.24	60

G.10b: Grade 9 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
9	RW	Female	6,166	686.79	686	36.48	51
9	RW	Male	7,290	686.87	686	37.37	51
9	RW	American Indian or Alaska Native	24	681.54	685	40.34	57
9	RW	Black or African American	1,209	686.51	683	37.50	51
9	RW	Asian	2,904	683.58	680	38.40	53
9	RW	Native Hawaiian/Other Pacific Islander	6	687.33	698.5	46.33	81
9	RW	Hispanic or Latino	7,980	685.89	686	34.88	51
9	RW	White	1,250	703.05	700	41.02	59
9	RW	Multiracial (not Hispanic origin)	59	675.88	670	34.60	45
9	RW	NYC	8,508	688.50	686	36.71	51
9	RW	Big 4 Cities	853	678.59	673	39.25	56
9	RW	High Need Urban/Suburban	1,594	683.34	683	36.52	57
9	RW	High Need Rural	70	683.39	683	33.16	50
9	RW	Average Need	985	688.25	690	35.55	48
9	RW	Low Need	516	695.22	693	39.29	57
9	RW	Charter Schools	48	706.08	712	30.62	36
9	RW	0 Years LEP	7,612	686.81	686	36.18	51
9	RW	1 Years LEP	2,352	674.39	670	35.91	49
9	RW	2 Years LEP	1,381	686.13	683	37.22	51
9	RW	3 Years LEP	601	693.67	693	35.57	47
9	RW	4 Years LEP	300	692.11	693	33.34	50
9	RW	5 Years LEP	233	699.33	704	32.19	44
9	RW	6 Years LEP or More	964	709.27	708	35.52	46
9	RW	Arabic	393	676.01	670	37.04	52
9	RW	Bengali	334	686.02	683	36.24	47
9	RW	Chinese	451	684.42	683	35.34	43
9	RW	English	369	690.78	686	36.04	53
9	RW	Spanish	7,733	685.81	686	35.00	51
9	RW	Other Language	4,192	689.72	686	40.40	56
9	RW	Students with Disabilities	2,305	691.29	690	36.44	45

G.11a: Grade 10 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
10	LS	Female	6,256	702.42	701	53.76	71
10	LS	Male	6,762	703.59	706	54.65	71
10	LS	American Indian or Alaska Native	28	704.82	711	52.77	69
10	LS	Black or African American	1,067	707.02	706	44.58	61
10	LS	Asian	3,509	688.09	685	49.86	63
10	LS	Native Hawaiian/Other Pacific Islander	2	719.00	719	4.24	6
10	LS	Hispanic or Latino	7,195	708.29	711	56.51	77
10	LS	White	1,193	711.39	711	52.47	70
10	LS	Multiracial (not Hispanic origin)	33	695.97	701	54.08	63
10	LS	NYC	8,789	702.81	701	54.67	71
10	LS	Big 4 Cities	570	708.60	701	58.53	77
10	LS	High Need Urban/Suburban	1,311	703.54	701	50.58	65
10	LS	High Need Rural	56	708.32	711	49.80	66.5
10	LS	Average Need	889	710.84	711	52.93	70
10	LS	Low Need	432	713.20	711	52.06	70
10	LS	Charter Schools	44	744.75	736	37.91	50
10	LS	0 Years LEP	7,140	704.71	706	56.88	80
10	LS	1 Years LEP	2,438	684.20	682	47.70	58
10	LS	2 Years LEP	1,617	700.28	697	46.94	60
10	LS	3 Years LEP	556	706.72	706	46.04	61
10	LS	4 Years LEP	343	726.87	722	48.27	63
10	LS	5 Years LEP	274	721.82	722	46.60	60
10	LS	6 Years LEP or More	639	738.11	736	49.43	40
10	LS	Arabic	322	710.76	708.5	48.93	67
10	LS	Bengali	294	703.92	699	50.65	68
10	LS	Chinese	649	676.97	675	48.23	61
10	LS	English	306	721.58	722	48.83	52
10	LS	Spanish	7,028	707.92	711	56.47	77
10	LS	Other Language	4,433	697.08	693	50.49	66
10	LS	Students with Disabilities	1,751	720.78	722	49.21	52

G.11b: Grade 10 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
10	RW	Female	6,094	694.30	693	37.50	54
10	RW	Male	6,419	692.58	693	36.92	50
10	RW	American Indian or Alaska Native	25	706.08	704	34.03	46
10	RW	Black or African American	1,049	693.58	693	35.63	50
10	RW	Asian	3,329	690.79	686	37.78	53
10	RW	Native Hawaiian/Other Pacific Islander	2	702.50	702.5	7.78	11
10	RW	Hispanic or Latino	6,928	693.13	693	36.01	50
10	RW	White	1,154	702.40	700	42.37	59
10	RW	Multiracial (not Hispanic origin)	30	688.93	688	40.31	40
10	RW	NYC	8,420	693.72	693	36.47	50
10	RW	Big 4 Cities	558	687.52	683	38.54	54
10	RW	High Need Urban/Suburban	1,272	691.05	690	36.12	53
10	RW	High Need Rural	52	694.69	697	36.19	46.5
10	RW	Average Need	874	698.34	700	38.82	48
10	RW	Low Need	424	707.56	708	41.59	60
10	RW	Charter Schools	43	715.37	717	29.01	47
10	RW	0 Years LEP	6,821	693.87	693	36.39	50
10	RW	1 Years LEP	2,333	686.56	683	37.66	54
10	RW	2 Years LEP	1,568	691.84	690	38.42	53
10	RW	3 Years LEP	540	697.78	697	39.85	59
10	RW	4 Years LEP	340	698.46	700	33.17	48
10	RW	5 Years LEP	272	703.64	704	34.32	43
10	RW	6 Years LEP or More	625	706.62	708	36.02	46
10	RW	Arabic	312	689.20	688	34.23	49.5
10	RW	Bengali	281	692.26	686	35.12	45
10	RW	Chinese	608	688.42	684.5	38.03	52.5
10	RW	English	301	703.52	704	34.68	46
10	RW	Spanish	6,771	692.89	693	36.12	50
10	RW	Other Language	4,250	694.65	693	39.10	54
10	RW	Students with Disabilities	1,691	695.87	697	38.79	51

G.12a: Grade 11 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
11	LS	Female	4,561	716.10	716	51.56	63
11	LS	Male	4,626	718.40	716	51.02	60
11	LS	American Indian or Alaska Native	17	715.06	722	51.20	67
11	LS	Black or African American	850	715.48	716	43.78	56
11	LS	Asian	2,341	709.06	706	47.37	61
11	LS	Native Hawaiian/Other Pacific Islander	8	734.13	741.5	48.90	85
11	LS	Hispanic or Latino	5,096	719.86	722	53.18	71
11	LS	White	865	725.93	728	54.37	67
11	LS	Multiracial (not Hispanic origin)	17	697.94	685	47.77	74
11	LS	NYC	6,190	716.77	716	50.97	63
11	LS	Big 4 Cities	410	719.43	716	57.44	78
11	LS	High Need Urban/Suburban	977	711.04	711	51.76	67
11	LS	High Need Rural	45	730.36	728	48.01	71
11	LS	Average Need	747	723.55	722	48.76	67
11	LS	Low Need	407	722.61	716	49.00	63
11	LS	Charter Schools	26	763.50	756	51.05	61
11	LS	0 Years LEP	4,579	722.42	722	52.82	67
11	LS	1 Years LEP	1,592	695.98	693	48.49	60
11	LS	2 Years LEP	1,365	709.48	706	45.81	58
11	LS	3 Years LEP	583	717.72	716	45.40	60
11	LS	4 Years LEP	294	729.02	722	45.22	55
11	LS	5 Years LEP	273	728.77	728	45.82	59
11	LS	6 Years LEP or More	494	743.03	745	46.06	56
11	LS	Arabic	209	723.16	722	47.24	71
11	LS	Bengali	231	715.28	711	42.93	51
11	LS	Chinese	347	702.62	701	42.82	53
11	LS	English	220	732.74	736	49.33	57
11	LS	Spanish	4,935	719.68	722	53.22	71
11	LS	Other Language	3,266	713.66	711	49.46	63
11	LS	Students with Disabilities	1,219	723.81	722	50.43	63

G.12b: Grade 11 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
11	RW	Female	4,496	705.42	704	36.77	46
11	RW	Male	4,468	705.69	704	36.66	52
11	RW	American Indian or Alaska Native	17	708.41	697	52.31	48
11	RW	Black or African American	834	702.40	700	35.18	49
11	RW	Asian	2,294	706.69	704	37.51	52
11	RW	Native Hawaiian/Other Pacific Islander	8	725.38	723.5	32.22	41
11	RW	Hispanic or Latino	4,962	704.28	704	35.79	46
11	RW	White	840	713.11	712	39.86	51
11	RW	Multiracial (not Hispanic origin)	15	705.47	690	54.01	65
11	RW	NYC	6,023	705.41	704	36.07	46
11	RW	Big 4 Cities	398	696.98	697	37.21	56
11	RW	High Need Urban/Suburban	959	700.90	700	36.13	49
11	RW	High Need Rural	43	712.30	712	38.12	51
11	RW	Average Need	740	710.53	712	36.07	46
11	RW	Low Need	400	711.76	712	37.31	51
11	RW	Charter Schools	26	726.00	723.5	30.02	32
11	RW	0 Years LEP	4,438	706.79	708	35.91	49
11	RW	1 Years LEP	1,554	701.19	697	41.12	56
11	RW	2 Years LEP	1,346	702.02	700	36.10	49
11	RW	3 Years LEP	573	706.80	708	35.08	49
11	RW	4 Years LEP	293	704.68	704	35.87	52
11	RW	5 Years LEP	269	709.97	712	32.26	42
11	RW	6 Years LEP or More	484	712.94	712	32.67	42
11	RW	Arabic	201	698.58	697	34.09	44
11	RW	Bengali	224	702.70	700	32.04	43.5
11	RW	Chinese	342	709.70	712	41.02	57
11	RW	English	215	713.26	712	38.11	47
11	RW	Spanish	4,806	704.30	704	35.97	46
11	RW	Other Language	3,196	707.07	708	37.59	52
11	RW	Students with Disabilities	1,168	705.30	704	40.38	46

G.13a: Grade 12 Listening and Speaking

Grade	Test	Group	N-count	Mean	Median	SD	IQR
12	LS	Female	3,204	732.02	728	44.87	55
12	LS	Male	2,878	732.10	728	46.37	55
12	LS	American Indian or Alaska Native	11	741.45	745	47.69	50
12	LS	Black or African American	564	732.91	728	37.60	45
12	LS	Asian	1,568	724.47	722	43.42	48
12	LS	Native Hawaiian/Other Pacific Islander	4	742.75	744	33.84	58.5
12	LS	Hispanic or Latino	3,342	734.87	736	46.72	50
12	LS	White	591	734.92	736	49.32	50
12	LS	Multiracial (not Hispanic origin)	9	737.22	706	52.24	71
12	LS	NYC	4,313	731.00	728	45.74	55
12	LS	Big 4 Cities	253	731.55	728	51.02	59
12	LS	High Need Urban/Suburban	552	730.57	728	41.83	50
12	LS	High Need Rural	31	744.16	736	40.91	40
12	LS	Average Need	490	735.86	736	48.42	45
12	LS	Low Need	321	738.32	736	40.96	61
12	LS	Charter Schools	17	745.71	745	45.83	56
12	LS	0 Years LEP	3,319	735.47	736	45.65	50
12	LS	1 Years LEP	622	716.56	716	47.15	60
12	LS	2 Years LEP	885	723.46	722	42.74	52
12	LS	3 Years LEP	263	731.75	728	42.16	45
12	LS	4 Years LEP	372	732.09	728	41.09	50
12	LS	5 Years LEP	229	740.94	736	38.12	40
12	LS	6 Years LEP or More	392	741.53	745	50.00	56
12	LS	Arabic	145	739.26	736	41.36	45
12	LS	Bengali	170	721.97	716	42.89	48
12	LS	Chinese	225	715.46	711	46.27	60
12	LS	English	149	750.51	745	45.27	50
12	LS	Spanish	3,262	734.55	736	46.60	50
12	LS	Other Language	2,145	728.88	728	43.53	55
12	LS	Students with Disabilities	749	723.88	722	45.69	44

G.13b: Grade 12 Reading and Writing

Grade	Test	Group	N-count	Mean	Median	SD	IQR
12	RW	Female	3,185	713.55	712	33.51	39
12	RW	Male	2,818	710.07	712	34.31	46
12	RW	American Indian or Alaska Native	11	716.36	717	48.42	60
12	RW	Black or African American	559	708.88	708	31.44	46
12	RW	Asian	1,556	713.84	712	35.84	47
12	RW	Native Hawaiian/Other Pacific Islander	4	728.25	726	32.83	56.5
12	RW	Hispanic or Latino	3,293	711.00	712	32.26	42
12	RW	White	578	714.70	717	39.68	54
12	RW	Multiracial (not Hispanic origin)	9	708.11	717	23.07	17
12	RW	NYC	4,251	711.34	712	33.26	42
12	RW	Big 4 Cities	250	704.98	704	35.38	49
12	RW	High Need Urban/Suburban	544	709.20	712	32.40	46
12	RW	High Need Rural	30	717.20	721	24.70	28
12	RW	Average Need	484	714.48	717	36.99	40
12	RW	Low Need	320	725.99	726	37.54	42
12	RW	Charter Schools	17	718.47	708	40.79	37
12	RW	0 Years LEP	3,266	711.19	712	33.07	42
12	RW	1 Years LEP	618	712.57	712	35.44	47
12	RW	2 Years LEP	878	712.77	712	35.00	47
12	RW	3 Years LEP	261	718.95	717	37.42	40
12	RW	4 Years LEP	371	709.92	708	32.42	42
12	RW	5 Years LEP	225	713.85	717	32.15	39
12	RW	6 Years LEP or More	383	711.61	717	35.60	44
12	RW	Arabic	146	704.95	704	33.76	46
12	RW	Bengali	167	699.14	697	31.99	37
12	RW	Chinese	222	713.54	712	36.12	47
12	RW	English	148	720.12	721	34.20	45.5
12	RW	Spanish	3,214	710.84	712	32.13	42
12	RW	Other Language	2,119	714.28	712	36.13	47
12	RW	Students with Disabilities	718	695.88	697	35.52	48

APPENDIX H: PROFICIENCY PERCENTAGES

H.1a: Kindergarten Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
K	LS	Female	12,578	6.02	24.71	35.87	33.40
K	LS	Male	13,289	4.91	27.90	35.59	31.61
K	LS	American Indian or Alaska Native	64	0.00	25.00	39.06	35.94
K	LS	Black or African American	1,225	5.80	28.00	36.65	29.55
K	LS	Asian	5,376	6.62	29.82	33.82	29.74
K	LS	Native Hawaiian/Other Pacific Islander	19	0.00	21.05	42.11	36.84
K	LS	Hispanic or Latino	15,705	3.37	25.90	37.47	33.26
K	LS	White	3,279	13.75	22.05	30.25	33.94
K	LS	Multiracial (not Hispanic origin)	188	1.60	21.81	42.55	34.04
K	LS	NYC	15,745	5.40	26.64	35.01	32.94
K	LS	Big 4 Cities	1,092	6.59	30.31	35.07	28.02
K	LS	High Need Urban/Suburban	3,431	7.96	30.25	34.83	26.96
K	LS	High Need Rural	164	4.88	27.44	39.63	28.05
K	LS	Average Need	3,034	3.99	21.62	38.33	36.06
K	LS	Low Need	1,358	2.28	23.27	41.24	33.21
K	LS	Charter Schools	277	1.81	17.69	34.30	46.21
K	LS	0 Years LEP	15,181	6.51	26.91	33.50	33.08
K	LS	1 Years LEP	10,675	3.92	25.60	38.91	31.57
K	LS	Arabic	587	8.18	30.83	31.69	29.30
K	LS	Bengali	682	3.96	24.78	39.00	32.26
K	LS	Chinese	1,250	6.16	33.12	32.16	28.56
K	LS	English	1,569	1.91	19.57	40.98	37.54
K	LS	Spanish	14,826	3.51	26.51	37.31	32.66
K	LS	Other Language	7,013	10.18	26.18	31.86	31.78
K	LS	Students with Disabilities	4,606	12.70	35.45	34.00	17.85

H.1b: Kindergarten Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
K	RW	Female	12,663	43.28	27.58	12.86	16.28
K	RW	Male	13,506	49.76	25.97	11.04	13.24
K	RW	American Indian or Alaska Native	65	35.38	32.31	13.85	18.46
K	RW	Black or African American	1,225	56.08	25.06	7.59	11.27
K	RW	Asian	5,585	32.41	26.12	15.95	25.51
K	RW	Native Hawaiian/Other Pacific Islander	19	31.58	31.58	5.26	31.58
K	RW	Hispanic or Latino	15,925	50.08	28.08	11.05	10.79
K	RW	White	3,149	51.32	21.59	10.73	16.35
K	RW	Multiracial (not Hispanic origin)	189	39.68	28.04	15.87	16.40
K	RW	NYC	15,903	43.64	26.06	12.83	17.46
K	RW	Big 4 Cities	1,116	63.80	23.57	6.99	5.65
K	RW	High Need Urban/Suburban	3,542	56.15	25.83	9.15	8.87
K	RW	High Need Rural	169	56.80	33.73	5.92	3.55
K	RW	Average Need	3,034	46.04	31.31	11.44	11.21
K	RW	Low Need	1,369	47.33	26.88	12.34	13.44
K	RW	Charter Schools	273	32.60	35.16	15.02	17.22
K	RW	0 Years LEP	15,386	43.68	26.22	12.89	17.20
K	RW	1 Years LEP	10,769	50.84	27.46	10.54	11.16
K	RW	Arabic	604	52.98	25.00	10.93	11.09
K	RW	Bengali	696	31.32	23.13	15.66	29.89
K	RW	Chinese	1,291	26.34	29.36	18.82	25.48
K	RW	English	1,567	48.31	29.42	10.59	11.68
K	RW	Spanish	15,046	49.97	27.91	11.14	10.99
K	RW	Other Language	7,021	43.98	23.61	12.29	20.11
K	RW	Students with Disabilities	4,382	65.29	20.10	7.83	6.78

H.1c: Kindergarten Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
K	Overall	Female	12,280	42.37	30.76	15.95	10.92
K	Overall	Male	13,087	48.90	29.30	12.84	8.97
K	Overall	American Indian or Alaska Native	64	34.38	34.38	21.88	9.38
K	Overall	Black or African American	1,191	55.00	27.29	8.65	9.07
K	Overall	Asian	5,351	31.58	32.31	20.65	15.46
K	Overall	Native Hawaiian/Other Pacific Islander	19	31.58	31.58	21.05	15.79
K	Overall	Hispanic or Latino	15,556	49.37	30.46	12.65	7.52
K	Overall	White	2,984	49.06	24.56	13.87	12.50
K	Overall	Multiracial (not Hispanic origin)	188	39.89	30.85	17.55	11.70
K	Overall	NYC	15,373	42.72	30.03	15.61	11.64
K	Overall	Big 4 Cities	1,072	62.69	25.56	8.40	3.36
K	Overall	High Need Urban/Suburban	3,386	54.78	28.44	10.87	5.91
K	Overall	High Need Rural	163	55.21	35.58	6.75	2.45
K	Overall	Average Need	2,996	45.79	32.71	13.08	8.41
K	Overall	Low Need	1,351	46.93	29.39	15.10	8.59
K	Overall	Charter Schools	273	32.60	37.00	16.48	13.92
K	Overall	0 Years LEP	14,798	42.50	30.27	15.55	11.68
K	Overall	1 Years LEP	10,555	50.30	29.60	12.68	7.43
K	Overall	Arabic	578	51.21	27.51	11.94	9.34
K	Overall	Bengali	679	30.93	27.54	24.89	16.64
K	Overall	Chinese	1,245	25.62	38.31	21.77	14.30
K	Overall	English	1,556	48.01	30.40	13.30	8.29
K	Overall	Spanish	14,681	49.21	30.38	12.79	7.62
K	Overall	Other Language	6,681	42.55	27.93	15.70	13.82
K	Overall	Students with Disabilities	4,210	64.54	22.99	8.95	3.52

H.2a: Grade 1 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
1	LS	Female	13,075	2.40	12.28	46.54	38.78
1	LS	Male	14,307	3.59	12.41	46.70	37.30
1	LS	American Indian or Alaska Native	44	4.55	13.64	31.82	50.00
1	LS	Black or African American	1,277	4.39	14.49	45.11	36.02
1	LS	Asian	5,278	3.92	12.18	42.99	40.91
1	LS	Native Hawaiian/Other Pacific Islander	16	0.00	12.50	37.50	50.00
1	LS	Hispanic or Latino	16,982	1.70	10.92	49.40	37.98
1	LS	White	3,633	7.24	18.00	40.13	34.63
1	LS	Multiracial (not Hispanic origin)	126	0.79	7.94	48.41	42.86
1	LS	NYC	16,740	2.70	12.06	46.55	38.69
1	LS	Big 4 Cities	1,233	5.03	16.22	50.04	28.71
1	LS	High Need Urban/Suburban	4,055	5.18	18.32	47.77	28.73
1	LS	High Need Rural	190	0.53	7.89	42.63	48.95
1	LS	Average Need	2,985	2.61	8.14	45.09	44.15
1	LS	Low Need	1,306	1.23	7.50	44.10	47.17
1	LS	Charter Schools	242	0.41	4.13	52.89	42.56
1	LS	0 Years LEP	14,618	3.57	13.20	44.92	38.32
1	LS	1 Years LEP	4,379	5.71	16.60	48.71	28.98
1	LS	2 Years LEP	8,423	0.75	8.92	48.44	41.90
1	LS	Arabic	646	7.12	14.40	43.19	35.29
1	LS	Bengali	716	2.65	10.34	44.69	42.32
1	LS	Chinese	1,123	2.58	9.53	43.81	44.08
1	LS	English	1,249	1.04	7.93	46.68	44.36
1	LS	Spanish	16,057	1.73	11.11	49.51	37.65
1	LS	Other Language	7,681	5.88	16.35	41.44	36.32
1	LS	Students with Disabilities	5,541	5.59	23.97	50.88	19.56

H.2b: Grade 1 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
1	RW	Female	13,311	14.60	18.26	28.31	38.83
1	RW	Male	14,533	21.56	18.45	27.59	32.39
1	RW	American Indian or Alaska Native	45	22.22	22.22	24.44	31.11
1	RW	Black or African American	1,308	24.92	21.87	26.61	26.61
1	RW	Asian	5,446	11.72	12.67	25.50	50.11
1	RW	Native Hawaiian/Other Pacific Islander	16	12.50	18.75	12.50	56.25
1	RW	Hispanic or Latino	17,233	19.53	19.89	27.85	32.74
1	RW	White	3,638	18.69	18.64	32.88	29.80
1	RW	Multiracial (not Hispanic origin)	129	14.73	13.95	27.91	43.41
1	RW	NYC	17,116	18.38	17.36	27.58	36.67
1	RW	Big 4 Cities	1,259	31.14	22.08	23.19	23.59
1	RW	High Need Urban/Suburban	4,091	20.04	21.58	28.97	29.41
1	RW	High Need Rural	194	15.46	19.07	32.99	32.47
1	RW	Average Need	2,996	13.65	18.46	28.57	39.32
1	RW	Low Need	1,311	9.38	16.17	31.73	42.72
1	RW	Charter Schools	242	9.50	20.66	33.06	36.78
1	RW	0 Years LEP	14,992	20.10	17.47	26.32	36.11
1	RW	1 Years LEP	4,449	22.95	18.50	30.64	27.92
1	RW	2 Years LEP	8,440	12.52	19.83	29.47	38.18
1	RW	Arabic	675	29.19	16.89	22.96	30.96
1	RW	Bengali	724	10.36	11.60	26.93	51.10
1	RW	Chinese	1,150	7.30	10.00	27.04	55.65
1	RW	English	1,251	13.91	18.23	27.90	39.97
1	RW	Spanish	16,310	19.84	19.96	27.80	32.40
1	RW	Other Language	7,824	17.09	17.05	28.95	36.91
1	RW	Students with Disabilities	5,513	29.33	23.27	28.77	18.63

H.2c: Grade 1 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
1	Overall	Female	13,030	13.97	21.73	40.34	23.96
1	Overall	Male	14,218	20.59	20.88	37.90	20.63
1	Overall	American Indian or Alaska Native	44	22.73	22.73	29.55	25.00
1	Overall	Black or African American	1,269	23.96	24.19	35.30	16.55
1	Overall	Asian	5,266	10.39	16.07	43.13	30.42
1	Overall	Native Hawaiian/Other Pacific Islander	16	12.50	18.75	31.25	37.50
1	Overall	Hispanic or Latino	16,944	18.63	22.16	38.93	20.27
1	Overall	White	3,551	18.98	24.25	35.48	21.29
1	Overall	Multiracial (not Hispanic origin)	126	12.70	16.67	41.27	29.37
1	Overall	NYC	16,674	17.18	20.13	39.59	23.10
1	Overall	Big 4 Cities	1,226	30.10	24.31	32.06	13.54
1	Overall	High Need Urban/Suburban	4,015	20.55	26.43	37.04	15.99
1	Overall	High Need Rural	190	13.68	22.11	38.42	25.79
1	Overall	Average Need	2,963	13.20	20.35	40.70	25.75
1	Overall	Low Need	1,303	9.21	18.96	42.75	29.09
1	Overall	Charter Schools	242	9.50	21.49	45.87	23.14
1	Overall	0 Years LEP	14,543	18.83	20.38	37.98	22.82
1	Overall	1 Years LEP	4,327	22.16	22.90	38.90	16.04
1	Overall	2 Years LEP	8,409	12.65	22.08	41.00	24.26
1	Overall	Arabic	643	26.44	18.66	34.37	20.53
1	Overall	Bengali	714	10.22	13.59	44.54	31.65
1	Overall	Chinese	1,119	5.99	13.40	47.27	33.33
1	Overall	English	1,247	13.79	20.37	39.21	26.62
1	Overall	Spanish	16,020	18.88	22.28	38.85	19.99
1	Overall	Other Language	7,588	16.67	21.60	38.09	23.64
1	Overall	Students with Disabilities	5,429	29.97	28.73	32.79	8.51

H.3a: Grade 2 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
2	LS	Female	11,037	0.99	6.41	32.89	59.71
2	LS	Male	12,476	1.39	7.81	34.44	56.36
2	LS	American Indian or Alaska Native	53	1.89	1.89	30.19	66.04
2	LS	Black or African American	1,259	1.99	8.58	34.31	55.12
2	LS	Asian	4,144	2.00	8.74	28.84	60.42
2	LS	Native Hawaiian/Other Pacific Islander	16	0.00	12.50	18.75	68.75
2	LS	Hispanic or Latino	14,929	0.66	4.80	33.61	60.93
2	LS	White	3,064	2.42	14.75	41.38	41.45
2	LS	Multiracial (not Hispanic origin)	59	1.69	10.17	32.20	55.93
2	LS	NYC	14,874	1.06	6.74	35.29	56.92
2	LS	Big 4 Cities	1,236	2.59	11.00	33.01	53.40
2	LS	High Need Urban/Suburban	3,445	2.00	10.54	36.17	51.29
2	LS	High Need Rural	134	0.75	5.22	29.10	64.93
2	LS	Average Need	2,304	0.91	5.08	25.82	68.19
2	LS	Low Need	1,040	0.19	2.98	26.15	70.67
2	LS	Charter Schools	202	0.00	0.99	23.27	75.74
2	LS	0 Years LEP	12,771	1.14	7.26	33.32	58.28
2	LS	1 Years LEP	2,440	4.47	18.61	34.71	42.21
2	LS	2 Years LEP	2,827	0.81	7.32	41.88	49.98
2	LS	3 Years LEP	5,544	0.05	2.02	30.28	67.64
2	LS	Arabic	588	2.38	11.56	28.40	57.65
2	LS	Bengali	601	0.50	7.65	29.78	62.06
2	LS	Chinese	827	1.21	6.17	29.38	63.24
2	LS	English	958	0.63	2.82	29.85	66.70
2	LS	Spanish	14,201	0.70	4.94	33.93	60.43
2	LS	Other Language	6,460	2.32	12.57	35.62	49.49
2	LS	Students with Disabilities	5,778	1.32	12.53	50.12	36.03

H.3b: Grade 2 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
2	RW	Female	10,606	11.31	37.19	34.29	17.21
2	RW	Male	11,502	15.18	38.27	31.42	15.13
2	RW	American Indian or Alaska Native	47	4.26	34.04	38.30	23.40
2	RW	Black or African American	1,155	17.84	42.25	28.83	11.08
2	RW	Asian	3,911	9.28	28.48	36.72	25.52
2	RW	Native Hawaiian/Other Pacific Islander	16	18.75	12.50	43.75	25.00
2	RW	Hispanic or Latino	14,176	14.10	41.53	32.44	11.94
2	RW	White	2,753	11.81	30.51	30.51	27.17
2	RW	Multiracial (not Hispanic origin)	56	16.07	35.71	28.57	19.64
2	RW	NYC	13,949	13.32	39.06	31.94	15.69
2	RW	Big 4 Cities	1,100	20.18	40.82	28.82	10.18
2	RW	High Need Urban/Suburban	3,223	14.18	38.01	34.35	13.47
2	RW	High Need Rural	124	8.06	38.71	33.87	19.35
2	RW	Average Need	2,229	12.20	32.44	34.99	20.37
2	RW	Low Need	1,016	6.40	26.38	36.52	30.71
2	RW	Charter Schools	199	4.52	39.70	39.20	16.58
2	RW	0 Years LEP	12,013	14.03	39.81	32.62	13.54
2	RW	1 Years LEP	2,051	20.62	37.54	27.21	14.63
2	RW	2 Years LEP	2,697	11.94	35.26	30.03	22.77
2	RW	3 Years LEP	5,416	9.66	34.73	36.39	19.22
2	RW	Arabic	528	18.75	34.85	32.58	13.83
2	RW	Bengali	581	9.29	31.84	34.94	23.92
2	RW	Chinese	802	7.11	25.06	38.78	29.05
2	RW	English	916	11.57	33.30	35.15	19.98
2	RW	Spanish	13,460	14.30	41.82	32.28	11.60
2	RW	Other Language	5,937	12.09	31.85	32.24	23.82
2	RW	Students with Disabilities	5,075	23.25	43.45	21.52	11.78

H.3c: Grade 2 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
2	Overall Female		10,533	11.01	37.80	36.29	14.91
2	Overall Male		11,420	15.01	38.34	33.59	13.06
2	Overall American Indian or Alaska Native		47	4.26	34.04	38.30	23.40
2	Overall Black or African American		1,143	17.50	42.26	29.40	10.85
2	Overall Asian		3,880	8.87	28.66	38.71	23.76
2	Overall Native Hawaiian/Other Pacific Islander		16	18.75	12.50	43.75	25.00
2	Overall Hispanic or Latino		14,100	13.83	41.65	33.10	11.42
2	Overall White		2,714	12.09	32.09	41.45	14.37
2	Overall Multiracial (not Hispanic origin)		56	16.07	35.71	32.14	16.07
2	Overall NYC		13,832	12.96	39.29	34.96	12.80
2	Overall Big 4 Cities		1,092	20.05	40.84	29.12	9.98
2	Overall High Need Urban/Suburban		3,198	14.32	38.62	35.62	11.44
2	Overall High Need Rural		124	8.06	38.71	34.68	18.55
2	Overall Average Need		2,222	12.15	33.39	35.19	19.26
2	Overall Low Need		1,014	6.21	26.43	37.67	29.68
2	Overall Charter Schools		199	4.52	39.70	40.20	15.58
2	Overall 0 Years LEP		11,918	13.63	40.06	33.95	12.37
2	Overall 1 Years LEP		2,020	20.89	38.02	27.28	13.81
2	Overall 2 Years LEP		2,676	11.85	36.36	38.34	13.45
2	Overall 3 Years LEP		5,403	9.61	34.85	37.89	17.66
2	Overall Arabic		525	18.29	35.24	33.33	13.14
2	Overall Bengali		579	9.15	31.78	36.79	22.28
2	Overall Chinese		797	7.15	24.97	41.78	26.10
2	Overall English		914	11.49	33.37	35.34	19.80
2	Overall Spanish		13,387	14.03	41.94	32.96	11.06
2	Overall Other Language		5,862	11.91	32.77	38.21	17.11
2	Overall Students with Disabilities		5,021	23.46	44.63	27.33	4.58

H.4a: Grade 3 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
3	LS	Female	10,198	1.58	5.24	32.03	61.16
3	LS	Male	11,628	2.92	7.32	31.53	58.23
3	LS	American Indian or Alaska Native	32	0.00	0.00	31.25	68.75
3	LS	Black or African American	1,165	2.15	10.21	36.91	50.73
3	LS	Asian	3,668	3.11	8.78	29.69	58.42
3	LS	Native Hawaiian/Other Pacific Islander	8	0.00	0.00	12.50	87.50
3	LS	Hispanic or Latino	14,029	1.23	3.66	30.96	64.15
3	LS	White	2,875	6.19	13.95	36.80	43.06
3	LS	Multiracial (not Hispanic origin)	29	3.45	13.79	34.48	48.28
3	LS	NYC	14,022	1.84	6.27	31.41	60.48
3	LS	Big 4 Cities	1,234	2.59	8.35	36.95	52.11
3	LS	High Need Urban/Suburban	3,206	4.09	8.05	35.09	52.78
3	LS	High Need Rural	132	3.03	2.27	31.82	62.88
3	LS	Average Need	2,057	3.01	4.96	26.54	65.48
3	LS	Low Need	854	1.29	3.16	28.57	66.98
3	LS	Charter Schools	130	0.00	0.77	26.92	72.31
3	LS	0 Years LEP	12,497	1.96	6.31	29.66	62.07
3	LS	1 Years LEP	1,829	10.50	16.95	38.71	33.84
3	LS	2 Years LEP	1,759	1.99	7.11	39.28	51.62
3	LS	3 Years LEP	1,389	1.37	6.77	43.05	48.81
3	LS	4 Years LEP	4,373	0.25	1.76	28.58	69.40
3	LS	Arabic	561	4.46	12.30	31.02	52.23
3	LS	Bengali	578	1.90	5.54	26.64	65.92
3	LS	Chinese	608	2.63	11.02	28.78	57.57
3	LS	English	794	0.50	1.39	29.35	68.77
3	LS	Spanish	13,386	1.23	3.80	31.23	63.74
3	LS	Other Language	5,962	4.71	11.88	34.38	49.03
3	LS	Students with Disabilities	5,790	3.01	9.69	46.48	40.83

H.4b: Grade 3 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
3	RW	Female	9,968	7.44	20.58	49.19	22.79
3	RW	Male	11,018	11.46	23.10	45.14	20.29
3	RW	American Indian or Alaska Native	32	6.25	9.38	59.38	25.00
3	RW	Black or African American	1,098	16.03	25.59	42.99	15.39
3	RW	Asian	3,521	9.94	16.30	44.93	28.83
3	RW	Native Hawaiian/Other Pacific Islander	8	0.00	0.00	50.00	50.00
3	RW	Hispanic or Latino	13,659	8.46	23.28	48.78	19.47
3	RW	White	2,623	10.83	21.35	43.16	24.67
3	RW	Multiracial (not Hispanic origin)	27	22.22	11.11	37.04	29.63
3	RW	NYC	13,487	9.18	22.46	47.38	20.98
3	RW	Big 4 Cities	1,162	16.87	25.13	41.48	16.52
3	RW	High Need Urban/Suburban	3,056	9.55	21.40	48.56	20.48
3	RW	High Need Rural	130	8.46	25.38	49.23	16.92
3	RW	Average Need	1,989	8.85	18.85	46.46	25.84
3	RW	Low Need	846	6.50	15.37	46.45	31.68
3	RW	Charter Schools	127	3.94	18.90	56.69	20.47
3	RW	0 Years LEP	12,090	9.63	22.42	47.68	20.28
3	RW	1 Years LEP	1,576	21.13	26.65	36.23	15.99
3	RW	2 Years LEP	1,693	10.51	22.80	43.89	22.80
3	RW	3 Years LEP	1,330	9.77	21.73	43.91	24.59
3	RW	4 Years LEP	4,309	4.64	18.29	51.75	25.32
3	RW	Arabic	506	17.39	22.73	44.47	15.42
3	RW	Bengali	560	6.07	16.25	48.39	29.29
3	RW	Chinese	581	9.98	15.32	43.37	31.33
3	RW	English	775	5.81	20.13	45.68	28.39
3	RW	Spanish	13,027	8.64	23.44	48.66	19.26
3	RW	Other Language	5,591	11.82	19.71	44.04	24.43
3	RW	Students with Disabilities	5,361	17.07	33.26	38.52	11.15

H.4c: Grade 3 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
3	Overall	Female	9,899	7.18	21.02	51.73	20.06
3	Overall	Male	10,953	11.30	23.39	47.16	18.15
3	Overall	American Indian or Alaska Native	32	6.25	9.38	68.75	15.63
3	Overall	Black or African American	1,094	15.90	25.78	44.24	14.08
3	Overall	Asian	3,481	9.11	16.60	48.00	26.29
3	Overall	Native Hawaiian/Other Pacific Islander	8	0.00	0.00	62.50	37.50
3	Overall	Hispanic or Latino	13,599	8.21	23.35	50.46	17.97
3	Overall	White	2,593	11.65	23.33	47.71	17.32
3	Overall	Multiracial (not Hispanic origin)	27	22.22	11.11	37.04	29.63
3	Overall	NYC	13,393	8.77	22.79	49.78	18.66
3	Overall	Big 4 Cities	1,160	16.81	25.26	42.76	15.17
3	Overall	High Need Urban/Suburban	3,028	10.11	22.13	50.46	17.31
3	Overall	High Need Rural	130	9.23	24.62	51.54	14.62
3	Overall	Average Need	1,983	8.72	19.31	48.56	23.40
3	Overall	Low Need	843	6.17	15.66	49.70	28.47
3	Overall	Charter Schools	127	3.94	18.90	56.69	20.47
3	Overall	0 Years LEP	11,998	9.18	22.69	49.46	18.68
3	Overall	1 Years LEP	1,556	21.53	26.80	37.79	13.88
3	Overall	2 Years LEP	1,682	10.64	24.14	46.37	18.85
3	Overall	3 Years LEP	1,323	10.05	23.05	50.34	16.55
3	Overall	4 Years LEP	4,305	4.67	18.40	54.17	22.76
3	Overall	Arabic	504	17.06	23.02	45.04	14.88
3	Overall	Bengali	560	6.25	16.43	50.18	27.14
3	Overall	Chinese	573	9.08	15.71	46.42	28.80
3	Overall	English	775	5.81	20.13	47.87	26.19
3	Overall	Spanish	12,968	8.39	23.50	50.33	17.77
3	Overall	Other Language	5,526	11.76	20.87	47.79	19.58
3	Overall	Students with Disabilities	5,331	17.54	34.40	41.27	6.79

H.5a: Grade 4 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
4	LS	Female	8,861	2.48	6.31	28.39	62.81
4	LS	Male	10,896	4.29	7.98	26.62	61.10
4	LS	American Indian or Alaska Native	20	0.00	0.00	15.00	85.00
4	LS	Black or African American	1,134	3.44	9.70	31.04	55.82
4	LS	Asian	3,240	5.52	11.57	26.60	56.30
4	LS	Native Hawaiian/Other Pacific Islander	9	0.00	22.22	33.33	44.44
4	LS	Hispanic or Latino	12,703	1.78	4.16	26.11	67.95
4	LS	White	2,634	8.69	14.58	34.47	42.26
4	LS	Multiracial (not Hispanic origin)	23	8.70	13.04	30.43	47.83
4	LS	NYC	12,909	3.00	6.80	27.13	63.07
4	LS	Big 4 Cities	1,140	4.56	9.47	31.93	54.04
4	LS	High Need Urban/Suburban	3,014	6.30	9.85	31.12	52.72
4	LS	High Need Rural	142	1.41	5.63	29.58	63.38
4	LS	Average Need	1,665	2.64	5.89	23.24	68.23
4	LS	Low Need	710	1.41	5.49	22.96	70.14
4	LS	Charter Schools	99	2.02	1.01	22.22	74.75
4	LS	0 Years LEP	11,384	3.10	6.68	24.79	65.43
4	LS	1 Years LEP	1,831	10.54	17.86	35.77	35.83
4	LS	2 Years LEP	1,211	3.47	8.67	36.75	51.11
4	LS	3 Years LEP	977	7.06	14.74	35.11	43.09
4	LS	4 Years LEP	1,154	2.25	3.55	33.80	60.40
4	LS	5 Years LEP	3,249	0.25	1.97	24.56	73.22
4	LS	Arabic	506	5.93	12.85	28.26	52.96
4	LS	Bengali	423	1.89	10.17	26.95	60.99
4	LS	Chinese	486	5.76	12.14	26.75	55.35
4	LS	English	685	0.88	2.63	24.82	71.68
4	LS	Spanish	12,217	1.86	4.23	26.15	67.76
4	LS	Other Language	5,512	7.13	13.43	30.95	48.49
4	LS	Students with Disabilities	5,771	4.05	8.63	37.76	49.56

H.5b: Grade 4 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
4	RW	Female	8,703	7.19	22.14	42.54	28.13
4	RW	Male	10,526	11.78	24.22	39.65	24.35
4	RW	American Indian or Alaska Native	20	0.00	30.00	40.00	30.00
4	RW	Black or African American	1,083	12.93	31.30	36.47	19.30
4	RW	Asian	3,114	11.50	19.30	37.83	31.37
4	RW	Native Hawaiian/Other Pacific Islander	9	0.00	33.33	33.33	33.33
4	RW	Hispanic or Latino	12,503	7.49	23.47	43.36	25.69
4	RW	White	2,499	16.29	24.01	35.41	24.29
4	RW	Multiracial (not Hispanic origin)	19	10.53	26.32	42.11	21.05
4	RW	NYC	12,600	9.24	23.82	40.93	26.02
4	RW	Big 4 Cities	1,078	14.10	28.66	39.61	17.63
4	RW	High Need Urban/Suburban	2,899	12.07	22.01	40.63	25.28
4	RW	High Need Rural	141	10.64	26.24	39.72	23.40
4	RW	Average Need	1,636	7.89	21.03	42.30	28.79
4	RW	Low Need	704	6.68	15.91	40.77	36.65
4	RW	Charter Schools	99	3.03	21.21	49.49	26.26
4	RW	0 Years LEP	11,113	8.75	23.39	41.71	26.16
4	RW	1 Years LEP	1,697	23.92	29.99	28.99	17.09
4	RW	2 Years LEP	1,181	12.28	25.83	35.73	26.16
4	RW	3 Years LEP	927	18.55	22.87	36.46	22.11
4	RW	4 Years LEP	1,129	5.31	21.52	43.49	29.67
4	RW	5 Years LEP	3,227	3.84	19.34	46.82	30.00
4	RW	Arabic	460	16.52	27.17	34.78	21.52
4	RW	Bengali	411	6.33	20.92	39.42	33.33
4	RW	Chinese	471	12.53	19.11	36.09	32.27
4	RW	English	675	6.07	20.15	42.67	31.11
4	RW	Spanish	12,016	7.51	23.62	43.21	25.66
4	RW	Other Language	5,263	14.73	23.14	36.71	25.42
4	RW	Students with Disabilities	5,541	14.69	34.09	37.65	13.57

H.5c: Grade 4 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
4	Overall	Female	8,624	6.89	22.62	45.56	24.93
4	Overall	Male	10,442	11.58	24.40	42.06	21.96
4	Overall	American Indian or Alaska Native	20	0.00	30.00	40.00	30.00
4	Overall	Black or African American	1,076	12.92	31.23	37.83	18.03
4	Overall	Asian	3,079	11.11	19.36	41.18	28.35
4	Overall	Native Hawaiian/Other Pacific Islander	9	0.00	33.33	44.44	22.22
4	Overall	Hispanic or Latino	12,421	7.11	23.54	45.85	23.50
4	Overall	White	2,461	16.86	26.13	38.81	18.20
4	Overall	Multiracial (not Hispanic origin)	19	10.53	26.32	47.37	15.79
4	Overall	NYC	12,482	8.79	24.07	43.74	23.40
4	Overall	Big 4 Cities	1,073	13.79	29.17	40.54	16.50
4	Overall	High Need Urban/Suburban	2,873	12.53	22.52	44.10	20.85
4	Overall	High Need Rural	141	10.64	26.24	41.13	21.99
4	Overall	Average Need	1,622	7.83	21.64	43.65	26.88
4	Overall	Low Need	703	6.54	16.22	44.38	32.86
4	Overall	Charter Schools	99	3.03	21.21	52.53	23.23
4	Overall	0 Years LEP	10,991	8.21	23.64	44.04	24.12
4	Overall	1 Years LEP	1,681	24.45	29.98	30.70	14.87
4	Overall	2 Years LEP	1,178	12.48	25.98	39.56	21.99
4	Overall	3 Years LEP	912	18.64	24.12	39.36	17.87
4	Overall	4 Years LEP	1,127	5.32	23.07	49.25	22.36
4	Overall	5 Years LEP	3,222	3.91	19.52	49.69	26.88
4	Overall	Arabic	454	15.64	27.75	36.78	19.82
4	Overall	Bengali	410	6.59	20.73	41.95	30.73
4	Overall	Chinese	463	11.66	19.87	38.66	29.81
4	Overall	English	674	6.08	20.18	44.81	28.93
4	Overall	Spanish	11,936	7.13	23.69	45.71	23.47
4	Overall	Other Language	5,195	14.86	24.08	39.96	21.10
4	Overall	Students with Disabilities	5,506	15.04	35.11	39.72	10.13

H.6a: Grade 5 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
5	LS	Female	7,144	3.91	9.38	40.23	46.49
5	LS	Male	9,191	6.08	9.87	37.61	46.44
5	LS	American Indian or Alaska Native	21	0.00	19.05	33.33	47.62
5	LS	Black or African American	1,025	4.20	13.95	41.56	40.29
5	LS	Asian	2,514	8.55	14.92	34.92	41.61
5	LS	Native Hawaiian/Other Pacific Islander	10	0.00	0.00	70.00	30.00
5	LS	Hispanic or Latino	10,449	3.16	6.44	39.09	51.31
5	LS	White	2,278	10.05	16.46	40.87	32.62
5	LS	Multiracial (not Hispanic origin)	21	9.52	23.81	28.57	38.10
5	LS	NYC	10,765	4.40	9.50	38.75	47.35
5	LS	Big 4 Cities	973	5.55	13.16	44.09	37.20
5	LS	High Need Urban/Suburban	2,448	9.56	11.19	38.77	40.48
5	LS	High Need Rural	112	3.57	8.93	42.86	44.64
5	LS	Average Need	1,231	4.47	8.69	35.26	51.58
5	LS	Low Need	506	1.58	7.71	35.77	54.94
5	LS	Charter Schools	171	0.58	0.58	36.26	62.57
5	LS	0 Years LEP	9,572	4.95	8.77	36.87	49.41
5	LS	1 Years LEP	1,691	11.12	23.06	41.45	24.36
5	LS	2 Years LEP	1,120	3.66	10.18	44.82	41.34
5	LS	3 Years LEP	606	0.33	13.70	50.66	35.31
5	LS	4 Years LEP	751	16.25	7.06	35.82	40.88
5	LS	5 Years LEP	767	1.56	8.74	47.59	42.11
5	LS	6 Years LEP or More	1,858	0.11	2.42	36.33	61.14
5	LS	Arabic	428	6.78	18.69	32.01	42.52
5	LS	Bengali	311	4.18	14.79	38.26	42.77
5	LS	Chinese	378	11.64	19.05	34.13	35.19
5	LS	English	510	1.57	4.90	39.80	53.73
5	LS	Spanish	10,018	3.26	6.57	39.03	51.14
5	LS	Other Language	4,736	8.89	15.03	39.15	36.93
5	LS	Students with Disabilities	5,333	4.65	9.28	48.49	37.58

H.6b: Grade 5 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
5	RW	Female	6,910	10.25	24.99	40.07	24.69
5	RW	Male	8,624	14.66	25.56	39.70	20.08
5	RW	American Indian or Alaska Native	20	10.00	35.00	40.00	15.00
5	RW	Black or African American	971	20.49	30.38	33.57	15.55
5	RW	Asian	2,266	13.33	21.84	34.86	29.96
5	RW	Native Hawaiian/Other Pacific Islander	10	10.00	20.00	10.00	60.00
5	RW	Hispanic or Latino	10,130	10.29	26.51	43.13	20.08
5	RW	White	2,102	18.74	21.69	32.87	26.69
5	RW	Multiracial (not Hispanic origin)	20	30.00	30.00	30.00	10.00
5	RW	NYC	10,185	11.31	25.43	40.57	22.69
5	RW	Big 4 Cities	912	20.94	33.22	34.43	11.40
5	RW	High Need Urban/Suburban	2,345	18.46	23.03	38.64	19.87
5	RW	High Need Rural	109	17.43	30.28	39.45	12.84
5	RW	Average Need	1,197	10.44	24.39	40.27	24.90
5	RW	Low Need	492	6.10	21.34	37.40	35.16
5	RW	Charter Schools	169	1.18	24.85	44.97	28.99
5	RW	0 Years LEP	9,096	11.60	24.65	41.14	22.61
5	RW	1 Years LEP	1,505	23.99	31.56	29.24	15.22
5	RW	2 Years LEP	1,075	14.98	27.53	33.49	24.00
5	RW	3 Years LEP	582	10.82	28.69	39.69	20.79
5	RW	4 Years LEP	721	23.44	23.30	33.15	20.11
5	RW	5 Years LEP	756	7.14	24.47	43.39	25.00
5	RW	6 Years LEP or More	1,826	6.35	22.89	46.82	23.93
5	RW	Arabic	386	21.24	29.79	31.87	17.10
5	RW	Bengali	294	14.97	20.75	36.73	27.55
5	RW	Chinese	330	16.36	23.64	30.61	29.39
5	RW	English	500	7.60	23.60	42.40	26.40
5	RW	Spanish	9,706	10.52	26.61	42.95	19.92
5	RW	Other Language	4,360	17.00	22.89	34.13	25.99
5	RW	Students with Disabilities	5,045	17.15	32.86	35.90	14.09

H.6c: Grade 5 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
5	Overall	Female	6,833	9.83	25.80	45.76	18.60
5	Overall	Male	8,565	14.45	25.99	43.76	15.80
5	Overall	American Indian or Alaska Native	20	10.00	35.00	45.00	10.00
5	Overall	Black or African American	959	19.92	30.87	37.12	12.10
5	Overall	Asian	2,242	13.02	22.08	40.81	24.09
5	Overall	Native Hawaiian/Other Pacific Islander	10	10.00	20.00	40.00	30.00
5	Overall	Hispanic or Latino	10,058	9.98	26.67	46.69	16.66
5	Overall	White	2,073	18.72	24.84	42.84	13.60
5	Overall	Multiracial (not Hispanic origin)	20	30.00	30.00	30.00	10.00
5	Overall	NYC	10,087	10.92	25.95	45.97	17.16
5	Overall	Big 4 Cities	905	20.77	33.37	36.24	9.61
5	Overall	High Need Urban/Suburban	2,321	18.31	24.30	42.74	14.65
5	Overall	High Need Rural	109	17.43	30.28	42.20	10.09
5	Overall	Average Need	1,190	10.59	24.96	43.45	21.01
5	Overall	Low Need	490	5.92	21.43	43.47	29.18
5	Overall	Charter Schools	169	1.18	24.85	47.93	26.04
5	Overall	0 Years LEP	8,999	11.15	24.95	46.65	17.26
5	Overall	1 Years LEP	1,480	23.92	32.70	31.35	12.03
5	Overall	2 Years LEP	1,074	14.99	27.93	37.80	19.27
5	Overall	3 Years LEP	580	10.69	31.90	43.28	14.14
5	Overall	4 Years LEP	717	23.57	24.41	36.26	15.76
5	Overall	5 Years LEP	748	6.95	27.01	49.73	16.31
5	Overall	6 Years LEP or More	1,825	6.36	22.85	50.90	19.89
5	Overall	Arabic	382	20.42	30.63	33.25	15.71
5	Overall	Bengali	291	14.43	21.65	39.52	24.40
5	Overall	Chinese	327	16.82	22.94	37.61	22.63
5	Overall	English	498	7.43	24.10	46.59	21.89
5	Overall	Spanish	9,634	10.21	26.76	46.54	16.48
5	Overall	Other Language	4,306	16.77	24.55	41.92	16.77
5	Overall	Students with Disabilities	5,017	17.18	34.32	41.32	7.18

H.7a: Grade 6 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
6	LS	Female	6,327	6.24	14.21	47.72	31.83
6	LS	Male	7,555	9.61	14.02	44.99	31.38
6	LS	American Indian or Alaska Native	27	7.41	11.11	55.56	25.93
6	LS	Black or African American	938	7.46	17.70	47.76	27.08
6	LS	Asian	2,279	15.36	20.10	39.58	24.97
6	LS	Native Hawaiian/Other Pacific Islander	10	0.00	0.00	80.00	20.00
6	LS	Hispanic or Latino	8,595	4.75	9.60	49.81	35.85
6	LS	White	1,984	13.05	25.00	37.95	23.99
6	LS	Multiracial (not Hispanic origin)	30	10.00	30.00	50.00	10.00
6	LS	NYC	9,072	7.82	14.64	46.65	30.90
6	LS	Big 4 Cities	888	8.00	18.36	44.59	29.05
6	LS	High Need Urban/Suburban	1,885	13.10	13.10	42.12	31.67
6	LS	High Need Rural	82	2.44	9.76	52.44	35.37
6	LS	Average Need	1,032	5.91	11.72	43.12	39.24
6	LS	Low Need	401	3.99	11.22	49.63	35.16
6	LS	Charter Schools	136	0.74	2.21	61.03	36.03
6	LS	0 Years LEP	8,062	8.26	13.09	46.55	32.10
6	LS	1 Years LEP	1,701	16.11	28.81	40.15	14.93
6	LS	2 Years LEP	1,111	5.04	16.83	49.50	28.62
6	LS	3 Years LEP	571	0.88	16.99	53.06	29.07
6	LS	4 Years LEP	460	19.35	6.52	41.30	32.83
6	LS	5 Years LEP	470	3.83	8.94	50.85	36.38
6	LS	6 Years LEP or More	1,530	1.44	4.71	45.75	48.10
6	LS	Arabic	389	11.31	17.48	49.61	21.59
6	LS	Bengali	283	6.36	19.43	44.52	29.68
6	LS	Chinese	300	23.67	21.67	39.00	15.67
6	LS	English	386	2.59	7.25	49.74	40.41
6	LS	Spanish	8,291	4.76	9.78	49.84	35.62
6	LS	Other Language	4,276	13.84	22.12	39.01	25.02
6	LS	Students with Disabilities	4,686	6.25	14.06	52.58	27.10

H.7b: Grade 6 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
6	RW	Female	6,154	12.90	27.54	37.46	22.10
6	RW	Male	7,090	18.62	31.54	33.74	16.11
6	RW	American Indian or Alaska Native	26	19.23	23.08	42.31	15.38
6	RW	Black or African American	893	18.81	35.39	30.24	15.57
6	RW	Asian	2,057	22.61	25.81	29.36	22.22
6	RW	Native Hawaiian/Other Pacific Islander	10	0.00	20.00	60.00	20.00
6	RW	Hispanic or Latino	8,404	13.43	31.68	38.20	16.69
6	RW	White	1,801	16.71	22.54	32.82	27.93
6	RW	Multiracial (not Hispanic origin)	29	27.59	37.93	24.14	10.34
6	RW	NYC	8,572	15.90	30.86	34.52	18.72
6	RW	Big 4 Cities	840	20.12	32.74	31.55	15.60
6	RW	High Need Urban/Suburban	1,837	18.78	25.80	36.15	19.27
6	RW	High Need Rural	78	6.41	26.92	46.15	20.51
6	RW	Average Need	1,013	12.73	26.85	39.29	21.13
6	RW	Low Need	388	6.70	22.42	42.01	28.87
6	RW	Charter Schools	135	0.74	22.96	62.96	13.33
6	RW	0 Years LEP	7,705	16.17	30.46	35.28	18.09
6	RW	1 Years LEP	1,490	26.78	34.90	24.83	13.49
6	RW	2 Years LEP	1,077	16.16	30.45	31.66	21.73
6	RW	3 Years LEP	562	10.68	25.44	38.61	25.27
6	RW	4 Years LEP	463	26.57	19.87	31.32	22.25
6	RW	5 Years LEP	462	9.74	22.94	46.32	21.00
6	RW	6 Years LEP or More	1,502	5.13	26.43	46.01	22.44
6	RW	Arabic	340	26.47	28.24	32.06	13.24
6	RW	Bengali	271	19.56	27.31	29.89	23.25
6	RW	Chinese	254	23.23	27.95	22.83	25.98
6	RW	English	382	9.42	29.58	39.27	21.73
6	RW	Spanish	8,103	13.79	31.65	38.06	16.50
6	RW	Other Language	3,933	19.60	25.91	31.10	23.39
6	RW	Students with Disabilities	4,424	16.52	35.17	33.52	14.78

H.7c: Grade 6 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
6	Overall	Female	6,064	12.73	30.08	43.98	13.21
6	Overall	Male	7,003	18.56	33.06	37.81	10.57
6	Overall	American Indian or Alaska Native	26	19.23	23.08	50.00	7.69
6	Overall	Black or African American	889	19.35	36.00	33.41	11.25
6	Overall	Asian	2,033	22.68	26.51	36.84	13.97
6	Overall	Native Hawaiian/Other Pacific Islander	10	0.00	20.00	70.00	10.00
6	Overall	Hispanic or Latino	8,286	12.94	32.09	43.36	11.61
6	Overall	White	1,770	17.80	34.24	36.95	11.02
6	Overall	Multiracial (not Hispanic origin)	28	25.00	42.86	28.57	3.57
6	Overall	NYC	8,447	15.50	33.40	39.64	11.47
6	Overall	Big 4 Cities	830	20.24	32.77	35.90	11.08
6	Overall	High Need Urban/Suburban	1,807	19.98	26.95	41.56	11.51
6	Overall	High Need Rural	78	6.41	29.49	55.13	8.97
6	Overall	Average Need	1,005	12.84	28.46	43.58	15.12
6	Overall	Low Need	388	6.96	24.48	51.55	17.01
6	Overall	Charter Schools	135	0.74	23.70	67.41	8.15
6	Overall	0 Years LEP	7,590	15.65	33.10	40.14	11.11
6	Overall	1 Years LEP	1,459	27.55	35.78	28.86	7.81
6	Overall	2 Years LEP	1,071	16.90	30.35	39.03	13.73
6	Overall	3 Years LEP	551	10.53	29.40	45.19	14.88
6	Overall	4 Years LEP	451	27.05	20.40	40.35	12.20
6	Overall	5 Years LEP	458	9.61	26.86	51.75	11.79
6	Overall	6 Years LEP or More	1,502	5.73	27.30	50.67	16.31
6	Overall	Arabic	338	26.63	28.11	35.80	9.47
6	Overall	Bengali	269	18.96	29.00	35.32	16.73
6	Overall	Chinese	250	24.40	28.80	37.60	9.20
6	Overall	English	378	9.52	29.10	44.44	16.93
6	Overall	Spanish	7,989	13.26	32.12	43.15	11.48
6	Overall	Other Language	3,878	20.22	31.72	36.10	11.96
6	Overall	Students with Disabilities	4,389	16.95	40.24	36.68	6.13

H.8a: Grade 7 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
7	LS	Female	5,923	4.47	12.41	30.25	52.86
7	LS	Male	7,279	6.24	13.70	30.21	49.86
7	LS	American Indian or Alaska Native	20	5.00	10.00	30.00	55.00
7	LS	Black or African American	927	3.13	12.08	38.19	46.60
7	LS	Asian	2,470	10.16	21.50	31.46	36.88
7	LS	Native Hawaiian/Other Pacific Islander	6	0.00	0.00	33.33	66.67
7	LS	Hispanic or Latino	8,078	3.87	9.53	28.66	57.94
7	LS	White	1,656	6.88	16.97	32.79	43.36
7	LS	Multiracial (not Hispanic origin)	20	15.00	35.00	25.00	25.00
7	LS	NYC	9,092	5.35	13.30	30.59	50.77
7	LS	Big 4 Cities	808	4.83	13.37	35.52	46.29
7	LS	High Need Urban/Suburban	1,674	8.42	13.62	28.38	49.58
7	LS	High Need Rural	69	5.80	5.80	28.99	59.42
7	LS	Average Need	861	4.30	13.12	27.53	55.05
7	LS	Low Need	405	1.98	8.64	27.90	61.48
7	LS	Charter Schools	90	1.11	1.11	23.33	74.44
7	LS	0 Years LEP	7,834	5.88	12.13	28.33	53.66
7	LS	1 Years LEP	1,718	9.14	28.81	38.13	23.92
7	LS	2 Years LEP	1,205	2.41	13.53	39.50	44.56
7	LS	3 Years LEP	426	0.23	7.04	39.20	53.52
7	LS	4 Years LEP	444	13.06	6.31	22.97	57.66
7	LS	5 Years LEP	291	0.34	2.06	24.40	73.20
7	LS	6 Years LEP or More	1,287	1.09	4.90	23.78	70.24
7	LS	Arabic	386	6.22	14.77	34.72	44.30
7	LS	Bengali	259	5.02	15.06	35.14	44.79
7	LS	Chinese	393	10.94	26.72	31.81	30.53
7	LS	English	310	2.58	2.90	24.52	70.00
7	LS	Spanish	7,853	3.93	9.75	28.89	57.42
7	LS	Other Language	4,029	8.04	18.94	32.49	40.53
7	LS	Students with Disabilities	3,891	3.08	8.35	35.18	53.38

H.8b: Grade 7 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
7	RW	Female	5,777	16.86	31.30	34.12	17.73
7	RW	Male	6,801	21.56	33.57	31.45	13.42
7	RW	American Indian or Alaska Native	20	10.00	40.00	25.00	25.00
7	RW	Black or African American	903	20.93	34.11	30.12	14.84
7	RW	Asian	2,267	26.47	29.11	27.66	16.76
7	RW	Native Hawaiian/Other Pacific Islander	6	0.00	33.33	0.00	66.67
7	RW	Hispanic or Latino	7,807	16.57	34.15	35.20	14.08
7	RW	White	1,535	20.65	28.53	29.97	20.85
7	RW	Multiracial (not Hispanic origin)	16	25.00	37.50	18.75	18.75
7	RW	NYC	8,601	19.42	33.24	32.46	14.88
7	RW	Big 4 Cities	752	24.34	32.98	29.79	12.90
7	RW	High Need Urban/Suburban	1,648	23.36	31.55	32.10	12.99
7	RW	High Need Rural	65	10.77	32.31	36.92	20.00
7	RW	Average Need	832	14.18	30.05	35.34	20.43
7	RW	Low Need	399	10.53	23.56	35.09	30.83
7	RW	Charter Schools	89	3.37	25.84	49.44	21.35
7	RW	0 Years LEP	7,472	18.96	33.15	33.32	14.56
7	RW	1 Years LEP	1,503	33.93	36.13	20.36	9.58
7	RW	2 Years LEP	1,178	20.80	30.31	31.32	17.57
7	RW	3 Years LEP	432	11.57	32.41	36.57	19.44
7	RW	4 Years LEP	436	20.87	27.98	31.65	19.50
7	RW	5 Years LEP	290	7.59	32.76	40.00	19.66
7	RW	6 Years LEP or More	1,272	8.88	27.75	41.75	21.62
7	RW	Arabic	343	25.07	31.20	28.57	15.16
7	RW	Bengali	247	23.48	27.94	30.36	18.22
7	RW	Chinese	354	28.81	32.49	22.88	15.82
7	RW	English	305	12.13	27.21	37.38	23.28
7	RW	Spanish	7,581	16.73	34.53	34.77	13.97
7	RW	Other Language	3,777	23.85	29.18	29.44	17.53
7	RW	Students with Disabilities	3,712	18.43	40.17	30.66	10.75

H.8c: Grade 7 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
7	Overall	Female	5,718	16.51	31.69	36.25	15.55
7	Overall	Male	6,753	21.40	33.79	33.20	11.61
7	Overall	American Indian or Alaska Native	20	10.00	40.00	25.00	25.00
7	Overall	Black or African American	892	20.40	34.53	31.84	13.23
7	Overall	Asian	2,252	26.24	29.40	29.31	15.05
7	Overall	Native Hawaiian/Other Pacific Islander	6	0.00	33.33	0.00	66.67
7	Overall	Hispanic or Latino	7,751	16.29	34.33	36.45	12.93
7	Overall	White	1,510	20.73	29.34	36.36	13.58
7	Overall	Multiracial (not Hispanic origin)	16	25.00	37.50	25.00	12.50
7	Overall	NYC	8,528	19.10	33.56	34.64	12.70
7	Overall	Big 4 Cities	749	24.30	32.98	31.11	11.62
7	Overall	High Need Urban/Suburban	1,623	23.11	31.85	33.58	11.46
7	Overall	High Need Rural	65	10.77	32.31	40.00	16.92
7	Overall	Average Need	828	14.13	30.68	36.23	18.96
7	Overall	Low Need	398	10.55	23.37	38.94	27.14
7	Overall	Charter Schools	89	3.37	25.84	50.56	20.22
7	Overall	0 Years LEP	7,404	18.63	33.44	35.37	12.56
7	Overall	1 Years LEP	1,487	33.56	36.52	21.92	8.00
7	Overall	2 Years LEP	1,174	20.87	30.58	32.28	16.27
7	Overall	3 Years LEP	421	11.40	32.78	39.19	16.63
7	Overall	4 Years LEP	433	20.55	28.18	33.49	17.78
7	Overall	5 Years LEP	290	7.59	32.76	41.38	18.28
7	Overall	6 Years LEP or More	1,265	8.93	28.22	44.66	18.18
7	Overall	Arabic	343	25.07	31.20	30.32	13.41
7	Overall	Bengali	246	23.58	27.64	31.30	17.48
7	Overall	Chinese	352	28.69	32.39	25.28	13.64
7	Overall	English	303	11.88	27.39	37.95	22.77
7	Overall	Spanish	7,528	16.46	34.71	36.00	12.83
7	Overall	Other Language	3,726	23.59	29.84	33.06	13.50
7	Overall	Students with Disabilities	3,676	18.44	40.61	33.68	7.26

H.9a: Grade 8 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
8	LS	Female	5,816	5.42	16.39	26.55	51.65
8	LS	Male	7,155	6.57	15.90	26.40	51.13
8	LS	American Indian or Alaska Native	14	7.14	21.43	35.71	35.71
8	LS	Black or African American	956	5.13	15.06	28.56	51.26
8	LS	Asian	2,613	9.64	25.14	27.17	38.04
8	LS	Native Hawaiian/Other Pacific Islander	4	0.00	25.00	25.00	50.00
8	LS	Hispanic or Latino	7,782	4.93	12.62	25.21	57.23
8	LS	White	1,533	6.26	17.22	30.72	45.79
8	LS	Multiracial (not Hispanic origin)	34	8.82	41.18	23.53	26.47
8	LS	NYC	9,136	5.88	15.55	27.05	51.52
8	LS	Big 4 Cities	836	8.85	21.89	29.78	39.47
8	LS	High Need Urban/Suburban	1,537	6.90	18.41	23.03	51.66
8	LS	High Need Rural	80	3.75	13.75	25.00	57.50
8	LS	Average Need	823	5.71	16.52	23.82	53.95
8	LS	Low Need	386	4.40	11.14	26.42	58.03
8	LS	Charter Schools	72	0.00	0.00	25.00	75.00
8	LS	0 Years LEP	7,606	6.23	14.25	25.27	54.25
8	LS	1 Years LEP	1,859	11.78	30.66	29.42	28.13
8	LS	2 Years LEP	1,296	3.32	18.98	34.49	43.21
8	LS	3 Years LEP	471	1.49	12.53	34.18	51.80
8	LS	4 Years LEP	393	5.09	18.83	22.14	53.94
8	LS	5 Years LEP	264	0.76	3.79	15.53	79.92
8	LS	6 Years LEP or More	1,074	1.96	4.66	21.04	72.35
8	LS	Arabic	404	9.41	22.03	25.50	43.07
8	LS	Bengali	278	3.24	19.78	28.78	48.20
8	LS	Chinese	427	9.13	26.46	32.08	32.32
8	LS	English	302	3.31	10.60	27.48	58.61
8	LS	Spanish	7,575	4.95	12.74	25.20	57.11
8	LS	Other Language	4,011	7.88	21.02	28.15	42.96
8	LS	Students with Disabilities	3,412	2.73	9.03	31.33	56.92

H.9b: Grade 8 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
8	RW	Female	5,630	17.02	35.33	26.32	21.33
8	RW	Male	6,792	21.82	36.43	23.66	18.09
8	RW	American Indian or Alaska Native	12	25.00	50.00	16.67	8.33
8	RW	Black or African American	927	20.60	37.97	24.60	16.83
8	RW	Asian	2,416	24.17	33.90	22.19	19.74
8	RW	Native Hawaiian/Other Pacific Islander	4	25.00	0.00	25.00	50.00
8	RW	Hispanic or Latino	7,574	17.96	38.09	25.89	18.06
8	RW	White	1,423	18.20	27.34	24.95	29.52
8	RW	Multiracial (not Hispanic origin)	30	53.33	30.00	6.67	10.00
8	RW	NYC	8,739	18.96	36.89	25.45	18.70
8	RW	Big 4 Cities	780	31.03	38.08	18.72	12.18
8	RW	High Need Urban/Suburban	1,493	23.24	31.08	25.18	20.50
8	RW	High Need Rural	77	14.29	37.66	19.48	28.57
8	RW	Average Need	787	16.65	34.69	23.13	25.54
8	RW	Low Need	373	9.65	28.42	26.54	35.39
8	RW	Charter Schools	72	0.00	30.56	38.89	30.56
8	RW	0 Years LEP	7,285	18.22	36.64	26.12	19.03
8	RW	1 Years LEP	1,702	32.43	38.48	16.04	13.04
8	RW	2 Years LEP	1,268	22.16	36.44	22.16	19.24
8	RW	3 Years LEP	467	15.42	32.98	28.48	23.13
8	RW	4 Years LEP	392	26.53	23.98	23.21	26.28
8	RW	5 Years LEP	261	4.60	29.50	36.78	29.12
8	RW	6 Years LEP or More	1,038	8.86	33.33	30.15	27.65
8	RW	Arabic	355	28.45	35.21	20.85	15.49
8	RW	Bengali	273	22.34	30.77	23.81	23.08
8	RW	Chinese	392	22.19	38.52	17.60	21.68
8	RW	English	291	15.12	32.99	25.43	26.46
8	RW	Spanish	7,377	18.11	38.10	25.88	17.91
8	RW	Other Language	3,756	21.75	32.00	24.07	22.18
8	RW	Students with Disabilities	3,271	17.95	41.03	25.25	15.77

H.9c: Grade 8 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
8	Overall	Female	5,581	16.97	35.55	28.78	18.71
8	Overall	Male	6,711	21.40	36.81	26.03	15.77
8	Overall	American Indian or Alaska Native	12	25.00	50.00	16.67	8.33
8	Overall	Black or African American	922	20.61	38.18	25.49	15.73
8	Overall	Asian	2,402	24.60	33.89	23.56	17.94
8	Overall	Native Hawaiian/Other Pacific Islander	4	25.00	0.00	25.00	50.00
8	Overall	Hispanic or Latino	7,478	17.40	38.39	27.32	16.89
8	Overall	White	1,409	18.10	28.32	35.56	18.03
8	Overall	Multiracial (not Hispanic origin)	30	53.33	30.00	6.67	10.00
8	Overall	NYC	8,640	18.63	37.27	28.08	16.02
8	Overall	Big 4 Cities	770	30.78	38.18	20.13	10.91
8	Overall	High Need Urban/Suburban	1,477	22.95	31.14	27.76	18.14
8	Overall	High Need Rural	77	14.29	37.66	20.78	27.27
8	Overall	Average Need	785	16.82	35.54	23.82	23.82
8	Overall	Low Need	372	9.68	28.49	28.76	33.06
8	Overall	Charter Schools	72	0.00	30.56	44.44	25.00
8	Overall	0 Years LEP	7,196	17.84	36.96	28.92	16.27
8	Overall	1 Years LEP	1,687	32.31	38.94	17.43	11.32
8	Overall	2 Years LEP	1,265	22.06	37.00	23.48	17.47
8	Overall	3 Years LEP	459	15.47	32.24	31.37	20.92
8	Overall	4 Years LEP	385	25.97	24.16	25.45	24.42
8	Overall	5 Years LEP	261	4.60	29.89	38.70	26.82
8	Overall	6 Years LEP or More	1,029	8.84	33.43	33.14	24.59
8	Overall	Arabic	353	28.61	34.84	21.53	15.01
8	Overall	Bengali	272	22.06	31.62	25.37	20.96
8	Overall	Chinese	388	22.42	39.18	19.07	19.33
8	Overall	English	290	14.83	33.10	26.90	25.17
8	Overall	Spanish	7,281	17.54	38.43	27.33	16.70
8	Overall	Other Language	3,729	21.99	32.31	28.80	16.89
8	Overall	Students with Disabilities	3,243	17.73	41.44	30.22	10.61

H.10a: Grade 9 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
9	LS	Female	6,490	10.40	26.22	20.97	42.40
9	LS	Male	7,950	10.88	25.04	20.39	43.69
9	LS	American Indian or Alaska Native	30	13.33	30.00	23.33	33.33
9	LS	Black or African American	1,299	7.78	28.71	24.02	39.49
9	LS	Asian	3,239	13.95	33.41	23.28	29.36
9	LS	Native Hawaiian/Other Pacific Islander	6	0.00	50.00	16.67	33.33
9	LS	Hispanic or Latino	8,458	10.29	21.60	19.06	49.05
9	LS	White	1,316	6.91	27.58	20.90	44.60
9	LS	Multiracial (not Hispanic origin)	67	7.46	40.30	28.36	23.88
9	LS	NYC	9,106	9.82	25.44	20.56	44.18
9	LS	Big 4 Cities	952	12.18	26.47	20.69	40.65
9	LS	High Need Urban/Suburban	1,707	12.24	26.95	21.91	38.90
9	LS	High Need Rural	79	8.86	16.46	29.11	45.57
9	LS	Average Need	1,050	11.52	22.38	20.10	46.00
9	LS	Low Need	539	9.46	22.82	22.08	45.64
9	LS	Charter Schools	48	0.00	4.17	12.50	83.33
9	LS	0 Years LEP	8,203	11.41	23.58	19.04	45.97
9	LS	1 Years LEP	2,621	18.35	38.57	22.28	20.79
9	LS	2 Years LEP	1,450	5.45	28.69	29.17	36.69
9	LS	3 Years LEP	618	2.59	17.48	29.94	50.00
9	LS	4 Years LEP	309	2.27	10.36	23.62	63.75
9	LS	5 Years LEP	234	0.43	5.13	20.09	74.36
9	LS	6 Years LEP or More	985	1.42	18.58	10.66	69.34
9	LS	Arabic	438	11.19	31.05	23.29	34.47
9	LS	Bengali	380	12.63	33.16	23.68	30.53
9	LS	Chinese	493	13.18	36.51	25.56	24.75
9	LS	English	391	7.42	21.23	20.20	51.15
9	LS	Spanish	8,194	10.39	21.48	19.18	48.95
9	LS	Other Language	4,560	10.96	31.03	22.28	35.72
9	LS	Students with Disabilities	2,424	3.26	15.80	21.95	58.99

H.10b: Grade 9 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
9	RW	Female	6,166	18.67	48.48	20.21	12.65
9	RW	Male	7,290	19.62	46.58	20.74	13.06
9	RW	American Indian or Alaska Native	24	29.17	37.50	25.00	8.33
9	RW	Black or African American	1,209	19.69	48.06	19.27	12.99
9	RW	Asian	2,904	22.45	48.28	16.67	12.60
9	RW	Native Hawaiian/Other Pacific Islander	6	33.33	16.67	33.33	16.67
9	RW	Hispanic or Latino	7,980	18.63	48.73	21.89	10.74
9	RW	White	1,250	12.00	38.16	21.92	27.92
9	RW	Multiracial (not Hispanic origin)	59	28.81	49.15	15.25	6.78
9	RW	NYC	8,508	17.48	48.62	20.22	13.68
9	RW	Big 4 Cities	853	28.25	44.78	16.65	10.32
9	RW	High Need Urban/Suburban	1,594	22.65	44.48	22.40	10.48
9	RW	High Need Rural	70	20.00	47.14	21.43	11.43
9	RW	Average Need	985	16.65	48.22	23.05	12.08
9	RW	Low Need	516	15.31	40.50	23.45	20.74
9	RW	Charter Schools	48	8.33	29.17	41.67	20.83
9	RW	0 Years LEP	7,612	18.79	48.27	21.02	11.93
9	RW	1 Years LEP	2,352	29.76	49.57	12.54	8.12
9	RW	2 Years LEP	1,381	19.12	49.60	18.61	12.67
9	RW	3 Years LEP	601	12.81	45.76	24.63	16.81
9	RW	4 Years LEP	300	13.33	45.67	28.33	12.67
9	RW	5 Years LEP	233	7.30	42.06	34.33	16.31
9	RW	6 Years LEP or More	964	5.60	35.58	30.08	28.73
9	RW	Arabic	393	29.26	47.07	15.52	8.14
9	RW	Bengali	334	19.16	52.99	16.17	11.68
9	RW	Chinese	451	19.96	52.99	14.41	12.64
9	RW	English	369	14.36	48.78	21.41	15.45
9	RW	Spanish	7,733	18.78	48.70	21.79	10.73
9	RW	Other Language	4,192	19.37	43.99	19.44	17.20
9	RW	Students with Disabilities	2,305	13.75	49.89	21.04	15.31

H.10c: Grade 9 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
9	Overall	Female	6,079	19.58	48.87	20.69	10.86
9	Overall	Male	7,221	20.36	47.21	21.23	11.20
9	Overall	American Indian or Alaska Native	24	29.17	45.83	20.83	4.17
9	Overall	Black or African American	1,206	20.65	48.01	19.49	11.86
9	Overall	Asian	2,872	23.19	48.08	17.51	11.21
9	Overall	Native Hawaiian/Other Pacific Islander	6	33.33	16.67	33.33	16.67
9	Overall	Hispanic or Latino	7,870	19.54	47.93	22.19	10.34
9	Overall	White	1,238	12.28	49.52	23.26	14.94
9	Overall	Multiracial (not Hispanic origin)	59	30.51	47.46	15.25	6.78
9	Overall	NYC	8,412	18.30	49.73	20.65	11.33
9	Overall	Big 4 Cities	845	28.28	44.73	16.92	10.06
9	Overall	High Need Urban/Suburban	1,582	23.51	44.75	23.32	8.41
9	Overall	High Need Rural	70	20.00	48.57	20.00	11.43
9	Overall	Average Need	976	17.21	48.05	22.95	11.78
9	Overall	Low Need	514	16.93	39.30	24.71	19.07
9	Overall	Charter Schools	48	8.33	29.17	43.75	18.75
9	Overall	0 Years LEP	7,493	19.62	47.63	21.59	11.16
9	Overall	1 Years LEP	2,330	31.50	48.63	12.75	7.12
9	Overall	2 Years LEP	1,378	19.38	49.78	19.23	11.61
9	Overall	3 Years LEP	600	13.17	46.33	26.50	14.00
9	Overall	4 Years LEP	298	13.42	45.64	28.19	12.75
9	Overall	5 Years LEP	231	6.93	41.99	35.06	16.02
9	Overall	6 Years LEP or More	955	5.65	49.84	29.74	14.76
9	Overall	Arabic	389	29.56	46.79	15.94	7.71
9	Overall	Bengali	334	20.36	52.10	17.07	10.48
9	Overall	Chinese	444	21.40	51.58	15.99	11.04
9	Overall	English	366	15.85	47.27	22.40	14.48
9	Overall	Spanish	7,627	19.65	47.93	22.08	10.33
9	Overall	Other Language	4,154	19.96	47.52	20.13	12.40
9	Overall	Students with Disabilities	2,290	13.89	56.11	21.53	8.47

H.11a: Grade 10 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
10	LS	Female	6,256	7.38	27.94	23.10	41.58
10	LS	Male	6,762	7.53	26.10	23.28	43.09
10	LS	American Indian or Alaska Native	28	10.71	17.86	25.00	46.43
10	LS	Black or African American	1,067	3.19	25.30	28.12	43.39
10	LS	Asian	3,509	9.03	37.48	25.59	27.90
10	LS	Native Hawaiian/Other Pacific Islander	2	0.00	0.00	0.00	100.00
10	LS	Hispanic or Latino	7,195	7.95	22.68	21.01	48.35
10	LS	White	1,193	3.94	23.64	24.39	48.03
10	LS	Multiracial (not Hispanic origin)	33	6.06	33.33	30.30	30.30
10	LS	NYC	8,789	7.81	26.95	22.48	42.76
10	LS	Big 4 Cities	570	6.14	25.61	25.26	42.98
10	LS	High Need Urban/Suburban	1,311	6.10	26.01	25.63	42.26
10	LS	High Need Rural	56	5.36	23.21	25.00	46.43
10	LS	Average Need	889	5.06	22.72	23.17	49.04
10	LS	Low Need	432	4.63	22.45	23.84	49.07
10	LS	Charter Schools	44	0.00	2.27	15.91	81.82
10	LS	0 Years LEP	7,140	8.77	25.01	20.80	45.42
10	LS	1 Years LEP	2,438	9.89	38.64	27.07	24.41
10	LS	2 Years LEP	1,617	4.27	30.36	28.51	36.86
10	LS	3 Years LEP	556	3.06	26.44	27.34	43.17
10	LS	4 Years LEP	343	0.87	15.16	27.11	56.85
10	LS	5 Years LEP	274	1.09	15.33	27.01	56.57
10	LS	6 Years LEP or More	639	2.35	7.82	13.62	76.21
10	LS	Arabic	322	3.11	23.29	27.95	45.65
10	LS	Bengali	294	4.42	29.93	26.53	39.12
10	LS	Chinese	649	13.71	40.06	26.50	19.72
10	LS	English	306	2.29	16.67	25.16	55.88
10	LS	Spanish	7,028	8.03	22.91	21.07	47.99
10	LS	Other Language	4,433	6.59	32.30	25.31	35.80
10	LS	Students with Disabilities	1,751	2.97	13.99	24.96	58.08

H.11b: Grade 10 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
10	RW	Female	6,094	16.16	46.64	23.15	14.05
10	RW	Male	6,419	17.03	47.22	23.52	12.23
10	RW	American Indian or Alaska Native	25	4.00	56.00	20.00	20.00
10	RW	Black or African American	1,049	15.06	48.62	24.50	11.82
10	RW	Asian	3,329	18.50	48.54	20.10	12.86
10	RW	Native Hawaiian/Other Pacific Islander	2	0.00	50.00	50.00	0.00
10	RW	Hispanic or Latino	6,928	16.43	46.82	24.75	11.99
10	RW	White	1,154	13.78	41.07	23.22	21.92
10	RW	Multiracial (not Hispanic origin)	30	20.00	56.67	16.67	6.67
10	RW	NYC	8,420	15.71	47.49	23.99	12.80
10	RW	Big 4 Cities	558	23.48	45.34	20.07	11.11
10	RW	High Need Urban/Suburban	1,272	18.24	47.72	22.96	11.08
10	RW	High Need Rural	52	13.46	46.15	28.85	11.54
10	RW	Average Need	874	14.42	44.16	26.09	15.33
10	RW	Low Need	424	11.32	38.44	23.35	26.89
10	RW	Charter Schools	43	0.00	34.88	37.21	27.91
10	RW	0 Years LEP	6,821	15.94	46.53	24.98	12.55
10	RW	1 Years LEP	2,333	21.65	50.19	17.57	10.59
10	RW	2 Years LEP	1,568	17.47	50.26	17.98	14.29
10	RW	3 Years LEP	540	18.15	39.81	24.81	17.22
10	RW	4 Years LEP	340	11.18	45.88	30.00	12.94
10	RW	5 Years LEP	272	7.72	45.96	30.51	15.81
10	RW	6 Years LEP or More	625	8.32	38.40	32.64	20.64
10	RW	Arabic	312	17.63	52.88	19.87	9.62
10	RW	Bengali	281	12.10	55.16	19.93	12.81
10	RW	Chinese	608	22.20	46.71	18.91	12.17
10	RW	English	301	6.98	46.51	27.57	18.94
10	RW	Spanish	6,771	16.72	46.82	24.66	11.80
10	RW	Other Language	4,250	16.52	46.21	22.05	15.22
10	RW	Students with Disabilities	1,691	14.13	47.37	24.25	14.25

H.11c: Grade 10 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
10	Overall	Female	6,042	17.15	47.60	23.50	11.75
10	Overall	Male	6,375	17.98	47.58	23.55	10.90
10	Overall	American Indian or Alaska Native	25	4.00	60.00	16.00	20.00
10	Overall	Black or African American	1,044	15.61	49.62	24.33	10.44
10	Overall	Asian	3,307	19.66	50.41	20.08	9.86
10	Overall	Native Hawaiian/Other Pacific Islander	2	0.00	50.00	50.00	0.00
10	Overall	Hispanic or Latino	6,865	17.48	46.34	25.01	11.17
10	Overall	White	1,149	14.10	44.73	23.93	17.23
10	Overall	Multiracial (not Hispanic origin)	30	20.00	56.67	16.67	6.67
10	Overall	NYC	8,351	16.88	48.17	23.90	11.04
10	Overall	Big 4 Cities	551	23.96	44.83	20.87	10.34
10	Overall	High Need Urban/Suburban	1,266	18.33	49.21	23.54	8.93
10	Overall	High Need Rural	52	13.46	48.08	30.77	7.69
10	Overall	Average Need	868	14.75	44.93	26.15	14.17
10	Overall	Low Need	424	12.74	38.21	25.71	23.35
10	Overall	Charter Schools	43	0.00	34.88	37.21	27.91
10	Overall	0 Years LEP	6,760	17.13	46.88	24.69	11.30
10	Overall	1 Years LEP	2,308	23.05	50.56	18.46	7.93
10	Overall	2 Years LEP	1,566	18.01	50.70	18.45	12.84
10	Overall	3 Years LEP	539	18.00	40.82	27.46	13.73
10	Overall	4 Years LEP	338	11.24	46.45	29.88	12.43
10	Overall	5 Years LEP	271	7.75	49.82	27.68	14.76
10	Overall	6 Years LEP or More	621	8.21	42.19	33.98	15.62
10	Overall	Arabic	311	17.68	53.70	18.97	9.65
10	Overall	Bengali	281	12.46	56.94	18.86	11.74
10	Overall	Chinese	602	24.09	49.17	18.94	7.81
10	Overall	English	300	8.00	46.00	28.33	17.67
10	Overall	Spanish	6,708	17.80	46.32	24.91	10.97
10	Overall	Other Language	4,225	17.28	48.43	22.27	12.02
10	Overall	Students with Disabilities	1,683	14.32	50.15	25.19	10.34

H.12a: Grade 11 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
11	LS	Female	4,561	4.17	24.97	29.29	41.57
11	LS	Male	4,626	4.17	22.78	29.27	43.77
11	LS	American Indian or Alaska Native	17	0.00	35.29	17.65	47.06
11	LS	Black or African American	850	2.12	22.71	36.24	38.94
11	LS	Asian	2,341	3.42	30.76	32.04	33.79
11	LS	Native Hawaiian/Other Pacific Islander	8	0.00	25.00	25.00	50.00
11	LS	Hispanic or Latino	5,096	4.96	21.33	27.53	46.17
11	LS	White	865	3.82	20.58	25.43	50.17
11	LS	Multiracial (not Hispanic origin)	17	0.00	52.94	17.65	29.41
11	LS	NYC	6,190	4.10	24.10	29.76	42.04
11	LS	Big 4 Cities	410	3.90	27.32	24.63	44.15
11	LS	High Need Urban/Suburban	977	5.73	26.61	27.43	40.23
11	LS	High Need Rural	45	2.22	13.33	33.33	51.11
11	LS	Average Need	747	2.41	21.69	28.25	47.66
11	LS	Low Need	407	4.18	16.22	34.64	44.96
11	LS	Charter Schools	26	0.00	3.85	19.23	76.92
11	LS	0 Years LEP	4,579	4.28	20.14	27.89	47.70
11	LS	1 Years LEP	1,592	7.85	36.49	30.65	25.00
11	LS	2 Years LEP	1,365	3.08	30.18	32.89	33.85
11	LS	3 Years LEP	583	1.37	24.19	35.51	38.94
11	LS	4 Years LEP	294	0.68	17.01	32.65	49.66
11	LS	5 Years LEP	273	0.73	18.68	27.47	53.11
11	LS	6 Years LEP or More	494	1.62	8.91	18.83	70.65
11	LS	Arabic	209	2.39	23.92	25.36	48.33
11	LS	Bengali	231	0.87	26.84	35.06	37.23
11	LS	Chinese	347	3.17	35.16	35.45	26.22
11	LS	English	220	2.27	11.82	28.64	57.27
11	LS	Spanish	4,935	5.01	21.50	27.56	45.94
11	LS	Other Language	3,266	3.49	27.04	31.05	38.43
11	LS	Students with Disabilities	1,219	3.12	19.28	28.63	48.97

H.12b: Grade 11 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
11	RW	Female	4,496	9.83	50.36	23.42	16.39
11	RW	Male	4,468	10.38	49.13	24.57	15.91
11	RW	American Indian or Alaska Native	17	11.76	52.94	11.76	23.53
11	RW	Black or African American	834	10.31	53.36	21.94	14.39
11	RW	Asian	2,294	9.55	49.87	23.06	17.52
11	RW	Native Hawaiian/Other Pacific Islander	8	0.00	37.50	37.50	25.00
11	RW	Hispanic or Latino	4,962	10.72	49.72	24.77	14.79
11	RW	White	840	8.10	45.24	24.29	22.38
11	RW	Multiracial (not Hispanic origin)	15	6.67	60.00	20.00	13.33
11	RW	NYC	6,023	9.85	50.22	24.19	15.74
11	RW	Big 4 Cities	398	15.83	51.51	20.85	11.81
11	RW	High Need Urban/Suburban	959	12.51	51.93	22.31	13.24
11	RW	High Need Rural	43	9.30	44.19	23.26	23.26
11	RW	Average Need	740	7.97	45.81	27.03	19.19
11	RW	Low Need	400	8.50	42.75	29.75	19.00
11	RW	Charter Schools	26	3.85	23.08	42.31	30.77
11	RW	0 Years LEP	4,438	9.73	48.26	25.82	16.18
11	RW	1 Years LEP	1,554	14.29	51.29	18.47	15.96
11	RW	2 Years LEP	1,346	9.96	54.75	20.73	14.56
11	RW	3 Years LEP	573	9.42	49.04	25.13	16.40
11	RW	4 Years LEP	293	9.22	50.17	23.21	17.41
11	RW	5 Years LEP	269	6.69	48.33	29.37	15.61
11	RW	6 Years LEP or More	484	4.75	46.28	29.55	19.42
11	RW	Arabic	201	12.94	57.21	16.92	12.94
11	RW	Bengali	224	6.25	59.38	22.77	11.61
11	RW	Chinese	342	11.70	43.57	21.35	23.39
11	RW	English	215	5.12	48.37	24.65	21.86
11	RW	Spanish	4,806	10.90	49.46	24.70	14.94
11	RW	Other Language	3,196	9.23	49.75	23.65	17.37
11	RW	Students with Disabilities	1,168	11.90	49.83	21.40	16.87

H.12c: Grade 11 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
11	Overall	Female	4,477	10.74	51.73	24.21	13.31
11	Overall	Male	4,457	10.97	50.28	25.56	13.19
11	Overall	American Indian or Alaska Native	17	11.76	58.82	5.88	23.53
11	Overall	Black or African American	833	10.80	54.50	21.97	12.73
11	Overall	Asian	2,292	10.30	52.49	24.30	12.91
11	Overall	Native Hawaiian/Other Pacific Islander	8	0.00	50.00	25.00	25.00
11	Overall	Hispanic or Latino	4,941	11.50	50.09	25.95	12.47
11	Overall	White	834	8.99	47.84	23.74	19.42
11	Overall	Multiracial (not Hispanic origin)	15	6.67	66.67	20.00	6.67
11	Overall	NYC	6,003	10.61	51.52	25.14	12.73
11	Overall	Big 4 Cities	393	15.78	52.42	20.36	11.45
11	Overall	High Need Urban/Suburban	958	13.26	53.76	22.44	10.54
11	Overall	High Need Rural	43	9.30	44.19	30.23	16.28
11	Overall	Average Need	739	8.80	47.09	28.15	15.97
11	Overall	Low Need	399	9.27	43.11	30.58	17.04
11	Overall	Charter Schools	26	3.85	23.08	46.15	26.92
11	Overall	0 Years LEP	4,422	10.43	49.37	26.62	13.59
11	Overall	1 Years LEP	1,547	15.38	53.59	19.46	11.57
11	Overall	2 Years LEP	1,344	10.79	55.51	21.80	11.90
11	Overall	3 Years LEP	572	9.79	49.13	27.45	13.64
11	Overall	4 Years LEP	291	9.28	50.86	25.77	14.09
11	Overall	5 Years LEP	269	7.06	53.16	27.88	11.90
11	Overall	6 Years LEP or More	482	5.81	46.89	29.25	18.05
11	Overall	Arabic	200	13.00	57.50	17.50	12.00
11	Overall	Bengali	224	6.25	60.71	24.55	8.48
11	Overall	Chinese	342	12.57	50.29	23.39	13.74
11	Overall	English	215	5.12	48.84	25.58	20.47
11	Overall	Spanish	4,786	11.68	49.83	25.95	12.54
11	Overall	Other Language	3,187	10.07	51.90	23.82	14.21
11	Overall	Students with Disabilities	1,160	12.67	52.24	21.38	13.71

H.13a: Grade 12 Listening and Speaking

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
12	LS	Female	3,204	1.34	20.47	23.85	54.34
12	LS	Male	2,878	2.08	20.08	22.65	55.18
12	LS	American Indian or Alaska Native	11	0.00	18.18	9.09	72.73
12	LS	Black or African American	564	0.53	14.89	28.37	56.21
12	LS	Asian	1,568	1.34	27.55	24.81	46.30
12	LS	Native Hawaiian/Other Pacific Islander	4	0.00	0.00	50.00	50.00
12	LS	Hispanic or Latino	3,342	1.65	18.76	22.11	57.48
12	LS	White	591	4.06	15.57	20.30	60.07
12	LS	Multiracial (not Hispanic origin)	9	0.00	11.11	44.44	44.44
12	LS	NYC	4,313	1.46	21.91	23.42	53.21
12	LS	Big 4 Cities	253	1.98	24.11	22.92	50.99
12	LS	High Need Urban/Suburban	552	1.81	18.12	24.64	55.43
12	LS	High Need Rural	31	0.00	6.45	25.81	67.74
12	LS	Average Need	490	4.69	12.04	21.63	61.63
12	LS	Low Need	321	0.62	15.26	22.74	61.37
12	LS	Charter Schools	17	0.00	11.76	17.65	70.59
12	LS	0 Years LEP	3,319	1.02	18.86	22.90	57.22
12	LS	1 Years LEP	622	3.70	32.64	23.63	40.03
12	LS	2 Years LEP	885	1.81	27.34	24.18	46.67
12	LS	3 Years LEP	263	1.90	18.25	21.29	58.56
12	LS	4 Years LEP	372	0.81	17.20	29.30	52.69
12	LS	5 Years LEP	229	0.00	9.17	23.58	67.25
12	LS	6 Years LEP or More	392	5.61	7.91	19.64	66.84
12	LS	Arabic	145	0.00	15.86	20.00	64.14
12	LS	Bengali	170	0.00	30.00	30.00	40.00
12	LS	Chinese	225	2.22	38.22	23.11	36.44
12	LS	English	149	1.34	6.71	19.46	72.48
12	LS	Spanish	3,262	1.62	18.98	22.19	57.20
12	LS	Other Language	2,145	2.05	20.93	24.90	52.12
12	LS	Students with Disabilities	749	4.01	20.43	26.03	49.53

H.13b: Grade 12 Reading and Writing

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
12	RW	Female	3,185	5.31	51.77	24.24	18.68
12	RW	Male	2,818	6.88	53.30	22.04	17.78
12	RW	American Indian or Alaska Native	11	18.18	36.36	27.27	18.18
12	RW	Black or African American	559	5.01	57.78	22.18	15.03
12	RW	Asian	1,556	6.30	50.64	22.17	20.89
12	RW	Native Hawaiian/Other Pacific Islander	4	0.00	50.00	0.00	50.00
12	RW	Hispanic or Latino	3,293	5.65	53.96	23.96	16.43
12	RW	White	578	8.82	43.77	22.15	25.26
12	RW	Multiracial (not Hispanic origin)	9	0.00	55.56	44.44	0.00
12	RW	NYC	4,251	5.72	53.66	23.01	17.62
12	RW	Big 4 Cities	250	9.60	56.00	21.20	13.20
12	RW	High Need Urban/Suburban	544	7.35	53.49	23.71	15.44
12	RW	High Need Rural	30	3.33	43.33	43.33	10.00
12	RW	Average Need	484	6.82	47.52	24.38	21.28
12	RW	Low Need	320	3.44	40.00	23.75	32.81
12	RW	Charter Schools	17	11.76	41.18	23.53	23.53
12	RW	0 Years LEP	3,266	6.03	53.40	23.09	17.48
12	RW	1 Years LEP	618	6.15	51.62	22.01	20.23
12	RW	2 Years LEP	878	5.58	53.64	21.98	18.79
12	RW	3 Years LEP	261	4.60	45.98	24.52	24.90
12	RW	4 Years LEP	371	6.74	52.02	24.26	16.98
12	RW	5 Years LEP	225	4.44	52.44	24.89	18.22
12	RW	6 Years LEP or More	383	8.36	47.78	25.59	18.28
12	RW	Arabic	146	8.22	59.59	17.81	14.38
12	RW	Bengali	167	12.57	63.47	13.77	10.18
12	RW	Chinese	222	6.31	50.90	22.97	19.82
12	RW	English	148	4.73	41.89	25.68	27.70
12	RW	Spanish	3,214	5.63	54.36	23.83	16.18
12	RW	Other Language	2,119	6.13	49.13	23.12	21.61
12	RW	Students with Disabilities	718	16.30	57.52	16.43	9.75

H.13c: Grade 12 Overall

Grade	Test	Group	N-count	Proficiency Levels			
				1	2	3	4
12	Overall	Female	3,182	5.85	53.93	24.73	15.49
12	Overall	Male	2,818	7.31	55.07	22.71	14.90
12	Overall	American Indian or Alaska Native	11	18.18	36.36	27.27	18.18
12	Overall	Black or African American	559	5.19	58.68	22.36	13.77
12	Overall	Asian	1,556	7.01	54.50	23.14	15.36
12	Overall	Native Hawaiian/Other Pacific Islander	4	0.00	50.00	25.00	25.00
12	Overall	Hispanic or Latino	3,292	6.20	55.07	24.30	14.43
12	Overall	White	576	8.68	46.88	23.09	21.35
12	Overall	Multiracial (not Hispanic origin)	9	0.00	55.56	44.44	0.00
12	Overall	NYC	4,250	6.38	55.79	23.41	14.42
12	Overall	Big 4 Cities	250	10.00	58.00	20.80	11.20
12	Overall	High Need Urban/Suburban	543	7.37	55.62	24.49	12.52
12	Overall	High Need Rural	30	3.33	43.33	46.67	6.67
12	Overall	Average Need	484	6.82	48.35	25.00	19.83
12	Overall	Low Need	320	3.44	42.50	26.56	27.50
12	Overall	Charter Schools	17	11.76	41.18	23.53	23.53
12	Overall	0 Years LEP	3,264	6.40	55.24	23.59	14.77
12	Overall	1 Years LEP	617	7.46	55.92	21.88	14.75
12	Overall	2 Years LEP	878	6.49	55.81	22.55	15.15
12	Overall	3 Years LEP	261	4.98	47.89	26.05	21.07
12	Overall	4 Years LEP	371	6.74	52.83	25.34	15.09
12	Overall	5 Years LEP	225	4.44	53.33	24.89	17.33
12	Overall	6 Years LEP or More	383	8.36	48.83	27.15	15.67
12	Overall	Arabic	145	7.59	60.69	20.00	11.72
12	Overall	Bengali	167	12.57	66.47	11.38	9.58
12	Overall	Chinese	222	6.31	56.31	23.42	13.96
12	Overall	English	148	4.73	41.89	26.35	27.03
12	Overall	Spanish	3,213	6.19	55.49	24.15	14.16
12	Overall	Other Language	2,118	6.70	52.17	24.22	16.90
12	Overall	Students with Disabilities	716	16.34	59.64	17.18	6.84

APPENDIX I: EXIT RATE BY SUBGROUP**I.1: Kindergarten**

Grade	Group	N-count	Percent
K	Female	1,341	10.92
K	Male	1,174	8.97
K	American Indian or Alaska Native	6	9.38
K	Black or African American	108	9.07
K	Asian	827	15.46
K	Native Hawaiian/Other Pacific Islander	3	15.79
K	Hispanic or Latino	1,170	7.52
K	White	373	12.50
K	Multiracial (not Hispanic origin)	22	11.70
K	NYC	1,789	11.64
K	Big 4 Cities	36	3.36
K	High Need Urban/Suburban	200	5.91
K	High Need Rural	4	2.45
K	Average Need	252	8.41
K	Low Need	116	8.59
K	Charter Schools	38	13.92
K	0 Years LEP	1,729	11.68
K	1 Years LEP	784	7.43
K	Arabic	54	9.34
K	Bengali	113	16.64
K	Chinese	178	14.30
K	English	129	8.29
K	Spanish	1,118	7.62
K	Other Language	923	13.82
K	Students with Disabilities	148	3.52

I.2: Grade 1

Grade	Group	N-count	Percent
1	Female	3,122	23.96
1	Male	2,933	20.63
1	American Indian or Alaska Native	11	25.00
1	Black or African American	210	16.55
1	Asian	1,602	30.42
1	Native Hawaiian/Other Pacific Islander	6	37.50
1	Hispanic or Latino	3,435	20.27
1	White	756	21.29
1	Multiracial (not Hispanic origin)	37	29.37
1	NYC	3,852	23.10
1	Big 4 Cities	166	13.54
1	High Need Urban/Suburban	642	15.99
1	High Need Rural	49	25.79
1	Average Need	763	25.75
1	Low Need	379	29.09
1	Charter Schools	56	23.14
1	0 Years LEP	3,318	22.82
1	1 Years LEP	694	16.04
1	2 Years LEP	2,040	24.26
1	Arabic	132	20.53
1	Bengali	226	31.65
1	Chinese	373	33.33
1	English	332	26.62
1	Spanish	3,202	19.99
1	Other Language	1,794	23.64
1	Students with Disabilities	462	8.51

I.3: Grade 2

Grade	Group	N-count	Percent
2	Female	1,570	14.91
2	Male	1,492	13.06
2	American Indian or Alaska Native	11	23.40
2	Black or African American	124	10.85
2	Asian	922	23.76
2	Native Hawaiian/Other Pacific Islander	4	25.00
2	Hispanic or Latino	1,610	11.42
2	White	390	14.37
2	Multiracial (not Hispanic origin)	9	16.07
2	NYC	1,770	12.80
2	Big 4 Cities	109	9.98
2	High Need Urban/Suburban	366	11.44
2	High Need Rural	23	18.55
2	Average Need	428	19.26
2	Low Need	301	29.68
2	Charter Schools	31	15.58
2	0 Years LEP	1,474	12.37
2	1 Years LEP	279	13.81
2	2 Years LEP	360	13.45
2	3 Years LEP	954	17.66
2	Arabic	69	13.14
2	Bengali	129	22.28
2	Chinese	208	26.10
2	English	181	19.80
2	Spanish	1,481	11.06
2	Other Language	1,003	17.11
2	Students with Disabilities	230	4.58

I.4: Grade 3

Grade	Group	N-count	Percent
3	Female	1,986	20.06
3	Male	1,988	18.15
3	American Indian or Alaska Native	5	15.63
3	Black or African American	154	14.08
3	Asian	915	26.29
3	Native Hawaiian/Other Pacific Islander	3	37.50
3	Hispanic or Latino	2,444	17.97
3	White	449	17.32
3	Multiracial (not Hispanic origin)	8	29.63
3	NYC	2,499	18.66
3	Big 4 Cities	176	15.17
3	High Need Urban/Suburban	524	17.31
3	High Need Rural	19	14.62
3	Average Need	464	23.40
3	Low Need	240	28.47
3	Charter Schools	26	20.47
3	0 Years LEP	2,241	18.68
3	1 Years LEP	216	13.88
3	2 Years LEP	317	18.85
3	3 Years LEP	219	16.55
3	4 Years LEP	980	22.76
3	Arabic	75	14.88
3	Bengali	152	27.14
3	Chinese	165	28.80
3	English	203	26.19
3	Spanish	2,305	17.77
3	Other Language	1,082	19.58
3	Students with Disabilities	362	6.79

I.5: Grade 4

Grade	Group	N-count	Percent
4	Female	2,150	24.93
4	Male	2,293	21.96
4	American Indian or Alaska Native	6	30.00
4	Black or African American	194	18.03
4	Asian	873	28.35
4	Native Hawaiian/Other Pacific Islander	2	22.22
4	Hispanic or Latino	2,919	23.50
4	White	448	18.20
4	Multiracial (not Hispanic origin)	3	15.79
4	NYC	2,921	23.40
4	Big 4 Cities	177	16.50
4	High Need Urban/Suburban	599	20.85
4	High Need Rural	31	21.99
4	Average Need	436	26.88
4	Low Need	231	32.86
4	Charter Schools	23	23.23
4	0 Years LEP	2,651	24.12
4	1 Years LEP	250	14.87
4	2 Years LEP	259	21.99
4	3 Years LEP	163	17.87
4	4 Years LEP	252	22.36
4	5 Years LEP	866	26.88
4	Arabic	90	19.82
4	Bengali	126	30.73
4	Chinese	138	29.81
4	English	195	28.93
4	Spanish	2,801	23.47
4	Other Language	1,096	21.10
4	Students with Disabilities	558	10.13

I.6: Grade 5

Grade	Group	N-count	Percent
5	Female	1,271	18.60
5	Male	1,353	15.80
5	American Indian or Alaska Native	2	10.00
5	Black or African American	116	12.10
5	Asian	540	24.09
5	Native Hawaiian/Other Pacific Islander	3	30.00
5	Hispanic or Latino	1,676	16.66
5	White	282	13.60
5	Multiracial (not Hispanic origin)	2	10.00
5	NYC	1,731	17.16
5	Big 4 Cities	87	9.61
5	High Need Urban/Suburban	340	14.65
5	High Need Rural	11	10.09
5	Average Need	250	21.01
5	Low Need	143	29.18
5	Charter Schools	44	26.04
5	0 Years LEP	1,553	17.26
5	1 Years LEP	178	12.03
5	2 Years LEP	207	19.27
5	3 Years LEP	82	14.14
5	4 Years LEP	113	15.76
5	5 Years LEP	122	16.31
5	6 Years LEP or More	363	19.89
5	Arabic	60	15.71
5	Bengali	71	24.40
5	Chinese	74	22.63
5	English	109	21.89
5	Spanish	1,588	16.48
5	Other Language	722	16.77
5	Students with Disabilities	360	7.18

I.7: Grade 6

Grade	Group	N-count	Percent
6	Female	801	13.21
6	Male	740	10.57
6	American Indian or Alaska Native	2	7.69
6	Black or African American	100	11.25
6	Asian	284	13.97
6	Native Hawaiian/Other Pacific Islander	1	10.00
6	Hispanic or Latino	962	11.61
6	White	195	11.02
6	Multiracial (not Hispanic origin)	1	3.57
6	NYC	969	11.47
6	Big 4 Cities	92	11.08
6	High Need Urban/Suburban	208	11.51
6	High Need Rural	7	8.97
6	Average Need	152	15.12
6	Low Need	66	17.01
6	Charter Schools	11	8.15
6	0 Years LEP	843	11.11
6	1 Years LEP	114	7.81
6	2 Years LEP	147	13.73
6	3 Years LEP	82	14.88
6	4 Years LEP	55	12.20
6	5 Years LEP	54	11.79
6	6 Years LEP or More	245	16.31
6	Arabic	32	9.47
6	Bengali	45	16.73
6	Chinese	23	9.20
6	English	64	16.93
6	Spanish	917	11.48
6	Other Language	464	11.96
6	Students with Disabilities	269	6.13

I.8: Grade 7

Grade	Group	N-count	Percent
7	Female	889	15.55
7	Male	784	11.61
7	American Indian or Alaska Native	5	25.00
7	Black or African American	118	13.23
7	Asian	339	15.05
7	Native Hawaiian/Other Pacific Islander	4	66.67
7	Hispanic or Latino	1,002	12.93
7	White	205	13.58
7	Multiracial (not Hispanic origin)	2	12.50
7	NYC	1,083	12.70
7	Big 4 Cities	87	11.62
7	High Need Urban/Suburban	186	11.46
7	High Need Rural	11	16.92
7	Average Need	157	18.96
7	Low Need	108	27.14
7	Charter Schools	18	20.22
7	0 Years LEP	930	12.56
7	1 Years LEP	119	8.00
7	2 Years LEP	191	16.27
7	3 Years LEP	70	16.63
7	4 Years LEP	77	17.78
7	5 Years LEP	53	18.28
7	6 Years LEP or More	230	18.18
7	Arabic	46	13.41
7	Bengali	43	17.48
7	Chinese	48	13.64
7	English	69	22.77
7	Spanish	966	12.83
7	Other Language	503	13.50
7	Students with Disabilities	267	7.26

I.9: Grade 8

Grade	Group	N-count	Percent
8	Female	1,044	18.71
8	Male	1,058	15.77
8	American Indian or Alaska Native	1	8.33
8	Black or African American	145	15.73
8	Asian	431	17.94
8	Native Hawaiian/Other Pacific Islander	2	50.00
8	Hispanic or Latino	1,263	16.89
8	White	254	18.03
8	Multiracial (not Hispanic origin)	3	10.00
8	NYC	1,384	16.02
8	Big 4 Cities	84	10.91
8	High Need Urban/Suburban	268	18.14
8	High Need Rural	21	27.27
8	Average Need	187	23.82
8	Low Need	123	33.06
8	Charter Schools	18	25.00
8	0 Years LEP	1,171	16.27
8	1 Years LEP	191	11.32
8	2 Years LEP	221	17.47
8	3 Years LEP	96	20.92
8	4 Years LEP	94	24.42
8	5 Years LEP	70	26.82
8	6 Years LEP or More	253	24.59
8	Arabic	53	15.01
8	Bengali	57	20.96
8	Chinese	75	19.33
8	English	73	25.17
8	Spanish	1,216	16.70
8	Other Language	630	16.89
8	Students with Disabilities	344	10.61

L.10: Grade 9

Grade	Group	N-count	Percent
9	Female	660	10.86
9	Male	809	11.20
9	American Indian or Alaska Native	1	4.17
9	Black or African American	143	11.86
9	Asian	322	11.21
9	Native Hawaiian/Other Pacific Islander	1	16.67
9	Hispanic or Latino	814	10.34
9	White	185	14.94
9	Multiracial (not Hispanic origin)	4	6.78
9	NYC	953	11.33
9	Big 4 Cities	85	10.06
9	High Need Urban/Suburban	133	8.41
9	High Need Rural	8	11.43
9	Average Need	115	11.78
9	Low Need	98	19.07
9	Charter Schools	9	18.75
9	0 Years LEP	836	11.16
9	1 Years LEP	166	7.12
9	2 Years LEP	160	11.61
9	3 Years LEP	84	14.00
9	4 Years LEP	38	12.75
9	5 Years LEP	37	16.02
9	6 Years LEP or More	141	14.76
9	Arabic	30	7.71
9	Bengali	35	10.48
9	Chinese	49	11.04
9	English	53	14.48
9	Spanish	788	10.33
9	Other Language	515	12.40
9	Students with Disabilities	194	8.47

I.11: Grade 10

Grade	Group	N-count	Percent
10	Female	710	11.75
10	Male	695	10.90
10	American Indian or Alaska Native	5	20.00
10	Black or African American	109	10.44
10	Asian	326	9.86
10	Native Hawaiian/Other Pacific Islander		
10	Hispanic or Latino	767	11.17
10	White	198	17.23
10	Multiracial (not Hispanic origin)	2	6.67
10	NYC	922	11.04
10	Big 4 Cities	57	10.34
10	High Need Urban/Suburban	113	8.93
10	High Need Rural	4	7.69
10	Average Need	123	14.17
10	Low Need	99	23.35
10	Charter Schools	12	27.91
10	0 Years LEP	764	11.30
10	1 Years LEP	183	7.93
10	2 Years LEP	201	12.84
10	3 Years LEP	74	13.73
10	4 Years LEP	42	12.43
10	5 Years LEP	40	14.76
10	6 Years LEP or More	97	15.62
10	Arabic	30	9.65
10	Bengali	33	11.74
10	Chinese	47	7.81
10	English	53	17.67
10	Spanish	736	10.97
10	Other Language	508	12.02
10	Students with Disabilities	174	10.34

I.12: Grade 11

Grade	Group	N-count	Percent
11	Female	596	13.31
11	Male	588	13.19
11	American Indian or Alaska Native	4	23.53
11	Black or African American	106	12.73
11	Asian	296	12.91
11	Native Hawaiian/Other Pacific Islander	2	25.00
11	Hispanic or Latino	616	12.47
11	White	162	19.42
11	Multiracial (not Hispanic origin)	1	6.67
11	NYC	764	12.73
11	Big 4 Cities	45	11.45
11	High Need Urban/Suburban	101	10.54
11	High Need Rural	7	16.28
11	Average Need	118	15.97
11	Low Need	68	17.04
11	Charter Schools	7	26.92
11	0 Years LEP	601	13.59
11	1 Years LEP	179	11.57
11	2 Years LEP	160	11.90
11	3 Years LEP	78	13.64
11	4 Years LEP	41	14.09
11	5 Years LEP	32	11.90
11	6 Years LEP or More	87	18.05
11	Arabic	24	12.00
11	Bengali	19	8.48
11	Chinese	47	13.74
11	English	44	20.47
11	Spanish	600	12.54
11	Other Language	453	14.21
11	Students with Disabilities	159	13.71

I.13: Grade 12

Grade	Group	N-count	Percent
12	Female	493	15.49
12	Male	420	14.90
12	American Indian or Alaska Native	2	18.18
12	Black or African American	77	13.77
12	Asian	239	15.36
12	Native Hawaiian/Other Pacific Islander	1	25.00
12	Hispanic or Latino	475	14.43
12	White	123	21.35
12	Multiracial (not Hispanic origin)		
12	NYC	613	14.42
12	Big 4 Cities	28	11.20
12	High Need Urban/Suburban	68	12.52
12	High Need Rural	2	6.67
12	Average Need	96	19.83
12	Low Need	88	27.50
12	Charter Schools	4	23.53
12	0 Years LEP	482	14.77
12	1 Years LEP	91	14.75
12	2 Years LEP	133	15.15
12	3 Years LEP	55	21.07
12	4 Years LEP	56	15.09
12	5 Years LEP	39	17.33
12	6 Years LEP or More	60	15.67
12	Arabic	17	11.72
12	Bengali	16	9.58
12	Chinese	31	13.96
12	English	40	27.03
12	Spanish	455	14.16
12	Other Language	358	16.90
12	Students with Disabilities	49	6.84

APPENDIX J: DIF CATEGORIES FOR OPERATIONAL ITEMS**J. 1 Results of DIF Analyses**

Grade Span	Modality	Item Seq. Number	DIF– Male/Female	DIF– Asian/Other	DIF– Hispanic/Other	DIF– White/Other
K-1	Listen/Speaking	1	A	<B	A	>B
		2	A	<B	A	>B
		4	A	A	C
		6	A	A	A	>B
		7	A	A	<C	>C
		8	A	A	<B	A
		10	A	>B	A	A
		12	A	A	A	<B
		13	A	A	A	<B
		16	A	A	C
		18	A	A	A	>B
		32	AA	AA	AA	>BB
		34	AA	AA	AA	>BB
		35	AA	AA	AA	>BB
38	AA	AA	AA	>BB		
40	AA	AA	AA	>CC		
K-1	Reading/Writing	1	A	A	A	>C
		17	A	A	A	>B
		22	AA	AA	AA	>BB
		23	AA	AA	AA	>BB
2-4	Listen/Speaking	1	A	A	<C	>C
		2	A	A	A	>B
		4	A	A	C
		5	A	A	<C	>C
		6	A	A	A	>C
		9	A	A	A	<B
		10	A	A	<C	>C
		11	A	A	A	>B
		13	A	A	A	<B
		19	A	A	A	<B
		20	A	A	A	>B
		25	AA	>BB	AA	AA
		38	AA	AA	AA	>BB
		2-4	Reading/Writing	1	A	<B
3	A			A	A	>B
5	A			A	A	>B
6	A			A	<C	>C
16	A			A	A	<B
22	A			>B	A	A
24	A			A	A	>B
26	A			A	A	>C
27	A			<C	A	A
30	A			A	>B	<B

J.1 Results of DIF Analyses (continued)

Grade Span	Modality	Item Seq. Number	DIF– Male/Female	DIF– Asian/Other	DIF– Hispanic/Other	DIF– White/Other
		31	A	A	A	<C
		37	AA	AA	AA	>BB
5–6	Listen/Speaking	1	<B	<C	<C	>C
		2	A	<C	A	>C
		4	A	A	A	>B
		5	<B	<B	A	>C
		6	A	A	B
		8	A	<B	A	A
		15	A	A	>B	<C
		20	A	A	A	>C
		21	A	A	A	<B
		24	A	A	A	>B
		25	A	<B	A	>B
		27	AA	>BB	AA	<BB
		29	AA	>BB	AA	<BB
		30	AA	AA	AA	<BB
		38	AA	>CC	AA	<BB
		41	AA	>BB	AA	AA
5–6	Reading/Writing	1	A	A	<C	>C
		2	A	<B	C
		3	>B	A	C
		4	<B	A	C
		6	A	A	A	>B
		14	A	A	A	<B
		20	A	<B	A	A
		25	A	A	A	>B
		28	A	>C	<C	>C
		29	A	>B	B
		33	A	A	A	<B
		34	A	<B	A	>C
		35	A	>C	<C	A
		36	A	>C	<B	A
		38	A	A	A	<B
		39	A	<B	A	A
		41	A	<C	A	A
		43	AA	AA	AA	>BB
		44	AA	AA	<BB	AA
7–8	Listen/Speaking	1	A	>C	<B	A
		2	A	>C	<C	>C
		6	A	<B	A	A
		8	A	A	C
		10	A	A	A	<B
		15	A	A	A	<B
		19	A	A	A	>C
		25	A	A	A	<B
		33	AA	AA	AA	>BB

J.1 Results of DIF Analyses (continued)

Grade Span	Modality	Item Seq. Number	DIF– Male/Female	DIF– Asian/Other	DIF– Hispanic/Other	DIF– White/Other
		38	AA	>BB	AA	AA
7–8	Reading/Writing	1	A	A	C
		2	A	A	C
		3	A	A	<C	>C
		6	A	>C	<C	>B
		14	A	A	A	>B
		20	A	B	A
		25	A	A	A	>B
		29	A	A	A	>B
		30	A	<B	A	>C
		33	A	A	A	<B
		34	A	<B	A	A
		35	A	<C	A	A
		36	A	A	A	<B
		40	A	>C	<C	A
		42	A	A	A	<B
43	AA	AA	AA	>CC		
44	AA	AA	AA	>BB		
45	AA	AA	AA	>BB		
9-12	Listen/Speaking	3	A	A	A	>B
		4	A	A	A	<B
		10	A	A	A	<B
		15	A	<B	A	A
		19	A	A	A	<B
		25	A	A	A	>B
		29	AA	>BB	AA	AA
9-12	Reading/Writing	1	A	A	A	>C
		2	A	>C	<C	A
		3	A	>C	<C	A
		6	A	<B	A	A
		14	A	A	A	>B
		17	A	A	B
		18	A	>B	A	A
		21	A	<B	A	A
		26	A	<B	A	A
		28	A	B	A
		29	A	>C	<C	A
		30	A	<C	>B	A
		32	A	<C	A	A
		33	A	>B	A	A
		35	A	>B	A	A
		37	A	A	A	<B
		39	A	<B	A	A
		40	A	A	A	<B
42	A	>C	A	<B		
45	AA	>BB	AA	AA		
46	<BB	AA	AA	AA		

Note: > indicates in favor of Males or Other groups while < indicates in favor of the focal groups (i.e., Female, Asian, Hispanic, and White).

APPENDIX K: REFERENCES

- Altman, D. G. (1991). *Practical statistics for medical research*. Chapman and Hall/CRC. 404.
- Andrich, A. & Luo, G. (2004). *Modern measurement and analysis in social science*. Perth, Western Australia: Murdoch University.
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, D.C.: American Psychological Association.
- Browne, M. W. (1982). "Covariance structures." In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis*. Cambridge: Cambridge University Press. 72–141.
- Camilli, G. & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando, FL: Holt, Rinehart and Winston.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16:297–334.
- Embretson, S. & Reise, S. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Haertel, E. H. (1996). "Estimating the decision consistency from a single administration of a performance assessment battery." *A Report on the National Board of Professional Teaching Standards McGEN Assessment*. Palo Alto, CA: Stanford University.
- Hanson, B. A. & Beguin, A. A. (1999). *Obtaining a common scale for IRT item parameters using separate versus concurrent estimation in the common item nonequivalent groups equating design*. Paper presented at the meeting of the National Council on Measurement in Education in Montreal.
- Hanson, B. A. (1995). *USmooth: A program for smoothing univariate test score distribution (Version 1.5)*. Iowa City, IA: American College Testing.
- Hanson, B. A. (1991). *Method of moments estimates for the four-parameter beta compound binomial model and the calculation of classification consistency indexes*. *ACT Research Report 91–5*. Iowa City, IA: American College Testing.

- Hu, M. & Bentler, P. M. (1999). "Cut off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives." *Structural Equation Modeling* 6:1–55.
- Huynh, H. (2000). Robust linking for PACT assessments. *South Carolina PACT Technical Report*.
- Jöreskog, K. G. & Sörbom, D. (1995a). *Liseral 8 user's reference guide*. Chicago: Scientific Software, Inc.
- Jöreskog, K. G. & Sörbom, D. (2006). *Liseral 8.8 for windows/computer software*. Lincolnwood, IL: Scientific Software International, Inc.
- Kolen, M. J. & Brennan, R. L. (1995). *Test equating methods and practices*. New York: Springer-Verlag.
- Kolen, M. J. & Brennan R. L. (2004). *Test equating, scaling, and linking: methods and practices* (2nd ed.). New York: Springer-Verlag.
- Linacre, J. M. & Wright, B. D. (2000). *A user's guide to WINSTEPS: Rasch-Model computer program*. Chicago: MESA Press
- Linacre, J. M. (2003). *A user's guide to WINSTEPS: MINISTEP: Rasch-Model computer program*. Chicago: MESA Press.
- Livingston, S. A. & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement* 32:179–1987.
- Lord, F. M. & Wingersky, M. S. (1984). Comparison of IRT true-score and equipercentile observed-score "equatings." *Applied Psychological Measurement* 8:452–461.
- MacCallum, R. C., Browne, M. & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods* 2(1):130–149.
- Mantel, N. & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* 22:719–748.
- Mantel, N. (1963). Chi-square tests with one degree of freedom: Extension of the Mantel-Haenszel procedure. *Journal of the American Statistical Association* 58:690–700.
- Masters, G. N. (1982). "A Rasch model for partial credit scoring." *Psychometrika* 47:149–174.
- Mitzel, H. C., Lewis, D. M. & Green, D. R. (2001). The bookmark procedure. In G. J. Cizek (Ed.), *Setting performance standards: concepts, methods, and perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates. 249–282.

- Nitko, A. J. (2004). *Educational assessment of students* (4th Ed.). Upper Saddle River, NJ: Pearson Education Inc.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin* 86(2):420–428.
- Steinley, D. & Wood, P. (2000). *Icc.sas—a program to calculate intraclass correlations and confidence intervals*. Columbia, MO: University of Missouri.
- Swygert, K. A., McLeod, L. D. & Thissen, D. (2001). Factor analysis for items or tests scored in more than two categories. In D. Thissen and H. Wainer (Eds.), *Test scoring*. Mahwah, NJ: Lawrence Erlbaum Associates. 217–250.
- Taherbhai, H. & Seo, D. (2007). Comparing concurrent vs. fixed parameter equating with common items, Using the Rasch dichotomous and the partial credit models in a mixed-item format test. *Journal of Applied Measurement* 8(1).
- Young, M. J. & Yoon, B. (1998, April). Estimating the consistency and accuracy of classifications in a standards-referenced assessment. *CSE technical report 475*. Los Angeles: University of California, Center for the Study of Evaluation, Standards, and Student Test.