New York State English as a Second Language Achievement Test (NYSESLAT)

2012 Operational Test Technical Report

Submitted by: Questar Assessment, Inc.
To: The New York State Education Department

April 12, 2013
Revised October 8, 2014

QUESTAR ASSESSMENT, INC.
Success By Your Standards®

5550 Upper 147th Street West
Apple Valley, MN 55124
(952) 997-2700
www.questarAI.com

TABLE OF CONTENTS

OVERVIEW 1

CHAPTER 1: INTRODUCTION 3

1.1 Background 3
1.2 Rationale and Purpose 3
1.3 Test Use 4
1.4 Test Accommodations 4
Large Type 4
Braille 4

CHAPTER 2: TEST DESIGN AND DEVELOPMENT 5

2.1 Test Blueprint and Specifications 5
2.2 Item Mapping and Specifications by New York State ESL Learning Standards 8
2.3 Item Development and Review 11
2.4 Field Test 11
2.5 Test Construction 12
Selecting Anchor Items 12
Selecting New Items 12
Selecting Research Items for K–1 and 2–4 13
Building the Item Selection Spreadsheets 13
Item Sequencing 13
Psychometric Review 13
Test Construction Meeting 14

CHAPTER 3: SCORING 15

3.1 Scoring Audit: 10% of the Operational Writing Responses 15
3.2 Scoring the Speaking Field Test Responses 15
3.3 Rangefinding 16

CHAPTER 4: CLASSICAL ITEM-LEVEL STATISTICS 17

4.1 Item-Level Descriptive Statistics 17
4.2 Differential Item Functioning (DIF) 20
Summary of the DIF Analyses 22

CHAPTER 5: RELIABILITY 24

5.1 Internal Consistency Reliability 24
5.2 Standard Error of Measurement (SEM) Based on Classical Test Theory 24
5.3 Standard Error of the Mean (SEMn) 25
5.4 Conditional Standard Error of Measurement (CSEM) Based on Item Response Theory 26
5.5 Summary of the Descriptive and Reliability Statistics 27
5.6 Inter-Rater Reliability 32
10% Scoring Audit 32
Intraclass Correlation 33
Kappa Coefficient 34
Summary of the 10% Audit Inter-Rater Reliability 37
5.7 Accuracy and Consistency of Performance Level Classifications 37

CHAPTER 6: VALIDITY 41
6.1 Content Validity 41
6.2 Internal Structure 42
Validity of the Scoring Structure via Confirmatory Factor Analyses (CFAs) 46
6.3 External Structure 47
Relationship with the New York State English Language Arts (ELA) Test 48
Proficiency Level Comparisons 50
Summary 53
Relationship with the New York State Regents Comprehensive Examination in English 54
Proficiency Level Classification 56
Summary 57
CHAPTER 7: CALIBRATION, EQUATING, AND SCALING 58
7.1 Item Response Model and Rationale for Use 58
7.2 Evidence of Model Fit 58
7.3 Description of the Calibration Sample and Process 59
7.4 Equating of the NYSESLAT 60
Common Item Linking Design 61
Anchor Item Evaluation 61
Stability of Anchor Items 61
Test Characteristic Curve (TCC) Comparison between 2011 and 2012 62
7.5 Scaling of the NYSESLAT 70
Raw Score-to-Scale Score Tables 70
7.6 Rasch Information 70
CHAPTER 8: STANDARD SETTING 72
8.1 Standard Setting Model 72
8.2 Committees of Panelists 72
8.3 Performance Levels and Cut Scores 73
8.4 Standard Setting Process 73
Three Rounds of Ratings 74
8.5 Summary Statistics for the Three Rounds of Ratings 75
8.6 Post-Standard Setting Analyses 75
8.7 Final Performance Level Cut Points 75
CHAPTER 9: SUMMARY OF THE OPERATIONAL TEST RESULTS 77
APPENDIX A: ITEM MAP 83
APPENDIX B: ITEM-LEVEL STATISTICS 118
APPENDIX C: RAW SCORE-TO-SCALE SCORE CONVERSIONS 128
APPENDIX D: IRT STATISTICS 143
APPENDIX E: ITEM INFORMATION AT DIFFERENT CUT POINTS 153
APPENDIX F: CONFIRMATORY FACTOR ANALYSIS (CFA) 179
APPENDIX G: SCALE SCORE SUMMARY BY SUBGROUP 188
APPENDIX H: PROFICIENCY PERCENTAGES BY SUBGROUP 214
APPENDIX I: EXIT RATE BY SUBGROUP 253
APPENDIX J: DIF STATISTICS 266
APPENDIX K: REFERENCES 269

LIST OF TABLES

Table 1—NYSESLAT 2012 Operational Test Blueprint 5
Table 2—2012 Test Specifications by Modality and Grade Band 7
Table 3—Maximum Number of Points by Modality and Grade Band 7
Table 4—Test Specifications for ESL Learning Standards by Grade Band 8
Table 5—Item Mapping by ESL Learning Standards and Performance Indicators 9
Table 6—N-Count by Grade Band 17
Table 7—Summary of Classical Item Difficulty and Item Discrimination Indices by Grade Band 19
Table 8—DIF Classification for Multiple-Choice Items 21
Table 9—DIF Classification for Constructed-Response Items 22
Table 10—Results of DIF Analyses for the 2012 Operational Test Items 23
Table 11—Descriptive Statistics and Reliability by Grade Band and Modality 28
Table 12—Descriptive Statistics and Reliability by Grade and Modality 29
Table 13—Audit Sample 33
Table 14—Rater Agreement for Pre-Writing and Writing Prompts 35
Table 15—Percentages of Score Difference Between Raters 36
Table 16—Comparison Between Local and Audit Raters 37
Table 17—Classification Accuracy and Consistency by Grade 40
Table 18—Intercorrelation Among the Modalities by Grade Band 43
Table 19—Intercorrelation Among the Modalities by Grade 43
Table 20—Global Fit Indices for the One- and Two-Factor Models (Grades K–1) 46
Table 21—Global Fit Indices for the One- and Two-Factor Models (Grades 2–4) 46
Table 22—Global Fit Indices for the One- and Two-Factor Models (Grades 5–6) 46
Table 23—Global Fit Indices for the One- and Two-Factor Models (Grades 7–8) 47
Table 24—Global Fit Indices for the One- and Two-Factor Models (Grades 9–12) 47
Table 25—Descriptive Statistics of the NYSESLAT Raw Scores of All Modalities and Their Correlation with the ELA Test Total Scale Scores 49
Table 26—Descriptive Statistics of the ELA Test and NYSESLAT Combined Scale Scores of Listening/Speaking and Reading/Writing 50
Table 27—Grade 3: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level 51
Table 28—Grade 4: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level 51
Table 29—Grade 5: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level 52
Table 30—Grade 6: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level 52
Table 31—Grade 7: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level 53
Table 32—Grade 8: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level 53

Table 33—Descriptive Statistics of the NYSESLAT Modalities' Raw Scores and Their Correlations with the Regents Comp Exam in English Total Scale Scores 54
Table 34—Descriptive Statistics of the NYSESLAT Modalities' Scale Scores and Total Scale Scores and Their Correlations with the Regents Comp Exam in English Total Scale Scores 55
Table 35—Grade 10: Percentage of ELL Students Passing/Failing the Regents Comp Exam in English at Each NYSESLAT Proficiency Level 56
Table 36—Grade 11: Percentage of ELL Students Passing/Failing the Regents Comp Exam in English at Each NYSESLAT Proficiency Level 56
Table 37—Grade 12: Percentage of ELL Students Passing/Failing the Regents Comp Exam in English at Each NYSESLAT Proficiency Level 57
Table 38—Common Items as Anchor Items by Modality Combination and Grade Band 62
Table 39—Average Rasch Difficulty by Grade Band and Modality 71
Table 40—Panel Composition for Standard Setting Committees 73
Table 41—Final Performance Level Cut Points for Listening/Speaking 76
Table 42—Final Performance Level Cut Points for Reading/Writing 76
Table 43—Raw Score Summary by Grade 78
Table 44—Raw Score Summary by Grade Band 78
Table 45—Scale Score Summary by Grade 79
Table 46—Scale Score Summary by Grade Band 79
Table 47—Percentage of Students in Each Proficiency Level by Grade 80
Table 48—Percentage of Students in Each Proficiency Level by Grade Band 81
Table 49—Exit Rate by Grade for the 2010, 2011, and 2012 Administrations 81
Table 50—Exit Rate by Grade Band for the 2010, 2011, and 2012 Administrations 82

LIST OF FIGURES

Figure 1—Classification Accuracy 38
Figure 2—Classification Consistency 38
Figure 3—2012 TCC for Listening/Speaking, Grades K–12 63
Figure 4—2012 TCC for Reading/Writing, Grades K–12 64
Figure 5—TCC for Listening/Speaking, Grade Band K–1 65
Figure 6—TCC for Listening/Speaking, Grade Band 2–4 65
Figure 7—TCC for Listening/Speaking, Grade Band 5–6 66
Figure 8—TCC for Listening/Speaking, Grade Band 7–8 66
Figure 9—TCC for Listening/Speaking, Grade Band 9–12 67
Figure 10—TCC for Reading/Writing, Grade Band K–1 67
Figure 11—TCC for Reading/Writing, Grade Band 2–4 68
Figure 12—TCC for Reading/Writing, Grade Band 5–6 68
Figure 13—TCC for Reading/Writing, Grade Band 7–8 69
Figure 14—2011 TCC for Reading/Writing, Grade Band 9–12 69

OVERVIEW

The following technical report for the 2012 operational New York State English as a Second Language Achievement Test (NYSESLAT) has nine chapters:

1. Introduction
2. Test Design and Development
3. Scoring
4. Classical Item-Level Statistics
5. Reliability
6. Validity
7. Calibration, Equating, and Scaling
8. Standard Setting
9. Summary of the Operational Test Results

Chapter 1: Introduction

Chapter 1 contains background information on the NYSESLAT, including the rationale, purpose, and recommended use of the assessment. It also describes the available test accommodations with information on the Large Type and Braille versions of the test.

Chapter 2: Test Design and Development

Chapter 2 details the development process of the assessment, including the test blueprint and specifications; item mapping; a description of the item development and review processes; field testing of the items; and test construction.

Chapter 3: Scoring

Chapter 3 describes the scoring audit conducted on a 10% sample of the operational Writing responses, as well as the scoring of the Speaking field test responses. It also includes a description of the 2012 rangefinding meeting in Albany, NY.

Chapter 4: Classical Item-Level Statistics

Chapter 4 presents the item-level descriptive statistics based on Classical Test Theory (CTT), including the p-value, point-biserial correlation, and Differential Item Functioning (DIF).

Chapter 5: Reliability

Chapter 5 relays reliability information of the assessment, including internal consistency reliability, the classical standard error of measurement (SEM), and the conditional standard error of measurement (CSEM) based on Item Response Theory (IRT). It also provides results of the inter-rater reliability of the 10% scoring audit, as well as the reliability of the performance level classifications.

Chapter 6: Validity

Chapter 6 describes the collected evidence of validity based on the test's content, internal structure, and external structure, which examines the relationships between the NYSESLAT and both the New York State English Language Arts (ELA) Test and the Regents Comprehensive Examination in English (Regents Comp Exam in English).

Chapter 7: Calibration, Equating, and Scaling

Chapter 7 explains the rationale for using the Rasch IRT model, contains test characteristic curves (TCCs) for the 2011 and 2012 administrations, and describes the calibration, equating, and scaling processes. Finally, this chapter includes summaries of the IRT model fit statistics and the average Rasch difficulty of the four subtests (Listening, Speaking, Reading, and Writing).

Chapter 8: Standard Setting

Chapter 8 describes the standard setting process followed by the previous vendor, which established performance level cut points. It includes the standard setting model, the standard setting process, summary statistics for the round-by-round ratings, post-standard setting analyses, and final performance level cut points as reported in previous technical reports.

Chapter 9: Summary of the Operational Test Results

Chapter 9 contains the raw score and scale score summaries, as well as the percentage of students in each performance category (Listening/Speaking and Reading/Writing). It also contains exit rates for the 2010, 2011, and 2012 administrations of the NYSESLAT.

CHAPTER 1: INTRODUCTION

1.1 Background

The New York State Education Department (NYSED) administers the New York State English as a Second Language Achievement Test (NYSESLAT) to English Language Learner (ELL) students in Grades K–12 in compliance with the federal No Child Left Behind Act of 2001 (NCLB). This technical report details the 2012 operational (OP) administration of the NYSESLAT, carried out by Questar Assessment, Inc. (Questar).

Title III of NCLB requires annual assessment of the English language skills of ELL students. According to section 3121(d)(1), 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" (No Child Left Behind [NCLB], 2002). 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." NCLB also requires that the annual assessment of ELL students be based on specific student achievement objectives. Section 3113(b)(2) states that the "agency will establish standards and objectives for raising the level of English proficiency that are derived from the four recognized domains of speaking, listening, reading, and writing, and that are aligned with achievement of the challenging State academic content and student academic achievement standards described in [Title I]" (NCLB, 2002).

NYSED developed the Learning Standards for English as a Second Language to meet the requirements of NCLB. These English as a Second Language (ESL) learning standards and performance indicators—i.e., achievement objectives—for ELL students derive from the Speaking, Listening, Reading, and Writing modalities; align with the state's English Language Arts standards; and are organized into four grade bands: K–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 ESL learning standards. Accordingly, the NYSESLAT was developed with four subtests—Speaking, Listening, Reading, and Writing—in each of the following five grade bands: K–1, 2–4, 5–6, 7–8, and 9–12. Individual test items match to specific ESL learning standards and performance indicators, and 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, as well as other applicable federal and state requirements.

1.2 Rationale and Purpose

The New York State Board of Regents established the ESL learning standards for all ELL students attending New York State schools. Based on NCLB mandates and because each student's annual progress toward proficiency must be tracked, all ELL students in Grades K–12 are assessed every year to measure academic English language proficiency in Listening, Speaking, Reading, and Writing. The NYSESLAT helps schools to determine which instructional standards teachers must devote time in order to ensure that their ELL students fully acquire the language proficiency that will prepare them for success in the classroom. The purpose of the NYSESLAT is to measure annual student improvement in achieving English

language proficiency in order for students to ultimately exit ELL status and move into English mainstream programs.

1.3 Test Use

The NYSESLAT is used when making decisions regarding language instructional programs and accountability determinations. First, the test measures the level of English proficiency of all ELL students in Listening, Speaking, Reading, and Writing. School districts then use this data to decide the type and amount of instructional services to which the students are entitled. Second, the NYSESLAT measures students' annual progress toward learning English and attaining academic 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

The test is consistent with the principles of Universal Test Design, which means it is as accessible as possible to all populations. 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, as well as provide 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, Questar utilized in-house content and fairness experts to ensure that the forms were pulled with concepts of Universal Test Design in mind. Questar stringently reviewed forms for special populations, such as visually or hearing-impaired students, to ensure that items were fair, reliable, and accessible to all.

Large Type

Questar applied large-type product specifications to ease the test-taking experience for visually impaired students from every grade band who required large type. Pages were printed in black only and on a cream-colored, non-glare vellum stock to ease readability of pages. Covers were printed on heavier stock to provide stiffness to the booklets and to protect the interior text pages.

Braille

Questar produced a Braille version of the NYSESLAT for visually impaired children, using certified and experienced transcribers who could work with the multiple codes, rules, and guidelines of Braille. Braille forms were available for each NYSESLAT subtest and grade band, although a checklist, rather than a Braille test, was provided for Grades K–1.

If a test question was difficult to Braille, Questar and content specialists determined if the construct could be worded or measured in other ways. To adapt some items, pictures were described, and care was taken not to convey the correct answer, but to give a description that would enable the student to ascertain the correct answer.

CHAPTER 2: TEST DESIGN AND DEVELOPMENT

2.1 Test Blueprint and Specifications

The NYSESLAT is divided into four modalities—Listening, Speaking, Reading, and Writing—for each of the five grade bands: K–1, 2–4, 5–6, 7–8, and 9–12. These subtests have either multiple-choice (MC) or constructed-response (CR) items, or a mixture of both; some of the constructed-response items are short spoken response (SS), and some are extended response (ER). The number of items varies by grade band: K–1 has 70 items, 2–4 has 80 items, and the remaining grade bands (5–6, 7–8, and 9–12) each have 87 items.

Across all grade bands, the Speaking modality has 16 constructed-response items. The Listening and Reading modalities are assessed solely by the use of multiple-choice items. Listening has 24 or 25 items depending on the grade band, and Reading has between 15 and 27 items depending on the grade band. Finally, Writing has between 15 and 19 items depending on the grade band and consists of the following parts:

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

The NYSESLAT consists entirely of items developed specifically for the assessment. The majority of items on the 2012 NYSESLAT were developed in 2010 and field tested in Spring 2011. Table 1 presents the operational test blueprint for the 2012 administration of the NYSESLAT. Table 2 shows the test specifications by modality and grade band, and Table 3 shows the maximum number of points by modality and grade band. Table 4 provides a breakdown of the number of items based on the ESL learning standards and grade bands.

Table 1—NYSESLAT 2012 Operational Test Blueprint

Strand	Cluster	Number of Items					Notes
		K–1	2–4	5–6	7–8	9–12	
Listening	First Sound	2					All items in this strand are multiple-choice items with 0–1 score points.
	Word/Sentence Comprehension	12	5	3	3	3	
	Comprehension of Conversational Language	10	14	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**	

Table 1 (continued)—NYSESLAT 2012 Operational Test Blueprint

Strand	Cluster	Number of Items					Notes
		K–1	**2–4**	**5–6**	**7–8**	**9–12**	
Reading	Word Reading	11					All items in this strand are MC 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**	
Writing	Phonemic Understanding	6	4	2	2	2	All items in this section (Writing Conventions) are MC items with 0–1 score points.
	Mechanics & Structure		8	13	13	13	
	Developmental Writing	8					Grade Band K–1 has 8 CR items with 0–2 score points.
	Pre-Writing		3	3	3	3	Grade Bands 2–4, 5–6, 7–8, and 9–12 each have 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 Grade Band K–1 has 0–2 score points. The ER items for Grade Bands 2–4, 5–6, 7–8, and 9–12 each have 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 CR items with 0–2 score points in this cluster.
	Storytelling	1	1	1	1	1	Each of the five levels has 1 CR item with 0–4 score points in this cluster.
	Picture Description	5	5	5	5	5	Each of the five levels has 5 CR items with 0–2 score points in this cluster.
	Social Interaction	5	5	5	5	5	Each of the five levels has 5 CR 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—2012 Test Specifications by Modality and Grade Band

Number of Items and Passages in the NYSESLAT Subtests							
				Writing			
	Speaking	Listening	Reading	Writing Conventions	Pre-Writing	Writing Prompt	Total Number of Items per Grade Band
Grade Band	CR	MC	MC/Passages	MC	CR	CR	
K–1	16	24	15/0	6	0	8 short CR (Dev. Writing) + 1 ER	70
2–4	16	24	24/4	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 Band

Maximum Number of Points in the NYSESLAT Subtests							
				Writing			
	Speaking	Listening	Reading	Writing Conventions	Pre-Writing	Writing Prompt	Total Number of Points per Grade Band
Grade Band	CR	MC	MC/Passages	MC	CR	CR	
K–1	34	24	15/0	6	0	16 short CR (Dev. Writing) + 2 ER	97
2–4	34	24	24/4	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 ESL Learning Standards by Grade Band

Grade Band	Standard*	Number of Items	Percentage of Total Items	Number of Score Points	Percentage of Total Points
K–1	1	42	60	52	54
	2	8	11	13	13
	3	6	9	9	9
	4	9	13	15	15
	5	5	7	8	8
	Total Test	**70**	**100**	**97**	**100**
2–4	1	39	49	45	43
	2	17	21	21	20
	3	8	10	15	14
	4	13	16	19	18
	5	3	4	4	4
	Total Test	**80**	**100**	**104**	**100**
5–6	1	43	49	45	41
	2	13	15	18	16
	3	10	11	17	15
	4	18	21	27	24
	5	3	03	4	04
	Total Test	**87**	**100**	**111**	**100**
7–8	1	33	38	35	32
	2	22	25	28	25
	3	17	20	26	23
	4	13	15	19	17
	5	2	2	3	3
	Total Test	**87**	**100**	**111**	**100**
9–12	1	32	37	34	31
	2	18	21	27	24
	3	17	20	24	22
	4	16	18	22	20
	5	4	5	4	4
	Total Test	**87**	**100**	**111**	**100**

*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.

2.2 Item Mapping and Specifications by New York State ESL Learning Standards

Table 5 provides an item-mapping summary, and Appendix A.1 provides, in detail, the item mapping for the 2012 NYSESLAT by ESL learning standards for each item within each grade band and modality. Appendix A.2 provides further item specifications.

Table 5—Item Mapping by ESL Learning Standards and Performance Indicators

Grade Band	Modality	Standard	Performance Indicators	Items in Subtest	Percentage of Subtest Score
K–1	Speaking	Standard 1	PI 9	3, 6	12
		Standard 2	PI 5, 7, 8	2, 7, 10	24
		Standard 3	PI 3, 9	4, 11	12
		Standard 4	PI 2, 3	1, 12–15, 18	35
		Standard 5	PI 3, 4	8, 9, 17	18
	Listening	Standard 1	PI 1, 4, 12	1–5, 7, 9, 13, 14, 16, 19–21	54
		Standard 2	PI 2	8, 10, 15, 18	17
		Standard 3	PI 1, 9	6, 22	8
		Standard 4	PI 7	17, 26, 28	13
		Standard 5	PI 2	23, 27	8
	Reading	Standard 1	PI 1, 3, 11, 16	1–8, 12–16	87
		Standard 2	PI 2	20	7
		Standard 3	PI 5	21	7
	Writing	Standard 1	PI 12	1–14	92
		Standard 3	PI 1	15	8
2–4	Speaking	Standard 1	PI 6, 9	2–4, 9, 10, 12	35
		Standard 2	PI 8	7, 11	18
		Standard 3	PI 3	16	6
		Standard 4	PI 2, 3, 9, 10	1, 6, 13–15, 18	35
		Standard 5	PI 4	8	6
	Listening	Standard 1	PI 1, 3, 4, 6, 16	2, 3, 6, 7, 9, 10, 12, 15, 20, 23, 28	46
		Standard 2	PI 1, 2, 5	1, 8, 11, 17, 19	21
		Standard 4	PI 7	18, 21, 24–27	25
		Standard 5	PI 2, 3	16, 22	8
	Reading	Standard 1	PI 1, 3, 6	1–3, 7–10, 25, 27–29	46
		Standard 2	PI 2, 4, 5, 12	14–17, 19, 20–23, 26	42
		Standard 3	PI 5	18, 30	8
		Standard 4	PI 7	24	4
	Writing	Standard 1	PI 1, 12, 15	1–11	50
		Standard 3	PI 1	12–16	50
5–6	Speaking	Standard 1	PI 5, 12	3, 5	12
		Standard 2	PI 5, 8	1, 6, 11	24
		Standard 3	PI 3	9	6
		Standard 4	PI 2, 3, 4, 5	2, 4, 8, 10, 12–16	53
		Standard 5	PI 2	7	6
	Listening	Standard 1	PI 1, 3, 4, 6, 7	1, 2, 4, 5, 12, 16–19, 22, 24	44
		Standard 2	PI 2	7, 10	8
		Standard 3	PI 5	3	4
		Standard 4	PI 3, 7, 8	9, 11, 13–15, 20, 21, 23, 25	36
		Standard 5	PI 1, 2	6, 8	8
	Reading	Standard 1	PI 1, 3, 4, 6, 16	1–5, 7, 9–12, 19–23	56
		Standard 2	PI 4, 5, 12	14, 16–18, 24–27	30
		Standard 3	PI 1, 3, 5	6, 8, 13, 15	15
	Writing	Standard 1	PI 1, 12, 15	1–15	60
		Standard 3	PI 1	16–19	40

Table 5 (continued)—Item Mapping by ESL Learning Standards and Performance Indicators

Grade Band	Modality	Standard	Performance Indicators	Items in Subtest	Percentage of Subtest Score
7–8	Speaking	Standard 1	PI 5, 6	7, 9	12
		Standard 2	PI 8	1, 2, 4, 6	29
		Standard 3	PI 1, 9	5, 8, 10	18
		Standard 4	PI 2, 3, 5, 9	3, 12–16	35
		Standard 5	PI 3	11	6
	Listening	Standard 1	PI 1, 3, 6, 9	1, 8–10, 15, 19–25	48
		Standard 2	PI 2, 5	7, 12, 13	12
		Standard 3	PI 9	2, 3	8
		Standard 4	PI 7	4–6, 14, 16–18	28
		Standard 5	PI 1	11	4
	Reading	Standard 1	PI 4, 6, 10, 16	2, 11, 13, 15	15
		Standard 2	PI 1, 2, 4, 5, 12	4–8, 18–27	56
		Standard 3	PI 1, 3, 5, 9	1, 3, 9, 10, 12, 14, 16, 17	30
	Writing	Standard 1	PI 1, 12, 15	1–15	60
		Standard 3	PI 1	16–19	40
9–12	Speaking	Standard 1	PI 6, 12	1, 9	12
		Standard 2	PI 5, 8	2, 4, 6–8, 10, 11	47
		Standard 3	PI 9	5	6
		Standard 4	PI 3, 5, 8, 9	3, 12–16	35
	Listening	Standard 1	PI 1, 3, 6, 9, 16	2, 4, 5, 7, 13, 14, 20, 21, 23–25	44
		Standard 2	PI 5	3	4
		Standard 3	PI 9	1	4
		Standard 4	PI 3, 6, 7, 8	6, 10–12, 15–19, 22	40
		Standard 5	PI 1	8, 9	8
	Reading	Standard 1	PI 4, 6, 16	2, 14–16	15
		Standard 2	PI 3, 4, 5, 7, 10	4–8, 19–23	37
		Standard 3	PI 1, 2, 3, 5, 9	1, 3, 10–12, 17, 18, 24–27	41
		Standard 5	PI 3, 5	9, 13	7
	Writing	Standard 1	PI 6, 12, 15	1–15	60
		Standard 3	PI 1	16–19	40

2.3 Item Development and Review

All items field tested in 2011 for use on the 2012 operational NYSESLAT or field tested in 2012 for use on the 2013 operational NYSESLAT were written by professional writers with experience writing assessment content for ELL students. Passage and item writers received guidelines to assist them in developing the passages and items, and content specialists were available via phone and e-mail to answer any questions. Subsequently, assessment specialists reviewed the passages and items sent in by the item writers and assessed them with respect to specifications and guidelines, ensuring the following:

- Absence of bias and sensitive topics in passages
- Factual accuracy of informational passages
- Item soundness, which can be interpreted as item validity and 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 band
- Alignment to the intended ESL learning standard

Committees of New York State ESL teachers reviewed the reading passages as a next step in the development process. Passage Review for the 2012 field test was held in Albany, NY, on June 7–8, 2011. Fifty-five passages were presented to the review committees, and, of those passages, 50 ultimately appeared on the 2012 field test.

Item Review for the 2012 field test was held in Albany, NY, on August 29–31, 2011. New York State ESL teachers reviewed each test item, accepting only those items judged to be of acceptable quality and fairness to students from various ethnic and cultural backgrounds.

2.4 Field Test

The items and data from the 2012 field test administration of the NYSESLAT will be used for forms construction of the 2013 operational NYSESLAT. The field test forms were constructed using only the items that had been accepted as-is or with edits during Item Review. In addition, because the 2013 operational NYSESLAT will be administered to six, rather than five, grade bands, the field test forms were constructed for six grade bands: K, 1–2, 3–4, 5–6, 7–8, and 9–12.

NYSED reviewed and approved the 2012 field test forms once they were constructed. Committees of New York State teachers also reviewed the forms during the Final Eyes review held in Albany, NY, on January 10–12, 2012, to ensure that they were error-free.

A stand-alone field test was administered to a representative sample of ELL students in the state of New York from March 19–30, 2012. Two field test forms were administered for the Listening, Speaking, and Reading modalities for each grade band (K, 1–2, 3–4, 5–6, 7–8, and 9–12). Two field test forms were administered for the Writing modality for Grades K and 1–2, whereas three field test forms were administered for 3–4, 5–6, 7–8, and 9–12. The forms for each grade band were distributed among the sample population, with each student given only one field test form testing only one modality. Details for the 2012 field test sampling plan and statistics are discussed in the *2012 NYSESLAT Field Test Technical Report*.

2.5 Test Construction

Selecting Anchor Items

At the beginning of the forms construction process, a selection of items from the 2011 operational test was identified as an anchor set. Anchor items provide statistics that help equate tests from one year to the next, and this process allows the scaled scores from one year to be comparable to those of previous years. The goal was to have anchor items make up approximately 25% of the test. Anchor item selection criteria are described in Section 7.4.

Selecting New Items

The majority of the 2012 NYSESLAT consisted primarily of new items—i.e., items that had never been used operationally. Questar's Assessment Development team selected these items from the set of items field tested in 2011 by following the blueprint as well as additional specifications outlined by Questar's Research and Development team.

The selection criteria for the 2012 NYSESLAT operational items included the following: item content, skill measured, item difficulty, gender balance, and ethnic balance. Content considerations included having a variety of settings and activities represented in items; having a mix of people, animals, and objects represented; and ensuring that no content overlapped within a subtest or across subtests in the same grade band. Skills measured varied by subtest and by sections within a subtest. In addition, the means and standard deviations of the modality and total scores, as well as Rasch values, were used as reference in order to facilitate item selection and forms construction.

Questar used the following guidelines to ensure the quality of the assessment:

- Use the blueprint to select the items with no variation greater than +/- 5% from the specified ESL learning standard.
- Following item selection, evaluate the target criteria for all content.
- To the extent possible, use items from the most recent field test.
- If the blueprint is met, use items with no DIF flags, a good discrimination index (.25 or greater), and an acceptable p-value range (between .30 and .95).
- Avoid items with higher than acceptable omission rates (5%).
- Assess the accuracy of content classifications of items and their associated answer keys.
- Avoid cueing.
- Attempt to have an even distribution of answer keys.
- Ensure that the set of Reading passages includes at least one of each NYSESLAT passage type—i.e., story, functional, informational, and poem.

Selecting Research Items for K–1 and 2–4

Research items were selected from the 2–4 test form for use on the K–1 form and from the K–1 test form for use on the 2–4 form. The purpose of these research items was to provide data that would help support the change in 2013 from five to six grade bands. The off-grade research items did not count toward a student's score.

Building the Item Selection Spreadsheets

Questar created an Item Selection Spreadsheet (ISS) for each test form. The ISS included the item ID number and item statistics associated with the item such as item type, prior usage, p-value, point biserial correlation coefficient, and score point value.

In building the ISS, Questar content specialists verified the following:

- The correct number of core, anchor, and research items (if applicable) was selected for each form.
- All required ISS fields were populated with accurate information and data.
- There was a balance of correct answer choices.
- There was a balance in the key—e.g., no more than three in a row of the same answer choice.
- The selected items reflected the appropriate range in terms of item type, content, level of difficulty, gender and ethnicity balance, etc.
- Score points were accurate.
- The items worked together in the form—e.g., no clueing of correct answers.

Item Sequencing

Questar specialists recommended the item positioning based on the following principles:

- Anchors should remain as close as possible to their previous administration position, avoiding first and last positions if possible.
- Item keys should not be the same more than three times in a row.
- Passages should be as varied as possible to avoid similar passages or the most difficult ones being placed next to each other.
- General sequencing of difficulty, when possible, should begin with the easiest items progressing to medium and difficult and tailing off to medium or easy items.

Psychometric Review

Questar's Research Development Team conducted a psychometric review of the proposed item selection and gave feedback to their Assessment Development Team. Iterations between the two departments occurred as necessary for each grade band.

Statistical considerations included item difficulty, item discrimination, and potential bias—i.e., p-values, point biserial correlation coefficients, and Differential Item Functioning (DIF) statistics, respectively. (See Chapter 4 for more information on item-level statistics.) Staff members from Questar were given the following instructions during the item selection and forms construction process:

1. Item Statistics:
    a. Check the range of item difficulty: Items should be flagged if the p-value is < 0.30 or > 0.95.
    b. Check the point-biserial range: Try to avoid items with a point biserial < 0.30.
    c. Check the omit rate: Watch for items with an omit rate > 5%.
    d. 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 for content during the Test Construction meeting.
2. There are no changes to an item once it has been field tested.
3. Tests were built using the statistical target that the average p-values between the Spring 2012 form and the Spring 2011 form should be similar.
4. The total number of items at each level, the number of items within each modality, and the number of items for each item type must follow the test blueprint.
5. In general, flagged items should be avoided. However, the match to blueprint and content considerations should take priority over statistical targets in most circumstances.

Test Construction Meeting

Finally, the proposed item selection for each test form was reviewed for content and psychometric characteristics at a formal Test Construction meeting. The objective of this meeting was to finalize the item selection for the operational test forms. Questar staff met with staff from NYSED to finalize the 2012 NYSESLAT test forms on Oct. 31–Nov. 4, 2011, in Albany, NY. A detailed discussion of the 2012 test construction is provided in the *2012 NYSESLAT Test Construction Report*.

CHAPTER 3: SCORING

This chapter details the scoring processes for the operational test audit and the field test. All scoring of 10% of the operational Writing constructed responses (i.e., the scoring audit) and the constructed-response field test items for Speaking and Writing was conducted at Questar's scoring center in Apple Valley, MN. Experienced Questar scoring directors and team leaders conducted the training and monitoring.

All Questar NYSESLAT readers had a minimum of a B.A. or B.S. degree and were trained to score according to the appropriate rubric to ensure accurate, consistent, and reliable results. Questar adhered to stringent criteria in its general screening and training procedures as preliminary measures for obtaining high levels of consistency and reliability. Only readers with proven ELL scoring experience were selected for NYSESLAT scoring.

3.1 Scoring Audit: 10% of the Operational Writing Responses

To measure inter-rater reliability and ensure that local teachers apply the same rigorous scoring standards across the state as intended by NYSED, Questar performed an audit on approximately 10% of all Writing tests at the request of NYSED by computing rater agreement based on local readers' ratings and Questar readers' ratings. A merged data file was created based on the local ratings (i.e., the final 2012 NYSESLAT data provided by NYSED) and Questar ratings (i.e., the sample of Writing items scored by Questar raters). The data was merged using Student State ID and Student Local ID to ensure perfect student match and removal of duplicates. (See Chapter 5 for more information on inter-rater reliability.) For a detailed discussion about the 10% Audit of the operational Writing responses, please refer to the *NYSESLAT 10% Audit Report for 2012*.

3.2 Scoring the Speaking Field Test Responses

All student responses were recorded on audiocassettes during the administration of the Speaking field test. Questar's scoring director, team leaders, and trained readers were all involved in scoring the speaking field test responses. Readers were trained using the 2012 Operational Scoring Guide and the Exemplar CD for Speaking. In addition, supplemental materials were created and distributed to accommodate minor changes between the 2012 and 2013 scoring rubrics and new items. These materials included a revision document used in coordination with the 2012 Speaking Scoring Guide, which highlighted changes made to annotations of student responses that matched the new rubric language. *Response to Graphic Information* was a new item type introduced during the 2013 field test, and a mini guide was created and approved by NYSED to use for reader training of those items. The student responses and audio files used for this guide were compiled using 2012 student field test responses. If an answer could not be heard or was unclear because of a mechanical error or malfunction during recording, the response was given a non-scorable code.[1]

During scoring, readers identified possible exemplars to be used in future scoring guides. Team leaders reviewed the identified responses and selected five of each score point for each field test item to upload as ".wav" files for future use. From those five sample responses, Questar

[1] A non-scorable code was treated as a missing value for item analysis.

identified the best two from each score point. Transcripts and annotations for these items were compiled and saved. When the 2013 operational forms for Speaking were constructed, those responses were used to develop the new scoring guides and CDs for teacher use.

3.3 Rangefinding

Rangefinding consists of using students' constructed-response answers to the 2012 Writing field test to create a set, or range, of scores associated with the different levels of performance. Prior to scoring the field test items, Questar facilitated rangefinding in Albany, NY, on June 25–28, 2012, to review a selected subset of the 2012 field test Writing responses and determine scoring standards for future Writing scoring guides and training sets. The participants included the following:

- Two Questar program managers, one Questar Assessment Development content manager, two Questar Performance Assessment managers, and six Questar scoring directors (one for each grade band) participated throughout the entire session.
- Two NYSED representatives attended throughout the session, and additional NYSED representatives visited frequently during the week.
- Forty-eight committee members with a variety of experience, including classroom teachers, principals, and administrators, also participated (7–10 members per grade band).

To prepare for the rangefinding meeting, a team consisting of Questar's Performance Assessment manager and six scoring directors pulled 30–60 sample responses from several hundred randomly selected responses for each field test item to score during the rangefinding session. These sample papers covered a range of score points and represented a variety of issues that might arise when scoring student responses.

Committee members independently scored Writing responses for each item per grade band. Each booklet contained 30–60 responses ranging from high to low score points in random order. Members discussed the merits of each response indicating the recommendations for use in the 2013 operational Writing scoring guide or training sets. Following independent review, the Questar Performance Assessment Facilitator led a group review and discussion. Group discussion notes were kept in a master binder for NYSED review and record.

Upon completion of each item, the facilitator submitted the master scoring log and master items binder to the "review room/office" for review by NYSED representatives for final anchor selection approval. This process continued until all items were scored and completed. Scoring guides for the 2012 field-test scoring and 2013 operational Writing scoring guides and training sets were developed using the approved anchor responses from rangefinding.

CHAPTER 4: CLASSICAL ITEM-LEVEL STATISTICS

This chapter contains the summary statistics for the raw scores of all items administered on the Spring 2012 operational NYSESLAT forms. The data file used for this analysis was the 100% student file compiled by the NYSED data team and provided to Questar in July 2012. The summary statistics are based in Classical Test Theory (CTT) and include information such as the p-values, the point-biserial correlations, and DIF statistics.

The data file obtained from NYSED had 248,335 students. The following exclusion rules were applied:

1. Removal of off-level students, carried out by Questar's IT team (14,098 removed)
2. Removal of duplicate records (44 removed)
3. Removal of students with an invalid raw score in any modality (1,141 removed)
4. Removal of students from non-public schools (24,632 removed)

Table 6 contains the number of students (208,420) who were valid and analyzed in each test.

Table 6—N-Count by Grade Band

Grade Band	Final Sample Size
K–1	51,942
2–4	58,036
5–6	28,623
7–8	25,115
9–12	44,704
Total	208,420

4.1 Item-Level Descriptive Statistics

A p-value is an indication of an item's difficulty and, for multiple-choice items, it is shown as the proportion of students that answers an item correctly. Higher p-values indicate that the items are easy, while lower p-values indicate that the items are difficult. For constructed-response items, the p-value is reported as a proportion of the maximum number of possible points. Operational 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.

The point-biserial correlation for each item is a measure of discrimination—i.e., how well an item discriminates, or distinguishes, between low-ability and high-ability students. It is an index of the association between the item score and the total test score. Students who did well on the test should, in theory, select the right answer to any given item, and students who did poorly on the test should select the wrong answer to any given item. An item with a high point-biserial correlation is better at discriminating between low-ability and high-ability students than an item with a low point-biserial correlation. As a correlation of the item to the total test, typically a larger point biserial is viewed as desirable for item inclusion. For form building, items were selected using the point biserial in two ways:

1. The strength of the association with the underlying trait—unidimensionality (high correlation)
2. The degree of local independence of the item (low correlation)

A very high correlation, in fact, may mean the item is actually redundant or assesses essentially the same aspect of the construct. A very low correlation may demonstrate a looser relationship between the underlying trait and the item than is acceptable. For the NYSESLAT, the point-biserial correlation for items is constrained to an upper and lower limit. The upper limit of the point biserial is fixed at 0.80, while the lower limit is fixed at 0.25. Occasionally, an upper limit and lower limit violation occurs. The exclusion of any items with a negative point-biserial correlation is stringently observed because this may indicate a wrong key during the scoring process or that an item has inappropriate or overly attractive distractor(s). Items with negative point-biserial correlations are recommended for exclusion from the form-building process. The upper and lower limits should be viewed as subjective guidelines or flags for further investigation. Upon investigation of the response pattern and distractors, an item may occasionally need to be rekeyed, or it may be determined that, based on content needs, the item needs to be included and is correctly keyed.

Item-level statistics for the 2012 operational NYSESLAT are presented in Appendix B by grade band. With the exception of few high p-values and low point-biserial values, all items fell well within the pre-set level of acceptance, both in terms of the p-value and point biserial. The tables are grouped by Listening/Speaking and Reading/Writing modality combinations, and the following item information and statistics are presented for each item:

- Item number
- Item type: multiple-choice (MC), constructed-response (CR), short spoken response (SS), or extended-response (ER)
- 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)
- Item difficulty index (p-value)
- Item discrimination index (point biserial correlation coefficient)

Items that are too easy or too difficult are flagged based on their p-values because items like these provide inadequate information. The inclusion of items outside of the predefined difficulty range is typically beneficial to the measurement process, but because the NYSESLAT is a standard-referenced assessment, content experts may determine that the inclusion of specific items is necessary.

Table 7 presents the mean p-values and point-biserial correlation coefficients. The mean p-values and point-biserial correlations for each grade band in the combined modalities of Listening/Speaking and Reading/Writing are between the upper and lower limits for an acceptable range of these values.

Table 7—Summary of Classical Item Difficulty and Item Discrimination Indices by Grade Band

			Item Difficulty (p-value)		Item Discrimination (point-biserial)	
Grade Band	**Modality**	**N-Count**	**Mean**	**SD**	**Mean**	**SD**
K–1	Listening	51,942	0.88	0.07	0.43	0.06
	Speaking	51,942	0.80	0.07	0.71	0.05
	Reading	51,942	0.78	0.08	0.54	0.06
	Writing	51,942	0.78	0.15	0.58	0.17
	Listening and Speaking	51,942	0.85	0.08	0.54	0.15
	Reading and Writing	51,942	0.78	0.12	0.56	0.13
2–4	Listening	58,036	0.81	0.13	0.43	0.09
	Speaking	58,036	0.86	0.03	0.74	0.04
	Reading	58,036	0.71	0.16	0.48	0.08
	Writing	58,036	0.75	0.13	0.54	0.08
	Listening and Speaking	58,036	0.83	0.11	0.55	0.17
	Reading and Writing	58,036	0.73	0.15	0.51	0.09
5–6	Listening	28,623	0.79	0.09	0.48	0.10
	Speaking	28,623	0.86	0.04	0.76	0.05
	Reading	28,623	0.69	0.12	0.50	0.07
	Writing	28,623	0.75	0.12	0.51	0.12
	Listening and Speaking	28,623	0.81	0.08	0.59	0.16
	Reading and Writing	28,623	0.72	0.12	0.51	0.09
7–8	Listening	25,115	0.72	0.12	0.40	0.08
	Speaking	25,115	0.82	0.03	0.78	0.03
	Reading	25,115	0.63	0.13	0.45	0.08
	Writing	25,115	0.72	0.13	0.50	0.13
	Listening and Speaking	25,115	0.76	0.11	0.55	0.20
	Reading and Writing	25,115	0.67	0.14	0.47	0.11
9–12	Listening	44,704	0.67	0.12	0.45	0.06
	Speaking	44,704	0.79	0.05	0.75	0.03
	Reading	44,704	0.61	0.15	0.43	0.07
	Writing	44,704	0.66	0.13	0.52	0.11
	Listening and Speaking	44,704	0.72	0.12	0.57	0.16
	Reading and Writing	44,704	0.63	0.14	0.47	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 208,420 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 invalid codes). “Research” items are not included in Grades K–1 and 2–4.

4.2 Differential Item Functioning (DIF)

Differential Item Functioning (DIF) analyses are statistical procedures used to flag items for potential bias. The results of DIF analyses are useful in reviewing items and tests for potential bias, but they are not proof of bias—rather, they are an indication that items need to be further scrutinized. In general, DIF statistics are used to compute the probability that one demographic group is more likely to correctly answer an item than another group or subgroup after having controlled for group differences in ability—e.g., an item that seems to be easy for female students but not for male students. Again, the presence of DIF does not necessarily indicate that an item is biased, and, therefore, flagged items should be evaluated for bias indication.

For the NYSESLAT, a contingency table approach was used for analyzing DIF. The categories outlined in this report are a modification of what are commonly referred to as the ETS DIF categories, which are widely accepted in the field of psychometrics for the categorization of DIF severity:

- "A" indicates negligible DIF.
- "B" indicates mild to moderate DIF.
- "C" indicates extreme DIF.

For multiple-choice items, the Mantel-Haenszel Chi-square (M-H χ^2) and the Mantel-Haenszel odds ratio were transformed into what ETS has named the Delta Scale (D). To calculate delta, the odds ratio was calculated as the odds of a correct response or *P/Q* (i.e., *P/[1-P])*. The odds ratio was then calculated as the odds of a correct response for the reference group divided by the odds of a correct response for the focal group. Thus, the odds ratio is as follows:

$$\alpha_{M-H} = \frac{P_r / Q_r}{P_f / Qf} \tag{1}$$

The null hypothesis for DIF is that the odds of correctly answering the item are equal for the two groups—i.e., the odds ratio is equal to 1:

$$H_0\text{: } \alpha_{M-H} = \frac{P_r / Q_r}{P_f / Qf} = 1 \tag{2}$$

Customarily, the odds ratio is then log-transformed and made symmetrical around zero with a range being in the interval $-\infty$ to $+\infty$. It is then called the log odds ratio, as follows:

$$\beta_{M-H} = \ln(\alpha_{\text{M-H}}) \tag{3}$$

The natural logarithm transformation of the odds ratio is symmetrical around zero. A zero value means either group has equal odds of correctly answering the item. A positive DIF measure means the reference group is favored, while a negative value indicates DIF is in favor of the focal group.

β_{M-H} can also be transformed in a linear fashion in order to convert to an interval scale (Camilli & Shepard, 1994). ETS used this process to transform β_{M-H} to their Delta Scale metric (D) via the following equation:

$$\mathrm{D} = -2.35 \cdot \beta_{M-H}. \tag{4}$$

On this scale, the quantity that D represents is the relative difference in the difficulty of the items for the two groups whose performance is being compared. Table 8 depicts the DIF classifications for multiple-choice items based on the M-H χ^2 and the item Delta Scale difference value (D).

Table 8—DIF Classification for Multiple-Choice Items

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

For polytomously scored constructed-response items, the standardized mean difference (SMD) statistic was used in addition to the MH test. SMD was developed as an extension to the Mantel procedure (Dorans & Schmitt, 1991; Zwick, Donoghue, & Grima, 1993):

$$SMD = \sum_{s=0}^{k} (M_{Fs} - M_{Rs}) P_{Fs} \tag{5}$$

where M_{Fs} and M_{Rs} are the mean item scores for the focal and reference groups conditional on *s*, and P_{Fs} is the proportion of the focal group members conditional on *s*. In this study, the corrected total score is used as the matching variable.

A statistic similar to effect size (ES) is then obtained by dividing the SMD by the standard deviation (SD) for the combined group:

$$ES = SMD / SD \tag{6}$$

A positive SMD or ES value indicates that the focal group has a higher mean item score than the reference group conditional on the matching variable. A negative SMD or ES value indicates that the focal group has a lower mean item score than the reference group conditional on the matching variable. Table 9 outlines the DIF classification for constructed-response items.

Table 9—DIF Classification for Constructed-Response Items

Category	Description	Criterion
A	No DIF	Non-significant Mantel χ^2 or Significant Mantel χ^2 and $\lvert SMD/SD \rvert \leq .17$
B	Moderate DIF	Significant Mantel χ^2 and $.17 < \lvert SMD/SD \rvert \leq .25$
C	Large DIF	Significant Mantel χ^2 and $.25 < \lvert SMD/SD \rvert$

Note*:* SD is the total group standard deviation of the item score.

NYSESLAT DIF analyses were completed on both gender and ethnic groups. For the gender DIF analysis, female students were considered the focal group and male students were considered the reference group. Three ethnic DIF analyses were conducted: Asian, Hispanic, and White 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. The same focal and reference groups were used in the DIF analysis of field test items if enough n-counts were available for the comparison.

Appendix J provides the DIF statistics for the 2012 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 the DIF Analyses

Table 10 on the following page contains the summary of the DIF analyses of operational test items, including n-counts for each DIF analysis group and the number of items flagged by the MH procedure or SMD (see details in Tables 8 and 9).

Table 10—Results of DIF Analyses for the 2012 Operational Test Items

DIF Type	Grade Band	Modality	Number of Items: All Item	DIF Items	Moderate DIF Items	Large DIF Items
Male vs. Female	K–1	Listening/Speaking	40	0	0	0
		Reading/Writing	30	0	0	0
	2–4	Listening/Speaking	40	0	0	0
		Reading/Writing	40	0	0	0
	5–6	Listening/Speaking	41	0	0	0
		Reading/Writing	46	0	0	0
	7–8	Listening/Speaking	41	0	0	0
		Reading/Writing	46	1	0	1
	9–12	Listening/Speaking	41	0	0	0
		Reading/Writing	46	2	0	2
White vs. Non-White	K–1	Listening/Speaking	46	0	0	0
		Reading/Writing	36	0	0	0
	2–4	Listening/Speaking	46	2	0	2
		Reading/Writing	46	0	0	0
	5–6	Listening/Speaking	41	0	0	0
		Reading/Writing	46	2	0	2
	7–8	Listening/Speaking	41	0	0	0
		Reading/Writing	46	2	0	2
	9–12	Listening/Speaking	41	0	0	0
		Reading/Writing	46	3	0	3
Hispanic vs. Non-Hispanic	K–1	Listening/Speaking	46	6	5	1
		Reading/Writing	36	3	0	3
	2–4	Listening/Speaking	46	3	3	0
		Reading/Writing	46	2	2	0
	5–6	Listening/Speaking	41	1	1	0
		Reading/Writing	46	6	5	1
	7–8	Listening/Speaking	41	2	1	1
		Reading/Writing	46	2	0	2
	9–12	Listening/Speaking	41	3	0	3
		Reading/Writing	46	6	1	5
Asian vs. Non-Asian	K–1	Listening/Speaking	40	16	3	13
		Reading/Writing	30	7	5	2
	2–4	Listening/Speaking	40	8	2	6
		Reading/Writing	40	3	0	3
	5–6	Listening/Speaking	41	5	1	4
		Reading/Writing	46	7	1	6
	7–8	Listening/Speaking	41	6	2	4
		Reading/Writing	46	7	4	3
	9–12	Listening/Speaking	41	5	2	3
		Reading/Writing	46	9	3	6

Note: "Research" items are not included in Grades K–1 and 2–4. Detailed results are located in Appendix J.

CHAPTER 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, and it is one way to provide support for a test's reliability; the more consistent the scores are, the more reliable the test is.

Several methods can be used to estimate internal consistency. One method is the "split-half" approach, which splits all test items into two groups and then correlates student scores on the two half-tests. The advantage of this method is that it avoids any implications of changes in individual students by administering only a single test. The theory is that if scores have high correlations on the two half-tests, then the test items complement one another, function well as a group, and measure the same construct. In addition, a positive result—i.e., a high correlation—would suggest that measurement error is lower. The problem with the split-half method is that the decision about which items belong in which half of the test can have a large impact on the resulting correlation. Therefore, Questar, like many other vendors, prefers to use Cronbach's coefficient alpha statistic (Cronbach, 1951) to sidestep this problem.

Rather than elect specific items to put into two halves, coefficient alpha is the average split-half correlation based on all possible divisions of a test into two parts. Coefficient alpha has the additional advantage in that it can be used to estimate the internal consistency of both dichotomous and polytomous items. Coefficient alpha is computed using the following formula:

$$\alpha = \frac{I}{I-1}\left(1 - \frac{\sum_i^I s_i^2}{S_X^2}\right) \tag{7}$$

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.

Questar calculated Cronbach's coefficient alpha reliability statistic, found in Tables 11 and 12 in Section 5.5. Most of the reliability values are close to or more than 0.90, which is considered high. Therefore, the NYSESLAT is considered a reliable test.

5.2 Standard Error of Measurement (SEM) Based on Classical Test Theory

Because perfect measurement of ability does not exist, it is important to analyze the amount of measurement errors on an assessment. Psychometricians sometimes think about reliability of an assessment as the ability of a test to produce the same results consistently. For example, if a student repeatedly took the same assessment and theoretically was unable to remember its contents, that student should earn roughly the same score. Obviously, this cannot be done in reality, but the spread, or standard deviation, of these hypothetical scores is what is known as the standard error of measurement (SEM).

The SEM is another measure of an assessment's reliability and provides an estimate of the amount of error that exists in an individual's observed test score. A given individual's observed total score, from this perspective, is thought of as an estimate of the student's true score. The SEM is inversely related to the reliability of a test because the greater the reliability, the lower the SEM will be. Hence, an observed test score has more accuracy or precision when the SEM is small. Measurement error is commonly expressed in terms of standard deviations—i.e., the SEM can be thought of as the standard deviation (SD) of the distribution of measurement error. The SEM is calculated using the following equation:

$$SEM = SD\sqrt{1-r_{xx}} \;\Leftrightarrow\; s_e = s_x\sqrt{1-\frac{s_t^2}{s_x^2}} \tag{8}$$

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

5.3 Standard Error of the Mean (SEMn)

Yet another way to conceptualize reliability is to examine the standard error of the mean (SEMn), which is an estimate of the sampling error of the sample mean as an estimate of the population mean. Since a given administration can be thought of as all possible students that could take the NYSESLAT, a single administration can similarly be thought of as a sample of all possible tests takers. It is defined as follows:

$$SEMn = \frac{\sigma}{\sqrt{n}} \tag{9}$$

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 Standard Error of Measurement (CSEM) Based on Item Response Theory

In Classical Test Theory (CTT), the SEM is the same for all students. However, in Item Response Theory (IRT), the SEM will not be the same for every student. For example, if a student gets either a few or many items correct—i.e., an extreme score—the standard error is greater than if the student gets a moderate number of items correct. In effect, the SEM depends on the total score (Andrich & Luo, 2004). Under the Rasch model, the SEM for each student is as follows:

$$\sigma_{\hat{\beta}} = \frac{1}{\sqrt{\sum_{i=1}^{L} p_{vi}(1-p_{vi})}} \tag{10}$$

where v is the subscript for a person, i is the subscript for an item, L is the length of the test, $\hat{\beta}$ is the ability estimate, and p_{vi} is the probability that a student answers an item correctly. It is defined as follows:

$$P_{vi} = \frac{e^{\beta_v - \delta_i}}{1 + e^{\beta_v - \delta_i}} \tag{11}$$

where β_v is student v's ability, and δ_i is the difficulty of the item. A confidence band can be used to interpret the ability estimate. For example, an approximate 68% confidence interval for $\hat{\beta}$ is given by:

$$\hat{\beta} \pm SEM \tag{12}$$

The SEM is smallest when the probability of passing an item is close to the probability of failing the item. In other words, if an item's difficulty is closely matched to the ability of many students in the sample, the standard error is small (Embretson & Reise, 2000).

One reason for the popularity of IRT over CTT in recent years is its ability to differentiate between students and item measurements. In CTT, item measurements for all students are assumed equivalent in a given test. IRT models vary student and item performances, which is taken into account when estimating the reliability of the measurement (Crocker & Algina, 1986). Equation 8 and its results can fluctuate in terms of the standard errors of students because 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 students perform, and they are the greatest for estimates in the extreme. Hence, fewer and fewer students in the extremes lead to less precise estimates of ability. For this reason, IRT estimates individual standard errors at score points, known as conditional standard errors of measurement (CSEMs). Essentially, the standard error is conditional on theta (student ability). This is preferred to classical SEMs because the CSEM allows for a more precise differentiation of students at different levels of performance. Appendix C provides the CSEMs in raw score-to-scale score conversion tables. Appendix D shows the IRT statistics.

5.5 Summary of the Descriptive and Reliability Statistics

Table 11 and Table 12 provide descriptive statistics for the raw score and reliabilities by grade band and grade. The tables contain the following:

- Number of items
- Maximum number of possible points
- N-count (number of students)
- The raw score (RS) mean and standard deviation
- Mean p-value
- Standard error of the mean (SEMn)
- 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 greater reliability is associated 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 bands are between 0.84 and 0.96, which is a strong indication that the test forms were constructed in good quality. The SEMn values of Listening/Speaking and Reading/Writing across the five grade bands ranged from 0.01 to 0.03, which is small and, thus, acceptable. The SEM values of Listening/Speaking and Reading/Writing across the five grade bands ranged from 1.37 to 3.15, which are also acceptable. If an item has, for example, an SEM 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 all grades are above 0.79, which is considered moderately high. The SEMn and SEM values are small and within acceptable ranges for each grade.

Table 11—Descriptive Statistics and Reliability by Grade Band and Modality

Grade Band	Test	Number of Items	Max. Points	N-Count	Mean (RS)	SD (RS)	Mean P-value	SEMn	Reliability	SEM
K–1	Listening	24	24	51,942	21.10	3.67	0.88	0.02	0.86	1.37
	Speaking	16	34	51,942	26.80	7.32	0.80	0.02	0.94	1.79
	Reading	15	15	51,942	11.76	3.48	0.78	0.02	0.86	1.30
	Writing	15	24	51,942	18.18	5.35	0.78	0.04	0.88	1.85
	Listening and Speaking	40	58	51,942	47.90	10.24	0.85	0.01	0.94	2.51
	Reading and Writing	30	39	51,942	29.94	8.41	0.78	0.02	0.93	2.23
2–4	Listening	24	24	58,036	19.48	4.07	0.81	0.03	0.84	1.63
	Speaking	16	34	58,036	29.20	6.91	0.86	0.01	0.95	1.55
	Reading	24	24	58,036	17.06	5.05	0.71	0.03	0.87	1.82
	Writing	16	22	58,036	15.73	4.64	0.75	0.03	0.85	1.80
	Listening and Speaking	40	58	58,036	48.67	10.16	0.83	0.02	0.94	2.49
	Reading and Writing	40	46	58,036	32.79	9.18	0.73	0.02	0.92	2.60
5–6	Listening	25	25	28,623	19.64	5.18	0.79	0.02	0.89	1.72
	Speaking	16	34	28,623	29.06	7.37	0.86	0.01	0.96	1.47
	Reading	27	27	28,623	18.71	6.25	0.69	0.02	0.90	1.98
	Writing	19	25	28,623	18.12	5.28	0.75	0.03	0.86	1.98
	Listening and Speaking	40	59	28,623	48.70	11.67	0.81	0.01	0.95	2.61
	Reading and Writing	46	52	28,623	36.83	10.89	0.72	0.02	0.93	2.88
7–8	Listening	25	25	25,115	18.12	5.00	0.72	0.02	0.85	1.94
	Speaking	16	34	25,115	27.69	8.48	0.82	0.01	0.96	1.70
	Reading	27	27	25,115	17.04	5.96	0.63	0.03	0.87	2.15
	Writing	19	25	25,115	17.27	5.51	0.72	0.03	0.86	2.06
	Listening and Speaking	41	59	25,115	45.81	12.39	0.76	0.02	0.95	2.77
	Reading and Writing	46	52	25,115	34.31	10.67	0.67	0.02	0.92	3.02
9–12	Listening	25	25	44,704	16.69	5.74	0.67	0.02	0.88	1.99
	Speaking	16	34	44,704	26.87	8.49	0.79	0.01	0.96	1.70
	Reading	27	27	44,704	16.34	5.78	0.61	0.03	0.86	2.16
	Writing	19	25	44,704	16.42	6.12	0.66	0.03	0.87	2.21
	Listening and Speaking	41	59	44,704	43.56	13.09	0.72	0.02	0.95	2.93
	Reading and Writing	46	52	44,704	32.76	11.13	0.63	0.02	0.92	3.15

Note: The data file used in this table is the final research file. The total n-counts for Grades K–12 are 208,420. 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 invalid codes).

Table 12—Descriptive Statistics and Reliability by Grade and Modality

Grade	Test	Number of Items	Max. Points	N-Count	Mean (RS)	SD (RS)	Mean P-value	SEMn	Reliability	SEM
K	Listening	24	24	26,266	20.10	3.95	0.84	0.02	0.84	1.58
	Speaking	16	34	26,266	25.13	7.59	0.75	0.02	0.94	1.86
	Reading	15	15	26,266	10.13	3.45	0.68	0.03	0.79	1.58
	Writing	15	24	26,266	15.78	5.23	0.69	0.05	0.86	1.96
	Listening and Speaking	40	58	26,266	45.23	10.64	0.80	0.02	0.93	2.82
	Reading and Writing	30	39	26,266	25.91	8.10	0.69	0.03	0.90	2.56
1	Listening	24	24	25,676	22.12	3.03	0.92	0.01	0.85	1.17
	Speaking	16	34	25,676	28.50	6.60	0.85	0.02	0.94	1.62
	Reading	15	15	25,676	13.43	2.62	0.90	0.01	0.86	0.98
	Writing	15	24	25,676	20.62	4.25	0.87	0.02	0.86	1.59
	Listening and Speaking	40	58	25,676	50.62	9.04	0.89	0.01	0.94	2.21
	Reading and Writing	30	39	25,676	34.05	6.52	0.89	0.01	0.92	1.84
2	Listening	24	24	21,159	18.26	4.04	0.77	0.03	0.81	1.76
	Speaking	16	34	21,159	28.65	6.90	0.85	0.01	0.95	1.54
	Reading	24	24	21,159	15.11	4.84	0.65	0.05	0.83	2.00
	Writing	16	22	21,159	14.15	4.66	0.68	0.04	0.84	1.86
	Listening and Speaking	40	58	21,159	46.91	10.04	0.81	0.02	0.93	2.66
	Reading and Writing	40	46	21,159	29.26	8.91	0.66	0.03	0.91	2.67
3	Listening	24	24	19,024	19.79	3.98	0.83	0.02	0.84	1.59
	Speaking	16	34	19,024	29.32	6.99	0.87	0.01	0.95	1.56
	Reading	24	24	19,024	17.53	4.90	0.74	0.03	0.86	1.83
	Writing	16	22	19,024	16.15	4.46	0.78	0.03	0.85	1.73
	Listening and Speaking	40	58	19,024	49.12	10.18	0.85	0.01	0.94	2.49
	Reading and Writing	40	46	19,024	33.68	8.83	0.76	0.02	0.92	2.50
4	Listening	24	24	17,853	20.58	3.82	0.86	0.02	0.85	1.48
	Speaking	16	34	17,853	29.72	6.80	0.88	0.01	0.95	1.52
	Reading	24	24	17,853	18.86	4.65	0.80	0.03	0.87	1.68
	Writing	16	22	17,853	17.14	4.24	0.82	0.03	0.85	1.64
	Listening and Speaking	40	58	17,853	50.30	9.96	0.87	0.01	0.95	2.23
	Reading and Writing	40	46	17,853	36.01	8.40	0.80	0.02	0.92	2.38

Note: The data file used in this table is the final research file. The total n-counts for Grades K–12 are 208,420. 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 invalid codes).

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

Grade	Test	Number of Items	Max. Points	N-Count	Mean (RS)	SD (RS)	Mean P-value	SEMn	Reliability	SEM
5	Listening	25	25	15,459	19.80	5.03	0.79	0.02	0.88	1.74
	Speaking	16	34	15,459	29.36	7.05	0.87	0.01	0.96	1.41
	Reading	27	27	15,459	18.82	6.23	0.70	0.02	0.90	1.97
	Writing	19	25	15,459	18.26	5.17	0.76	0.03	0.86	1.93
	Listening and Speaking	40	59	15,459	49.17	11.22	0.82	0.01	0.95	2.51
	Reading and Writing	46	52	15,459	37.08	10.77	0.72	0.02	0.93	2.85
6	Listening	25	25	13,164	19.45	5.33	0.78	0.02	0.89	1.77
	Speaking	16	34	13,164	28.70	7.71	0.85	0.01	0.96	1.54
	Reading	27	27	13,164	18.58	6.28	0.69	0.02	0.90	1.99
	Writing	19	25	13,164	17.96	5.40	0.75	0.03	0.87	1.95
	Listening and Speaking	40	59	13,164	48.16	12.16	0.81	0.01	0.96	2.43
	Reading and Writing	46	52	13,164	36.53	11.02	0.71	0.02	0.93	2.92
7	Listening	25	25	12,609	17.93	4.99	0.72	0.02	0.84	2.00
	Speaking	16	34	12,609	27.77	8.47	0.82	0.01	0.96	1.69
	Reading	27	27	12,609	16.64	5.93	0.62	0.03	0.86	2.22
	Writing	19	25	12,609	17.13	5.54	0.71	0.03	0.86	2.07
	Listening and Speaking	41	59	12,609	45.70	12.37	0.76	0.02	0.95	2.77
	Reading and Writing	46	52	12,609	33.77	10.66	0.66	0.02	0.92	3.02
8	Listening	25	25	12,506	18.31	5.00	0.73	0.02	0.85	1.94
	Speaking	16	34	12,506	27.61	8.50	0.82	0.01	0.96	1.70
	Reading	27	27	12,506	17.44	5.96	0.65	0.03	0.87	2.15
	Writing	19	25	12,506	17.40	5.48	0.72	0.03	0.86	2.05
	Listening and Speaking	41	59	12,506	45.92	12.42	0.77	0.02	0.95	2.78
	Reading and Writing	46	52	12,506	34.85	10.64	0.68	0.02	0.92	3.01
9	Listening	25	25	14,288	16.02	6.05	0.64	0.02	0.89	2.01
	Speaking	16	34	14,288	25.41	9.90	0.75	0.01	0.97	1.71
	Reading	27	27	14,288	15.26	5.87	0.57	0.03	0.86	2.20
	Writing	19	25	14,288	15.40	6.57	0.63	0.03	0.89	2.18
	Listening and Speaking	41	59	14,288	41.42	14.89	0.68	0.02	0.96	2.98
	Reading and Writing	46	52	14,288	30.66	11.68	0.59	0.02	0.93	3.09

Note: The data file used in this table is the final research file. The total n-counts for Grades K–12 are 208,420. 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 invalid codes).

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

Grade	Test	Number of Items	Max. Points	N-Count	Mean (RS)	SD (RS)	Mean P-value	SEMn	Reliability	SEM
10	Listening	25	25	13,289	16.30	5.75	0.65	0.02	0.87	2.07
	Speaking	16	34	13,289	26.43	8.42	0.78	0.01	0.96	1.68
	Reading	27	27	13,289	15.98	5.68	0.59	0.03	0.85	2.20
	Writing	19	25	13,289	16.12	6.03	0.65	0.03	0.87	2.17
	Listening and Speaking	41	59	13,289	42.74	13.04	0.70	0.02	0.95	2.92
	Reading and Writing	46	52	13,289	32.10	10.92	0.62	0.02	0.92	3.09
11	Listening	25	25	9,775	17.30	5.31	0.69	0.02	0.86	1.99
	Speaking	16	34	9,775	27.98	7.07	0.83	0.01	0.94	1.73
	Reading	27	27	9,775	17.45	5.43	0.65	0.03	0.84	2.17
	Writing	19	25	9,775	17.42	5.48	0.70	0.03	0.85	2.12
	Listening and Speaking	41	59	9,775	45.28	11.18	0.74	0.02	0.94	2.74
	Reading and Writing	46	52	9,775	34.87	10.14	0.67	0.02	0.91	3.04
12	Listening	25	25	7,352	17.86	5.36	0.71	0.02	0.87	1.93
	Speaking	16	34	7,352	29.05	6.49	0.86	0.01	0.94	1.59
	Reading	27	27	7,352	17.57	5.76	0.65	0.03	0.86	2.16
	Writing	19	25	7,352	17.64	5.74	0.71	0.03	0.87	2.07
	Listening and Speaking	41	59	7,352	46.91	10.58	0.77	0.02	0.93	2.80
	Reading and Writing	46	52	7,352	35.21	10.72	0.67	0.02	0.92	3.03

Note: The data file used in this table is the final research file. The total n-counts for Grades K–12 are 208,420. 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 invalid codes).

5.6 Inter-Rater Reliability

Another source of measurement error can occur in handscoring. Inter-rater reliability investigates the extent to which students would receive the same score if they were scored again by either the same rater or by different raters. Inter-rater reliability is commonly evaluated in two ways. First, the correlation of two raters scoring each student's paper can be used as an estimate of this type of reliability. Using this approach, the consistency of ratings—e.g., a higher correlation—is the measure of reliability. Another type of evidence for inter-rater reliability is to calculate the percentage of agreement between raters. In a scenario of perfect agreement between raters, the result would be a 100% match in their assignment of scores. The lower limit to agreement rates—i.e., raters never agree with each other—is 0% agreement.

10% Scoring Audit

NYSED has chosen to carry out a yearly audit of the NYSESLAT because the audit seeks to ensure that local teachers apply the same rigorous scoring standards as intended by NYSED. This audit also provides evidence of inter-rater reliability. To achieve this, NYSED requires Questar to rescore 10% of all students' Writing responses after each test administration. This 10% audit sample within each grade band is selected as a stratified random sample based on the Need/Resource Category (NRC). The NRC for the state schools is divided into six categories:

1. New York City
2. Big 4 Cities (Buffalo, Rochester, Syracuse, and Yonkers)
3. High Need Urban/Suburbs
4. High Need Rural
5. Average Need
6. Low Need

Target values were calculated for each level using the percentage of enrollment data for each of the above classifications and were applied to the level targets for "10% Plus 40% Over Sample." Because the sample selection level was by school, enrollment data collected for the Spring 2012 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 band. These schools returned their booklets to Questar for rescoring. The following sections and tables indicate the procedures that Questar conducted to ensure reliability and accurate scoring of the items.

To perform the audit, scoring personnel and team leaders were typically the same people who had participated in the rangefinding process and were familiar with the NYSESLAT. All training was done using the same scoring materials as those used by the New York State teachers for scoring the operational test. Readers had to qualify for scoring the audit by scoring sample papers prior to the actual audit scoring. Following this, team leaders rechecked readers throughout the process to maintain their accuracy.

Table 13 provides the audit sample information regarding the number of students obtained for the rater analyses after merging the auditing file and the final research file. The audit data and the research file were merged by student ID (SID). Questar selected the stratified random sample to achieve the desired percentage (about 10%) of the full population.

Table 13—Audit Sample

Grade Band	#Students in Audit Sample after Merging	#Students in Population by Grade Band	Percentage of Total Population
K–1	5,861	51,942	11.28%
2–4	6,386	58,036	11.00%
5–6	3,260	28,623	11.39%
7–8	2,672	25,115	10.64%
9–12	5,616	44,704	12.56%
Total	23,795	208,420	11.42%

Table 14 provides, by grade band and item, the rater agreement for the Pre-Writing and Writing constructed-response items between local raters and Questar raters. The Questar 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 following are the descriptions of the rater-agreement variable:

- Exact: 0 score point difference between local and audit raters
- Adjacent: +/– 1 score point difference between local and audit raters
- Non-Adjacent: +/– 2 score point difference between local and audit raters

In Grades K–1, all nine items are worth a maximum of two points. In the remaining grade bands, the first three items are worth a maximum of two points while the remaining item in each grade band is worth a maximum of four points. For the two-point items, the percent of exact agreement ranged from 68.68% to 96.02%. The exact agreement for the four-point items ranged from 45.71% to 57.86%. The percent of adjacent agreement ranged from 2.61% to 29.15% for the two-point items and 38.22% to 43.54% for the four-point items. The total percent exact and adjacent agreement ranged from 96.23% to 99.80% for the two-point items and 89.25% to 96.08% for the four point items. The exact agreement is lower for the four-point items because there is a wider range of score points.

In addition to agreement ratings, Table 14 also includes the intraclass correlations between the local raters and Questar raters, the weighted Kappa coefficient, the asymptotic standard error (ASE) of the weighted Kappa coefficient, and the lower and upper 95% confidence limits of the Kappa coefficient.

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 subjects. The intraclass correlations in Table 14 were calculated using the random and fixed effects model introduced by Shrout and Fleiss (1979) to measure the agreement of the local and Questar 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 0.57 to 0.89 with the lowest correlation coming from Grades 7–8 and the highest coming from Grades K–1. The results indicate that the raters generally agreed relatively well.

Kappa Coefficient

The Kappa coefficient is an index measuring rater agreement. The value of Kappa is always less than or equal to 1. A value of 1 signifies perfect agreement while values less than 1 indicate less-than-perfect agreement. In extremely rare situations, a weighted Kappa value can be negative, which indicates that the two observers agreed less than would be expected just by chance. It is also rare that there is perfect agreement. There is not complete agreement on the interpretation as to what degree of agreement can be considered an acceptable level, but one guideline to the interpretation of the weighted Kappa coefficient is as follows (Altman, 1991, p. 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 bands indicate that the rater agreement ranged 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% confidence limits indicate that the errors of the weighted Kappa coefficient are very small and within a limited range.

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

Grade Band	Item	Max. Points	N-count	% Exact	% Adjacent	% Non-Adjacent	Intraclass Correlation	Weighted Kappa	ASE	95% Confidence Limits	
										Lower	Upper
K–1	1	2	5,861	85.43%	13.94%	0.63%	0.68	0.69	0.01	0.67	0.70
	2	2	5,861	83.76%	15.88%	0.36%	0.67	0.66	0.01	0.64	0.68
	3	2	5,861	94.80%	3.99%	1.21%	0.84	0.83	0.01	0.81	0.85
	4	2	5,861	96.02%	2.61%	1.36%	0.84	0.81	0.01	0.79	0.84
	5	2	5,861	87.03%	12.17%	0.80%	0.89	0.83	0.01	0.82	0.84
	6	2	5,861	85.67%	13.82%	0.51%	0.88	0.83	0.01	0.82	0.84
	7	2	5,861	83.86%	15.88%	0.26%	0.86	0.79	0.01	0.78	0.80
	8	2	5,861	83.93%	15.87%	0.20%	0.86	0.80	0.01	0.78	0.81
	9	2	5,861	76.90%	22.81%	0.29%	0.79	0.71	0.01	0.69	0.72
2–4	1	2	6,386	93.78%	5.26%	0.96%	0.83	0.79	0.01	0.77	0.81
	2	2	6,386	72.38%	25.76%	1.86%	0.65	0.56	0.01	0.54	0.58
	3	2	6,386	74.22%	25.16%	0.61%	0.60	0.55	0.01	0.53	0.57
	4	4	6,386	57.86%	38.22%	3.91%	0.67	0.53	0.01	0.52	0.55
5–6	1	2	3,260	92.58%	7.06%	0.37%	0.76	0.69	0.02	0.66	0.73
	2	2	3,260	77.27%	20.80%	1.93%	0.70	0.62	0.01	0.60	0.65
	3	2	3,260	72.21%	25.89%	1.90%	0.68	0.60	0.01	0.57	0.62
	4	4	3,260	52.06%	41.87%	6.07%	0.65	0.49	0.01	0.47	0.51
7–8	1	2	2,672	84.88%	14.37%	0.75%	0.78	0.72	0.01	0.70	0.75
	2	2	2,672	68.68%	29.15%	2.17%	0.57	0.49	0.02	0.46	0.52
	3	2	2,672	79.12%	18.64%	2.25%	0.73	0.66	0.01	0.63	0.68
	4	4	2,672	51.91%	41.32%	6.77%	0.68	0.52	0.01	0.50	0.54
9–12	1	2	5,616	85.51%	12.70%	1.80%	0.78	0.73	0.01	0.72	0.75
	2	2	5,616	81.71%	16.01%	2.28%	0.68	0.51	0.01	0.48	0.54
	3	2	5,616	71.60%	24.63%	3.77%	0.63	0.54	0.01	0.52	0.57
	4	4	5,616	45.71%	43.54%	10.75%	0.66	0.49	0.01	0.47	0.51

Note: The statistics in the table are based on the merged file between audit and Questar's final research file. The audit data and research file were merged by SIDs.

Table 15 provides the proportional distribution of the score point differences between local and Questar ratings by grade band and item. The percentage of zero difference ranged from 68.68% to 96.02% for the two-point items and 45.71% to 57.86% for the four-point items. By definition, zero difference is exact agreement. Hence, these proportions are the same as the exact agreement percentages shown in Table 14.

Table 15—Percentages of Score Difference Between Raters

Grade Band	Item #	Max.	Percentage of Point Difference Score Difference (local score minus Questar score)								
			–4	–3	–2	–1	0	1	2	3	4
K–1	1	2			0.02	2.59	85.43	11.35	0.61		
	2	2			0.03	4.57	83.76	11.31	0.32		
	3	2			0.55	1.38	94.80	2.61	0.67		
	4	2			0.80	0.96	96.02	1.66	0.56		
	5	2			0.32	1.89	87.03	10.27	0.48		
	6	2			0.09	3.60	85.67	10.22	0.43		
	7	2			0.14	7.97	83.86	7.92	0.12		
	8	2			0.10	9.64	83.93	6.23	0.10		
	9	2			0.17	10.63	76.90	12.18	0.12		
2–4	1	2			0.17	1.61	93.78	3.65	0.78		
	2	2			0.34	7.16	72.38	18.60	1.52		
	3	2			0.53	9.25	74.22	15.91	0.08		
	4	4		0.09	0.50	12.34	57.86	25.88	3.19	0.13	
5–6	1	2			0.31	5.52	92.58	1.53	0.06		
	2	2			0.46	6.84	77.27	13.96	1.47		
	3	2			0.89	11.29	72.21	14.60	1.01		
	4	4		0.06	0.74	12.82	52.06	29.05	5.15	0.06	0.06
7–8	1	2			0.19	4.12	84.88	10.25	0.56		
	2	2			1.09	11.71	68.68	17.44	1.09		
	3	2			1.31	8.31	79.12	10.33	0.94		
	4	4		0.04	1.31	12.39	51.91	28.93	4.64	0.52	0.26
9–12	1	2			0.28	2.69	85.51	10.01	1.51		
	2	2			1.28	6.73	81.71	9.28	1.00		
	3	2			1.82	9.26	71.60	15.37	1.96		
	4	4	0.02	0.09	0.75	10.49	45.71	33.05	8.49	0.96	0.45

Table 16 provides the mean and standard deviation (SD) of each item for the local raters, the audit (Questar) raters, and the difference between the local and Questar raters. The largest mean difference (0.38) occurred for item 4 in Grades 9–12. The SD ratio ranged from 0.74 to 1.31, with the largest SD ratio occurring for item 4 in Grades 5–6. For all items, the mean difference is close to 0 and the SD ratio is close to 1, which is a good indication of rater agreement.

Table 16—Comparison Between Local and Audit Raters

Grade Band	Item #	Local		Questar		Differences	
		Mean	SD	Mean	SD	Mean	SD Ratio
K–1	1	1.69	0.48	1.59	0.53	0.10	0.90
	2	1.65	0.49	1.57	0.53	0.08	0.94
	3	1.79	0.50	1.77	0.53	0.03	0.94
	4	1.85	0.49	1.83	0.51	0.01	0.96
	5	1.45	0.79	1.35	0.89	0.10	0.89
	6	1.14	0.81	1.06	0.84	0.08	0.97
	7	1.37	0.77	1.36	0.77	0.01	1.00
	8	1.32	0.77	1.34	0.79	-0.01	0.99
	9	1.03	0.78	0.99	0.73	0.04	1.06
2–4	1	1.83	0.47	1.79	0.55	0.03	0.87
	2	1.54	0.65	1.33	0.74	0.21	0.88
	3	1.42	0.58	1.28	0.57	0.14	1.03
	4	1.91	0.99	1.72	0.82	0.19	1.21
5–6	1	1.84	0.45	1.88	0.39	-0.04	1.15
	2	1.54	0.65	1.39	0.73	0.15	0.90
	3	1.33	0.72	1.24	0.74	0.09	0.97
	4	2.09	1.08	1.84	0.83	0.25	1.31
7–8	1	1.65	0.59	1.58	0.65	0.07	0.91
	2	1.45	0.66	1.29	0.69	0.16	0.96
	3	1.61	0.59	1.47	0.67	0.14	0.87
	4	2.12	1.11	1.89	0.89	0.23	1.24
9–12	1	1.64	0.62	1.54	0.70	0.10	0.89
	2	1.77	0.53	1.59	0.71	0.18	0.74
	3	1.57	0.62	1.36	0.72	0.21	0.87
	4	2.37	1.19	1.98	0.97	0.38	1.23

Summary of the 10% Audit Inter-Rater Reliability

Inter-rater reliability statistics and evidence collected and presented here demonstrate that the local raters and Questar raters in general had a high degree of agreement on the NYSESLAT scores. A few items had greater disagreement—e.g., item 4 in Grade Band 5–6. In these cases, the local raters generally gave higher scores than those given by Questar raters.

5.7 Accuracy and Consistency of Performance Level Classifications

The NYSESLAT scale scores are used to categorize student performance into one of four proficiency levels (Beginning, Intermediate, Advanced, and Proficient). (See Chapter 8 for details on standard setting.) Although it is important to know that student scores are highly reliable in any examination, it is perhaps of even greater importance to evaluate the reliability of the decisions based on these scores. The previous vendor assessed the reliability of classification decisions via an analysis of the probabilities of correct and consistent classification of student performance. Livingston and Lewis (1995) developed procedures to do this, and these procedures were applied to derive measures of the accuracy and consistency of the classifications for the NYSESLAT.

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, which would be defined as equal in difficulty and covering the same content as the form they actually took. These notions are graphically displayed in Figures 1 and 2.

Figure 1—Classification Accuracy

		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 2—Classification Consistency

		Decision made on the second form taken	
		Does Not Achieve Proficient Status	Achieves Proficient Status
Decision made on the first form taken	Does Not Achieve Proficient Status	Correct Classification	Misclassification
	Achieves Proficient Status	Misclassification	Correct Classification

Note: Adapted from Young and Yoon (1998)

In Figure 1, accurate classifications occur when the decision made based on the average of all-theoretical forms agrees with the decision made on the form actually taken. Misclassifications are defined as the case where a student who achieves a score of "Does Not Achieve Proficient Status" for his or her all-theoretical forms average is classified incorrectly as "Achieves Proficient Status." In Figure 2, consistent classification occurs 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.

Decision consistency and accuracy techniques are outlined and implemented by Hanson (1991), Haertel (1996), Livingston and Lewis (1995), and Young and Yoon (1998). The previous vendor used the BB-Class software developed by Hanson (1995) to conduct these 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.

Table 17 contains the proportions of False Positive and False Negative classifications. The sum of the Accuracy, False Positive, and False Negative values should be equal to 1.00. However,

because of rounding, the table values may not always equal 1.00. False Positive and False Negative classifications refer to the mismatch between student true scores and observed scores. The proportions of student scores misclassified into the category of Achieves Proficient Status are labeled as False Positive. The proportion of student scores misclassified into the category Does Not Achieve Proficient Status when student scores meet proficient status is contained in the False Negative column.

Table 17 also presents the decision accuracy and consistency results for the Achieves Proficient Status cut scores by the paired modalities of Listening/Speaking and Reading/Writing. The table has the following information:

- Accuracy
- False positives
- False negatives
- Consistency

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, the chance that Kindergarten students were classified the same was 77% for Listening/Speaking. For the Listening/Speaking proficiency level classifications, the consistency ranged from 66% to 91% across all grades. For the Reading/Writing proficiency level classifications, the consistency ranged from 73% to 93% across all grades. In most cases, decision accuracy and consistency for the proficiency level were higher for Reading/Writing than those for Listening/Speaking.

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—i.e., the true score—of the examination. For example, for students in Kindergarten, 83% of the classifications for Listening/Speaking at the proficiency level were accurate, while 7% percent of the students were classified as proficient when they were not. Ten percent of students were classified as not proficient when they were proficient. For the Listening/Speaking proficiency level classification, the accuracy of the classification ranged from 67% to 93% across all grades. For the Reading/Writing proficiency level classification, the accuracy ranged from 78% to 95% across all grades.

Table 17 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 33% for Listening/Speaking and 4% to 14% for Reading/Writing across all grades. The false negative ranged from 0% to 14% for Listening/Speaking and 0% to 12% for Reading/Writing across all the grades. More attention should be paid to the high false positive values because it is less acceptable to see a high percentage of students who were not proficient be classified as proficient and exit the program.

The NYSESLAT shows much higher percentages of consistency and accuracy 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	0.83	0.07	0.10	0.77
	Reading and Writing	0.95	0.04	0.02	0.93
1	Listening and Speaking	0.67	0.33	0.00	0.66
	Reading and Writing	0.78	0.10	0.12	0.73
2	Listening and Speaking	0.89	0.03	0.08	0.83
	Reading and Writing	0.93	0.05	0.02	0.91
3	Listening and Speaking	0.85	0.03	0.11	0.78
	Reading and Writing	0.85	0.08	0.07	0.81
4	Listening and Speaking	0.87	0.03	0.10	0.79
	Reading and Writing	0.86	0.14	0.00	0.81
5	Listening and Speaking	0.81	0.04	0.14	0.73
	Reading and Writing	0.89	0.05	0.06	0.84
6	Listening and Speaking	0.82	0.05	0.13	0.75
	Reading and Writing	0.88	0.05	0.06	0.84
7	Listening and Speaking	0.88	0.04	0.08	0.83
	Reading and Writing	0.92	0.06	0.03	0.89
8	Listening and Speaking	0.90	0.04	0.06	0.86
	Reading and Writing	0.90	0.06	0.04	0.87
9	Listening and Speaking	0.93	0.03	0.03	0.91
	Reading and Writing	0.93	0.04	0.03	0.90
10	Listening and Speaking	0.93	0.04	0.04	0.89
	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	0.91	0.05	0.04	0.87
12	Listening and Speaking	0.87	0.05	0.08	0.81
	Reading and Writing	0.86	0.09	0.05	0.82

CHAPTER 6: VALIDITY

Assessments constructed by Questar support the standards set forth in the Standards for Educational and Psychological Testing (AERA, APA, & NCME, 1999). The judgments made in this technical report regarding test validity are based upon the following[2]:

- 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

Test content as evidence of validity is evidenced by the relationship between test material and items in that they should represent an appropriate sampling of the knowledge, skills, and understanding of the construct or domain being assessed. The definitions of those skills, knowledge, and understanding come from the New York ESL learning standards. As such, the standards define what is expected of ELL students in the state of New York. To ensure the content validity of the assessment, multiple steps are taken to align the NYSESLAT items and test forms to these standards.

With input from NYSED assessment and content specialists, psychometricians and ELL assessment specialists developed a test blueprint (see Chapter 2 for further details) that includes a range of item types that can measure the appropriate standards in terms of which standards and performance indicators lend themselves to assessment on a standardized test. The item types can also measure performance indicators from the ESL learning standards. This determination was made by ELL assessment specialists who worked with committees of New York State ESL teachers in order to identify which performance indicators are suited to classroom use only and, consequently, would not become part of the assessment.

Item writers for the NYSESLAT received item-writing assignments and were instructed to use NYSESLAT item type specifications and the ESL learning standards to write their items. Specifically, they were asked to use these materials to guide their writing and identify which standards and performance indicators were the best match to what the items were intended to measure. Committees of New York State teachers then reviewed the items as part of the review process, which included discussion and agreement upon the standards and performance indicators assigned to each item.

Performance standards and indicators were part of several criteria—including item statistics, content diversity, and gender and ethnic balance of item subjects—that were taken into consideration by ELL assessment specialists and NYSED assessment and content specialists

[2] The page numbers in parentheses refer to the page numbers in the Standards for Educational and Psychological Testing, 1999.

during forms construction to ensure that forms reflected broad coverage of the ESL learning standards. Following forms construction, assessment specialists developed item maps for each form to show this coverage. ELL assessment specialists and content specialists at NYSED reviewed the item maps to confirm the alignment of a given form to the learning standards. This item mapping is detailed in Table 5 and in Appendices A.1 and A.2 as evidence for the alignment to the ESL learning standards.

6.2 Internal Structure

A coherent assessment is not a random collection of assessment tasks or test items. The tasks selected for inclusion in an assessment are intended to contribute positively to the total result. One way of providing evidence for this is to assess the interrelationship of the tasks for the test, referred to as the internal structure of the assessment. The following questions are often posed to investigate the internal structure of education assessments (Nitko, 2004):

- 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?

Correlations based on raw scores of the four modalities (Speaking, Listening, Reading, and Writing) are calculated to investigate the answers to these questions. Tables 18 and 19 show the intercorrelation of the four modalities by grade band and grade. The evidence of internal structure of the 2012 NYSESLAT can also be illustrated by examining the point-biserial correlation coefficients and fit statistics provided in Appendices B and D.

Table 18—Intercorrelation Among the Modalities by Grade Band

Grade Band	Modality	Correlation Coefficient			
		Listening	Speaking	Reading	Writing
K–1	Listening	1.00			
	Speaking	0.71	1.00		
	Reading	0.58	0.51	1.00	
	Writing	0.62	0.57	0.81	1.00
2–4	Listening	1.00			
	Speaking	0.69	1.00		
	Reading	0.69	0.55	1.00	
	Writing	0.73	0.66	0.79	1.00
5–6	Listening	1.00			
	Speaking	0.72	1.00		
	Reading	0.71	0.58	1.00	
	Writing	0.76	0.72	0.78	1.00
7–8	Listening	1.00			
	Speaking	0.67	1.00		
	Reading	0.69	0.55	1.00	
	Writing	0.74	0.73	0.73	1.00
9–12	Listening	1.00			
	Speaking	0.68	1.00		
	Reading	0.72	0.58	1.00	
	Writing	0.75	0.72	0.75	1.00

Table 19—Intercorrelation Among the Modalities by Grade

Grade	Modality	Correlation Coefficient			
		Listening	Speaking	Reading	Writing
K	Listening	1.00			
	Speaking	0.66	1.00		
	Reading	0.49	0.40	1.00	
	Writing	0.55	0.49	0.73	1.00
1	Listening	1.00			
	Speaking	0.72	1.00		
	Reading	0.61	0.56	1.00	
	Writing	0.63	0.61	0.79	1.00
2	Listening	1.00			
	Speaking	0.66	1.00		
	Reading	0.62	0.51	1.00	
	Writing	0.68	0.62	0.76	1.00

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

		Correlation Coefficient			
Grade	**Modality**	**Listening**	**Speaking**	**Reading**	**Writing**
3	Listening	1.00			
	Speaking	0.70	1.00		
	Reading	0.68	0.57	1.00	
	Writing	0.73	0.69	0.78	1.00
4	Listening	1.00			
	Speaking	0.74	1.00		
	Reading	0.72	0.61	1.00	
	Writing	0.76	0.72	0.79	1.00
5	Listening	1.00			
	Speaking	0.71	1.00		
	Reading	0.71	0.57	1.00	
	Writing	0.76	0.71	0.78	1.00
6	Listening	1.00			
	Speaking	0.73	1.00		
	Reading	0.70	0.58	1.00	
	Writing	0.76	0.73	0.78	1.00
7	Listening	1.00			
	Speaking	0.67	1.00		
	Reading	0.68	0.54	1.00	
	Writing	0.74	0.73	0.73	1.00
8	Listening	1.00			
	Speaking	0.67	1.00		
	Reading	0.70	0.55	1.00	
	Writing	0.74	0.73	0.73	1.00
9	Listening	1.00			
	Speaking	0.73	1.00		
	Reading	0.74	0.62	1.00	
	Writing	0.79	0.77	0.76	1.00
10	Listening	1.00			
	Speaking	0.68	1.00		
	Reading	0.71	0.56	1.00	
	Writing	0.74	0.71	0.74	1.00
11	Listening	1.00			
	Speaking	0.62	1.00		
	Reading	0.69	0.53	1.00	
	Writing	0.70	0.65	0.73	1.00
12	Listening	1.00			
	Speaking	0.59	1.00		
	Reading	0.69	0.50	1.00	
	Writing	0.71	0.61	0.74	1.00

Observations of the language proficiency assessment subtests in Table 18 are as follows:

- Listening and Speaking are moderately correlated for grade bands ranging from 0.67–0.72.
- Listening and Reading are moderately correlated for grade bands ranging from 0.58–0.72.
- Listening and Writing are moderately correlated for grade bands ranging from 0.62–0.76.
- Speaking and Reading are moderately correlated for grade bands ranging from 0.51–0.58.
- Speaking and Writing are moderately correlated for grade bands ranging from 0.57–0.73.
- Reading and Writing are moderately correlated for grade bands ranging from 0.73–0.81.
- The language domain pairs of Listening and Speaking and Reading and Writing are moderately to highly correlated, while Speaking and Reading are not as highly correlated.

The results of these internal correlations follow what is theoretically expected, that the Reading and Writing subtests have the highest correlation range (from 0.73 to 0.81) across the five grade bands. The Speaking and Reading subtests have the lowest correlation range (from 0.51 to 0.58).

The following bullet lists contain some observations regarding the NYSESLAT subtests and the results found in Table 19:

<u>Kindergarten</u>

- Students in this age group cannot usually read or write but can have Listening and Speaking skills. The expected outcome is that neither Reading nor Writing will correlate highly with Listening or Speaking.
- Speaking and Listening have a correlation of 0.66, and Reading and Writing have a correlation of 0.73. Speaking and Reading have a low correlation of 0.4; Listening and Reading have a low correlation of 0.49; Speaking and Writing have a low correlation of 0.49; and Listening and Writing have a low correlation of 0.55.

<u>Grades 1–8</u>

- Correlation between Listening and Writing varies between 0.67 and 0.73.
- A possible explanation for this trend is that students in this age group are required to expand the use of and development of their writing skills. At the same time, there are limited increases in the demands placed upon their listening skills.

<u>Grades 9–12</u>

- The correlation between Listening and Writing steadily decreases from 0.73 to 0.59. The same pattern was observed in 2011.
- One plausible explanation is that high school content contains an increased focus on the use of writing skills, especially an increased focus on academic content. Listening skills may also increase, but not to the same degree as writing.
- 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 0.49 to 0.73) for Grades 1–8, and there is a steady decrease in the correlation from (0.77 to 0.61) for Grades 9–12.

Again, the internal correlations, and, thus, the internal structure of the test, follow expectations for the most part. Across all grades, Reading and Writing have the highest correlation range from 0.73 to 0.79, and Speaking and Reading have the lowest correlation range from 0.5 to 0.62.

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

A confirmatory factor analysis (CFA) was conducted in 2012. The 2012 NYSESLAT scoring dimensionality tests produced results that substantiated the use of the two-factor model in reporting IRT scores. The two-factor model provided a better fit to the data as the one-factor unidimensional model as measured by global indices of fit, especially for Grades K–1, 2–4, and 5–6. This pattern is justifiable by content experts as a distinct demarcation of modalities for younger students for whom a marked differentiation existed 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. Tables 20–24 show the global fit indices for the one- and two-factor models by grade band. Appendix F provides a detailed summary of the 2012 CFA.

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

Fit Indices	One-Factor model	Two-Factor Model
χ^2	13955.411	209.103
GFI	0.884	0.998
AGFI	0.421	0.980
RMSR	0.079	0.005
RMSEA	0.367	0.063

Note: χ^2 with 2 *df*, p <.001 for the one-factor model and χ^2 with 1 *df*, p<.001 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
χ^2	5907.201	1494.695
GFI	0.953	0.987
AGFI	0.762	0.875
RMSR	0.035	0.014
RMSEA	0.226	0.160

Note: χ^2 with 2 *df*, p <.001 for the one-factor model and χ^2 with 1 *df*, p<.001 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
χ^2	2796.791	1184.526
GFI	0.956	0.980
AGFI	0.781	0.803
RMSR	0.030	0.018
RMSEA	0.221	0.203

Note: χ^2 with 2 *df*, p <.001 for the one-factor model and χ^2 with 1 *df*, p<.001 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
χ^2	1649.991	1503.533
GFI	0.971	0.972
AGFI	0.856	0.718
RMSR	0.028	0.025
RMSEA	0.181	0.245

Note: χ^2 with 2 *df*, p <.001 for the one-factor model and χ^2 with 1 *df*, p<.001 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
χ^2	2266.212	1853.231
GFI	0.977	0.980
AGFI	0.887	0.801
RMSR	0.022	0.019
RMSEA	0.159	0.204

Note: χ^2 with 2 *df*, p <.001 for the one-factor model and χ^2 with 1 *df*, p<.001 for the two-factor model

6.3 External Structure

An additional source of validity evidence is sometimes called external criterion validity. To provide evidence of this, the external structure of the NYSESLAT with an external criterion was assessed. In particular, the data from two other tests were used. For ELL students in Grades 3–8, data from those who took both the NYSESLAT in May 2012 and the New York State English Language Arts (ELA) Test in January 2012 were examined. In the state of New York, all ELL students, with the exception of first-year ESL students, must take the ELA test. For ELL students in Grades 10–12, data from those who took both the NYSESLAT in May 2012 and the New York State Regents Comprehensive Examination in English (Regents Comp Exam in English) in June 2012 were analyzed. The Regents English exam is an end-of-course exam that all high school students, including ELL students, are required to pass as a high school graduation requirement. Students may take the exam in any grade, but most students take the exam in Grade 11.

Because the NYSESLAT is intended to be a measure of annual student progress in the achievement of academic English language proficiency, it should be expected that ELL students who score proficient on the NYSESLAT would be able to perform well in mainstream classrooms. Therefore, they should, in theory, have a similar chance to pass the ELA Test and Regents English exam as native English speakers who are required to take the state examinations. Hence, there should be a positive relationship between the NYSESLAT and the ELA Test and Regents Comp Exam in English, wherein those who perform well on the NYSESLAT are generally expected to perform well on the other two tests.

Relationship with the New York State English Language Arts (ELA) Test

Table 25 gives the sample size, minimum and maximum observed scores, means, standard deviations, and the correlation between the ELA Test and the NYSESLAT raw score total of Listening, Reading, and Writing for each grade from Grades 3–8. Following this, Table 26 contains the sample size, minimum and maximum observed scores, means, standard deviations, and the correlation between the ELA Test and the NYSESLAT combined scale scores of Listening/Speaking and Reading/Writing for each grade from Grades 3–8.

The correlations between the ELA test and the NYSESLAT raw score total of the three modalities (Listening, Reading, and Writing) ranged from 0.43 to 0.79 across the six grades. The correlations are positive between the two tests, which is logical since the modalities tested in the NYSESLAT are the same as those in the ELA test. The correlations between the ELA Test and the NYSESLAT combined scale scores of the four modalities (Listening, Speaking, Reading, and Writing) ranged from 0.7 to 0.74 in Grades 3–8. The correlations in Table 26 are positive but slightly lower than the correlations in Table 25 because the NYSESLAT total combined scale scores include the Speaking modality, while the ELA test does not include a Speaking portion.

In sum, a positive relationship exists between the NYSESLAT and the ELA Test, which provides evidence of the external structure check and demonstrates a positive relationship (moderate) between the two exams. It can be reasonably concluded that higher scores on the NYSESLAT are associated with higher scores on the ELA Test for each grade examined in this study.

Table 25—Descriptive Statistics of the NYSESLAT Raw Scores of All Modalities and Their Correlation with the ELA Test Total Scale Scores

Grade	N-Count	Modality	Max. Points	Mean	SD	Sample Min.	Sample Max.	Correlation with ELA
3	16,957	Listening	24	20.56	2.95	2	24	0.59
	16,957	Speaking	34	30.93	3.95	0	34	0.43
	16,957	Reading	24	18.26	4.32	1	24	0.75
	16,957	Writing	22	16.95	3.58	0	22	0.70
	16,957	LI/RD/WR	70	55.77	9.45	8	70	0.79
4	15,841	Listening	24	21.36	2.68	1	24	0.58
	15,841	Speaking	34	31.28	3.86	0	34	0.44
	15,841	Reading	24	19.61	3.96	2	24	0.73
	15,841	Writing	22	17.96	3.26	0	22	0.67
	15,841	LI/RD/WR	70	58.93	8.68	8	70	0.77
5	13,384	Listening	25	20.89	3.82	3	25	0.59
	13,384	Speaking	34	31.10	4.11	0	34	0.44
	13,384	Reading	27	19.85	5.55	2	27	0.74
	13,384	Writing	25	19.32	4.08	0	25	0.68
	13,384	LI/RD/WR	77	60.06	11.86	12	77	0.77
6	11,130	Listening	25	20.68	4.11	0	25	0.57
	11,130	Speaking	34	30.69	4.81	0	34	0.46
	11,130	Reading	27	19.67	5.62	1	27	0.71
	11,130	Writing	25	19.14	4.30	0	25	0.68
	11,130	LI/RD/WR	77	59.49	12.37	10	77	0.75
7	10,534	Listening	25	18.99	4.18	0	25	0.59
	10,534	Speaking	34	29.90	5.87	0	34	0.48
	10,534	Reading	27	17.58	5.55	1	27	0.69
	10,534	Writing	25	18.35	4.61	0	25	0.68
	10,534	LI/RD/WR	77	54.93	12.56	3	77	0.75
8	10,407	Listening	25	19.30	4.27	1	25	0.61
	10,407	Speaking	34	29.54	6.36	0	34	0.50
	10,407	Reading	27	18.32	5.53	2	27	0.69
	10,407	Writing	25	18.51	4.63	0	25	0.66
	10,407	LI/RD/WR	77	56.13	12.72	4	77	0.75
2–4*	32,798	Listening	24	20.94	2.85	1	24	0.56
	32,798	Speaking	34	31.10	3.91	0	34	0.43
	32,798	Reading	24	18.91	4.21	1	24	0.71
	32,798	Writing	22	17.43	3.46	0	22	0.66
	32,798	LI/RD/WR	70	57.29	9.22	8	70	0.74
5–6	24,514	Listening	25	20.80	3.95	0	25	0.56
	24,514	Speaking	34	30.91	4.45	0	34	0.44
	24,514	Reading	27	19.77	5.58	1	27	0.70
	24,514	Writing	25	19.24	4.18	0	25	0.65
	24,514	LI/RD/WR	77	59.80	12.10	10	77	0.73
7–8	20,941	Listening	25	19.15	4.23	0	25	0.56
	20,941	Speaking	34	29.72	6.12	0	34	0.47
	20,941	Reading	27	17.95	5.55	1	27	0.63
	20,941	Writing	25	18.43	4.62	0	25	0.63
	20,941	LI/RD/WR	77	55.53	12.65	3	77	0.69

*Grade band 2–4 does not include Grade 2.

Table 26—Descriptive Statistics of the ELA Test and NYSESLAT Combined Scale Scores of Listening/Speaking and Reading/Writing

Grade	N-Count	Test	Mean	SD	Sample Min.	Sample Max.	Correlation
3	16,957	NYSESLAT Scale Score	1,338	77.63	992	1,605	0.74
		NYS ELA Scale Score	646	20.25	475	722	
4	15,841	NYSESLAT Scale Score	1,367	82.26	1,015	1,605	0.70
		NYS ELA Scale Score	646	30.20	430	754	
5	13,384	NYSESLAT Scale Score	1,378	76.69	1,077	1,642	0.73
		NYS ELA Scale Score	651	16.56	495	712	
6	11,130	NYSESLAT Scale Score	1,375	81.82	1,048	1,642	0.71
		NYS ELA Scale Score	642	14.00	480	691	
7	10,534	NYSESLAT Scale Score	1,376	74.32	1,077	1,666	0.70
		NYS ELA Scale Score	642	17.97	470	707	
8	10,407	NYSESLAT Scale Score	1,382	79.35	1,080	1,666	0.71
		NYS ELA Scale Score	629	19.28	430	697	

Proficiency Level Comparisons

As an additional way to demonstrate the positive relationship of student success on both exams, a crosstabs of the proficiency levels of the ELA Test and NYSESLAT was performed. Tables 27–32 contain the percentage distributions of the ELA proficiency level by the NYSESLAT proficiency level for each grade from Grades 3–8.

In Table 27, nearly 100% of the Grade 3 students who were classified as beginning students on the NYSESLAT were also classified as not proficient (Level 1 and Level 2) on the ELA Test. Of the 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. Of the students who were classified as advanced students on the NYSESLAT, 84% of them were classified as not proficient (Level 1 and Level 2) on the ELA Test while 15% of them were classified as proficient on the ELA Test. Of the students who were classified as proficient students on the NYSESLAT, 50% of them were classified as not proficient (Level 1 and Level 2) on the ELA Test, and 50% of them were classified as proficient on the ELA Test.

Table 27—Grade 3: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
ELA Test	**Beginning (N=1,214)**	**Intermediate (N=4,117)**	**Advanced (N=7,806)**	**Proficient (N=3,820)**
Level 1	98.02%	79.89%	24.03%	3.40%
Level 2	1.65%	18.90%	60.61%	46.81%
Level 3	0.33%	1.21%	15.27%	48.82%
Level 4	0.00%	0.00%	0.09%	0.97%
Total	100.00%	100.00%	100.00%	100.00%

In Table 28, nearly 100% of the Grade 4 students who were classified as beginning students on the NYSESLAT were also classified as not proficient (Level 1 and Level 2) on the ELA Test. Of the 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. Of the students who were classified as advanced students on the NYSESLAT, 83% of them were classified as not proficient (Level 1 and Level 2) on the ELA Test, and 17% of them were classified as proficient on the ELA Test. Of the students who were classified as proficient students on the NYSESLAT, 52% of them were classified as not proficient (Level 1 and Level 2) on the ELA Test, and 48% of them were classified as proficient on the ELA Test.

Table 28—Grade 4: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
ELA Test	**Beginning (N=936)**	**Intermediate (N=3,225)**	**Advanced (N=8,682)**	**Proficient (N=2,998)**
Level 1	97.44%	71.81%	15.71%	2.94%
Level 2	2.35%	26.76%	67.76%	49.33%
Level 3	0.21%	1.43%	16.46%	47.33%
Level	0.00%	0.00%	0.07%	0.40%
Total	100.00%	100.00%	100.00%	100.00%

In Table 29, nearly 100% of the Grade 5 students who were classified as beginning students on the NYSESLAT were also classified as not proficient (Level 1 and Level 2) on the ELA Test. Of the 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. Of the students who were classified as advanced students on the NYSESLAT, 92% of them were classified as not proficient (Level 1 and Level 2) on the ELA Test, and 8% of them were classified as proficient on the ELA Test. Of the students who were classified as proficient students on the NYSESLAT,69% of them were classified as not proficient (Level 1 and Level 2) on the ELA Test, and 31% of them were classified as proficient on the ELA Test.

Table 29—Grade 5: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
ELA Test	**Beginning (N=833)**	**Intermediate (N=2,300)**	**Advanced (N=6,078)**	**Proficient (N=4,173)**
Level 1	97.48%	86.74%	34.96%	5.68%
Level 2	2.16%	12.65%	57.40%	63.05%
Level 3	0.36%	0.61%	7.62%	31.10%
Level 4	0.00%	0.00%	0.02%	0.17%
Total	100.00%	100.00%	100.00%	100.00%

In Table 30, 100% of the Grade 6 students who were classified as beginning students on the NYSESLAT were also classified as not proficient (Level 1 and Level 2) on the ELA Test. Of the 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. Of the students who were classified as advanced students on the NYSESLAT, 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. Of the 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 16% of them were classified as proficient on the ELA Test.

Table 30—Grade 6: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
ELA Test	**Beginning (N=893)**	**Intermediate (N=2,552)**	**Advanced (N=5,184)**	**Proficient (N=2,501)**
Level 1	97.31%	87.97%	42.28%	12.16%
Level 2	2.58%	11.83%	54.13%	71.65%
Level 3	0.11%	0.20%	3.59%	16.19%
Level 4	0.00%	0.00%	0.00%	0.00%
Total	100.00%	100.00%	100.00%	100.00%

In Table 31, 100% of the Grade 7 students who were classified as beginning students on the NYSESLAT were also classified as not proficient (Level 1 and Level 2) on the ELA Test. Of the 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. Of the students who were classified as advanced students on the NYSESLAT, 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. Of the students who were classified as proficient students on the NYSESLAT, 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.

Table 31—Grade 7: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
ELA Test	**Beginning (N=1,447)**	**Intermediate (N=4,143)**	**Advanced (N=3,717)**	**Proficient (N=1,227)**
Level 1	95.92%	59.88%	13.45%	2.93%
Level 2	4.08%	39.92%	81.68%	73.76%
Level 3	0.00%	0.19%	4.87%	23.23%
Level 4	0.00%	0.00%	0.00%	0.08%
Total	100.00%	100.00%	100.00%	100.00%

In Table 32, 100% of the Grade 8 students who were classified as beginning students on the NYSESLAT were also classified as not proficient (Level 1 and Level 2) on the ELA Test. Of the 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. Of the students who were classified as advanced students on the NYSESLAT, 97% of them were classified as not proficient (Level 1 and Level 2) on the ELA Test, and 3% of them were classified as proficient on the ELA Test. Of the 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.

Table 32—Grade 8: Percentage of English Language Learners Scoring at Each ELA Level at NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
ELA Test	**Beginning (N=1,451)**	**Intermediate (N=4,051)**	**Advanced (N=3,342)**	**Proficient (N=1,563)**
Level 1	94.83%	60.06%	16.82%	4.03%
Level 2	5.10%	39.62%	80.58%	82.41%
Level 3	0.07%	0.32%	2.60%	13.56%
Level 4	0.00%	0.00%	0.00%	0.08%
Total	100.00%	100.00%	100.00%	100.00%

Summary

These analyses and resulting correlations demonstrate that a moderately positive relationship exists between the NYSESLAT and the ELA Test. Positive correlations like these are evidence of external validity of an assessment. The positive correlations also show that higher scores on the NYSESLAT are associated with higher scores on the ELA Test. Since the ELA Test and the NYSESLAT were developed for different populations of students, these relationships will also be less than complete correlation. Hence, it is of no surprise that there is only a moderate, rather than a high, positive correlation between the two tests. Similarly, the analyses also confirm the validity of inferences made from the NYSESLAT cut scores because about 50% to 86% (across the six grades) of those who scored at the proficiency level on the NYSESLAT were also classified as Level 2 or Level 3 (proficient) by the ELA Test. Moreover, the high percentages of students 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.

Relationship with the New York State Regents Comprehensive Examination in English

Table 33 contains the sample size, maximum score points, minimum and maximum observed scores, means, standard deviations, and the correlation between the New York State Regents Comprehensive Examination in English (Regents Comp Exam in English) and each of the NYSESLAT modalities' raw scores, as well as the total raw scores of Listening, Reading, and Writing for Grades 10–12. Table 34 presents the sample size, minimum and maximum observed scores, means, standard deviations, the correlation between the 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 10–12.

Table 33—Descriptive Statistics of the NYSESLAT Modalities' Raw Scores and Their Correlations with the Regents Comp Exam in English Total Scale Scores

Grade/ Grade Band	N-Count	Test	Max. Points	Mean	SD	Sample Min.	Sample Max.	Correlation with Regents English
10	3,653	Listening	25	17.91	5.12	1	25	0.48
10	3,653	Speaking	34	28.84	6.62	0	34	0.41
10	3,653	Reading	27	17.49	5.18	2	27	0.57
10	3,653	Writing	25	17.98	5.14	1	25	0.55
10	3,653	L/R/W Total	77	53.38	13.65	9	77	0.60
11	10,785	Listening	25	17.87	4.83	0	25	0.45
11	10,785	Speaking	34	28.80	6.13	0	34	0.38
11	10,785	Reading	27	18.00	4.92	1	27	0.56
11	10,785	Writing	25	18.09	4.75	0	25	0.57
11	10,785	L/R/W Total	77	53.96	12.66	7	77	0.60
12	7,447	Listening	25	17.71	4.70	1	25	0.35
12	7,447	Speaking	34	28.78	5.76	0	34	0.28
12	7,447	Reading	27	17.55	4.86	3	27	0.45
12	7,447	Writing	25	17.71	4.57	0	25	0.44
12	7,447	L/R/W Total	77	52.97	11.92	13	77	0.49
10–12	21,885	Listening	25	17.82	4.84	0	25	0.42
10–12	21,885	Speaking	34	28.80	6.09	0	34	0.35
10–12	21,885	Reading	27	17.76	4.95	1	27	0.52
10–12	21,885	Writing	25	17.94	4.76	0	25	0.52
10–12	21,885	L/R/W Total	77	53.53	12.60	7	77	0.56

Table 34—Descriptive Statistics of the NYSESLAT Modalities' Scale Scores and Total Scale Scores and Their Correlations with the Regents English Exam Total Scale Scores

Grade	N-Count	Test	Mean	SD	Sample Min.	Sample Max.	Correlation with Regents English
10	3,653	Regents English	48.16	19.40	0	100	--
		NYSESLAT Listening and Speaking	46.76	10.73	1	59	0.48
		NYSESLAT Reading and Writing	35.47	9.47	3	52	0.61
		Total NYSESLAT	1,423.00	82.52	1,111	1,719	0.56
11	10,785	Regents English	53.64	18.76	1	99	--
		NYSESLAT Listening and Speaking	46.67	9.74	0	59	0.46
		NYSESLAT Reading and Writing	36.09	8.83	4	52	0.61
		Total NYSESLAT	1,421.00	74.84	1,061	1,719	0.55
12	7,447	Regents English	55.69	17.55	0	100	--
		NYSESLAT Listening and Speaking	46.49	9.10	5	59	0.36
		NYSESLAT Reading and Writing	35.26	8.37	6	52	0.50
		Total NYSESLAT	1,416.00	69.73	1,173	1,670	0.44
10–12	21,885	Regents English	53.42	18.64	0	100	--
		NYSESLAT Listening and Speaking	46.62	9.70	0	59	0.43
		NYSESLAT Reading and Writing	35.71	8.80	3	52	0.57
		Total NYSESLAT	1,420.00	74.56	1,061	1,719	0.51

The correlation coefficients for the Regents Comp Exam in English scale scores and the NYSESLAT raw score total of the three modalities (Listening, Reading, and Writing) ranged from 0.35 to 0.6 across the three grades (10, 11, and 12). Within a grade, the combined modalities have higher correlations between the scale scores of the Regents Comp Exam in English than any individual modality of the NYSESLAT raw scores. Additionally, it is logical that the correlations between the scale scores of the Regents Comp Exam in English and the combined raw scores of Reading/Writing of the NYSESLAT are higher than the correlations between the scale scores of the Regents Comp Exam in English and the NYSESLAT combined raw scores of Listening/Speaking because the Regents Comp Exam in English does not have a Speaking component.

The correlations between the scale scores of the Regents Comp Exam in English and the NYSESLAT total scale scores for all four modalities combined ranged from 0.44 to 0.56 across the three grades. All correlations in Table 34 are positive but slightly lower than the correlations in Table 33 because of the NYSESLAT total combined scale scores including Speaking, which is not a portion of the Regents Comp Exam in English. For this same reason, the correlations

between the scale scores of the Regents Comp Exam in English and the NYSESLAT combined scale scores of Reading/Writing are higher than the correlations between the scale scores of the Regents Comp Exam in English and the NYSESLAT combined scale scores of Listening/Speaking.

Proficiency Level Classification

The classification percentages on the NYSESLAT by proficiency level were also compared to passing and failing on the Regents Comp Exam in English for Grades 10–12. The results are presented in Tables 35–37.

Table 35 shows that 22% of the Grade 10 students who scored proficient on the NYSESLAT passed at 75 on the Regents Comp Exam in English, and 29% of them passed at 65 on the Regents Comp Exam in English. For students who scored advanced on the NYSESLAT, the percentage of those who also passed at 75 on the Regents Comp Exam in English decreased to about 10%, and to about 23% for those passed at 65 on the Regents Comp Exam in English.

Table 35—Grade 10: Percentage of ELL Students Passing/Failing the Regents Comp Exam in English at Each NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
Regents English Exam	**Beginning (N=289)**	**Intermediate (N=1,476)**	**Advanced (N=1,181)**	**Proficient (N=707)**
Pass at 75	1.38%	2.10%	9.82%	22.49%
Pass at 65	0.69%	10.03%	23.37%	29.00%
Fail	97.92%	87.87%	66.81%	48.51%
Total	100.00%	100.00%	100.00%	100.00%

Table 36 shows that of the Grade 11 students who were classified as proficient on the NYSESLAT, 36% passed at 75, and 33% passed the Regents Comp Exam in English exam at 65. Of the students who were classified as advanced students on the NYSESLAT, 31% passed. For those students classified at the beginning and intermediate levels on the NYSESLAT, the percentage of students passing the Regents Comp Exam in English were 2% and 16%, respectively.

Table 36—Grade 11: Percentage of ELL Students Passing/Failing the Regents Comp Exam in English at Each NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
Regents English Exam	**Beginning (N=776)**	**Intermediate (N=5,447)**	**Advanced (N=3,077)**	**Proficient (N=1,485)**
Pass at 75	0.52%	6.08%	19.14%	35.56%
Pass at 65	1.93%	15.97%	30.74%	33.40%
Fail	97.55%	77.95%	50.11%	31.04%
Total	100.00%	100.00%	100.00%	100.00%

Table 37 shows that of the Grade 12 students who were classified as proficient on the NYSESLAT, 35% passed at 75, and 33% passed the Regents Comp Exam in English at 65. Of the students who were classified as advanced students on the NYSESLAT, 33% passed. For those students classified at the beginning and intermediate levels on the NYSESLAT, the percentage of students passing the Regents Comp Exam in English were 8% and 22%, respectively.

Table 37—Grade 12: Percentage of ELL Students Passing/Failing the Regents Comp Exam in English at Each NYSESLAT Proficiency Level

	NYSESLAT Proficiency Levels			
Regents English Exam	**Beginning (N=540)**	**Intermediate (N=4,647)**	**Advanced (N=1,473)**	**Proficient (N=787)**
Pass at 75	1.48%	9.30%	22.20%	35.32%
Pass at 65	7.96%	22.36%	33.20%	32.66%
Fail	90.56%	68.35%	44.60%	32.02%
Total	100.00%	100.00%	100.00%	100.00%

Summary

The correlation analyses detailed above provide additional evidence of a moderately positive relationship between the NYSESLAT and the Regents English exam. Furthermore, the evidence suggests, as other analyses do, that higher NYSESLAT scores are associated with higher scores on the Regents English exam.

In addition to the evidence of a positive correlation between the two assessments, these analyses also confirm the validity of inferences made based upon the NYSESLAT cut scores. In other words, ELL students classified as proficient on the NYSESLAT have a higher degree of passing scores on the Regents English exam as opposed to ELL students who scored below the proficient level on the NYSESLAT.

CHAPTER 7: CALIBRATION, EQUATING, AND SCALING

This chapter describes the Item Response Theory (IRT) models and the processes used to calibrate, equate, and scale the NYSESLAT. 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 was used for polytomous items. These IRT 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 previous vendor used WINSTEPS Version 3.27 to calibrate and scale the NYSESLAT (Linacre & Wright, 2000). Questar performed its anchor stability checks using Version 3.68.1, and it has continued to use this version since 2011 (Linacre 2009).

7.1 Item Response Model and Rationale for Use

The Rasch model for dichotomous items and the PCM for polytomous items were used for developing, scoring, and reporting the NYSESLAT and were recommended for several reasons.

1. The NYSESLAT vertical scale was created based on the pre-existing Stanford English Language Proficiency (SELP) vertical scale developed using the Rasch model. By using SELP items with known Rasch item difficulties, the vertical scale for the NYSESLAT could be created in a timely fashion.

2. 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 students per form for equating versus approximately 1,500 students per form under the 3PL IRT model (Kolen and Brennan, 2004).

3. For the requirements of the NYSESLAT program, the Rasch model has 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 led the previous vendor 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 INFIT statistics are used 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 subgroups. It is a standardized outlier-sensitive mean-square fit statistic and is 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 that 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 student 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 fit 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 the 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 flagged for INFIT and OUTFIT varied depending on the grade band and subtest. In general, around 0–2% of the items were flagged for INFIT, but the percentage of misfit was greater for OUTFIT across grade bands.

The OUTFIT and the INFIT statistics are presented in the item statistics tables in Appendix D. These numbers are updated based upon the 2012 anchor stability results.

7.3 Description of the Calibration Sample and Process

An equating sample (nearly 100% of the population) used for the equating analyses represented the ELL population in all six Need/Resource Categories (NRCs): NYC, Large City, and High Need Urban/Suburban, Rural, Average Need, and Low Need districts.

The separate scales, one for Listening/Speaking and one for Reading/Writing, were obtained by taking the item parameters for those separate combinations and using them to create raw score-to-scale score tables. Anchor items were used to scaling the item parameter estimates to the existing NYSESLAT scale.

Specifically, the WINSTEPS software program was used to conduct the item calibration, fixing the item parameters for the anchor items to the current NYSESLAT scale. The results of this calibration were used as the operational item parameters to create the preliminary scales for the NYSESLAT Spring 2012 administration. Item estimates of the common items were used as anchors to place the current NYSESLAT scale.

7.4 Equating of the NYSESLAT

Historically, the common-item, non-equivalent groups design was used for the 2009 and 2008 equating because it is one of the most practical 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). This makes the common-item non-equivalent groups design a popular choice for large-scale assessments.

The common-item, non-equivalent groups design necessitates the inclusion of items common to both administrations that would allow the linking of the tests and/or forms to be placed on the same scale. 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 common practice among psychometricians is to use a method called the anchor or fixed parameter method in which 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, or 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 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). These thetas for the different test forms are on the same scale as established by the scale transformation method discussed above.

Thus, by fixing the known parameters of the common set of items (anchor items) between 2012 and 2011, the items on the 2012 operational form were calibrated and then located on the 2011 (current) NYSESLAT scale. Once the scale locations of the 2012 NYSESLAT were known, IRT

true score equating was used to relate the raw scores on the 2012 NYSESLAT to the existing NYSESLAT scale. In this process, the true score on the NYSESLAT in 2012 with a given level of student ability was considered an equivalent true score on the 2011 test associated with that level of student 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 common on two administrations—i.e., between the assessment that is already placed on the established scale and the one that is to be placed on that scale. Alternatively, items on the current examination that have already been placed on an established scale can be used. These items then act as anchor items to place the new form on the existing scale.

For the NYSESLAT, the existing NYSESLAT scale was used to create the NYSESLAT vertical scale of 2012. For the 2012 administration, the common items between 2012 and 2011 were the anchor items on the 2012 NYSESLAT and were, as explained earlier, fixed to the parameter values from the pre-existing vertical scale of NYSESLAT. 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

The general guidelines in the use of anchor items are as follows:

- An adequate number of anchor items need to be used.
- 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 being equated.
- Anchor items are expected to have strong statistical properties—e.g., point-biserials > 0.30—and span the range of difficulty including easy, medium, and difficult items. The mean difficulty (p-value/Rasch difficulty values) of a new form's anchor items should be very similar to the mean difficulty of the previous operational test.

Stability of Anchor Items

To assess the psychometric stability of the 2012 NYSESLAT, Questar psychometricians conducted an anchor stability study aimed at determining if a significant shift had occurred in the stability of the item parameters. The anchor stability check was conducted for every grade band on the Listening/Speaking and Reading/Writing portions of the NYSESLAT. For the 2012 NYSESLAT across all grade bands, the displacement value of 0.30 for an item (used in many large-scale assessments) was used in order to identify potential outliers. For all grade bands, the overall percentage of anchor items was around 25% of the total items. Following anchor evaluation, the general rule is that the anchor set should be no smaller than 20% of the test (Huynh, 2000). Table 38 shows the number of items and the final number of anchors for each modality combination by grade band for 2011 and 2012. Items excluded from being anchor items in the NYSESLAT have their parameters re-estimated, and the new estimate becomes the adjusted item parameter in the item bank for future use.

Table 38—Common Items as Anchor Items by Modality Combination and Grade Band

Grade Band	Modality Combination	Total Number of Items in 2012	Common Items (2011 and 2012)	Percentage of Total Items	Total Number of Items in 2011	Anchor Items (2010 and 2011)	Percentage of Total Items
K–1	Listening/Speaking	40	10	25.0%	40	10	25.0%
	Reading/Writing	30	8	26.7%	30	8	26.7%
2–4	Listening/Speaking	40	9	22.5%	40	10	25.0%
	Reading/Writing	40	10	25.0%	40	10	25.0%
5–6	Listening/Speaking	41	10	24.4%	41	10	24.4%
	Reading/Writing	46	11	23.9%	46	11	23.9%
7–8	Listening/Speaking	41	11	26.8%	41	11	26.8%
	Reading/Writing	46	11	23.9%	46	11	23.9%
9–12	Listening/Speaking	41	11	26.8%	41	11	26.8%
	Reading/Writing	46	11	23.9%	46	11	23.9%

Note: Common items are the same items appearing in both the 2011 and 2012 operational tests.

Test Characteristic Curve (TCC) Comparison between 2011 and 2012

In order to ascertain the maintenance of the scale across years, comparisons between the 2011 and 2012 administrations' Test Characteristic Curves (TCCs) were examined with respect to their alignment at each level of administration. For TCCs, the *x*-axis represents the thetas, or students' abilities, and the *y*-axis represents 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.

Figures 3–4 show the TCC plots of Listening/Speaking and Reading/Writing, respectively, for Grades 9–12 for the 2012 administration. Figures 5–9 show the TCC plots of Listening/Speaking for Grade Bands K–1, 2–4, 5–6, 7–8, and 9–12, respectively, for the 2011 and 2012 administrations, and Figures 10–14 show the TCC plots of Reading/Writing for Grade Bands K–1, 2–4, 5–6, 7–8, and 9–12, respectively, for the 2011 and 2012 administrations.

The 2012 TCC plots for all grade bands (see Figures 3 and 4) indicate that the TCC shifts to the right with each progression of grade band across the Listening/Speaking and Reading/Writing vertical scales. However, TCCs for Grade Bands 2–4 and 5–6 for Listening/Speaking and Grade Bands 5–6 and 7–8 in Reading/Writing are very close or even overlap, which indicates that there is a lack of separation between these grade bands. Overall, the TCCs of 2011 and 2012 align well.

Figure 3—2012 TCC for Listening/Speaking, Grades K–12

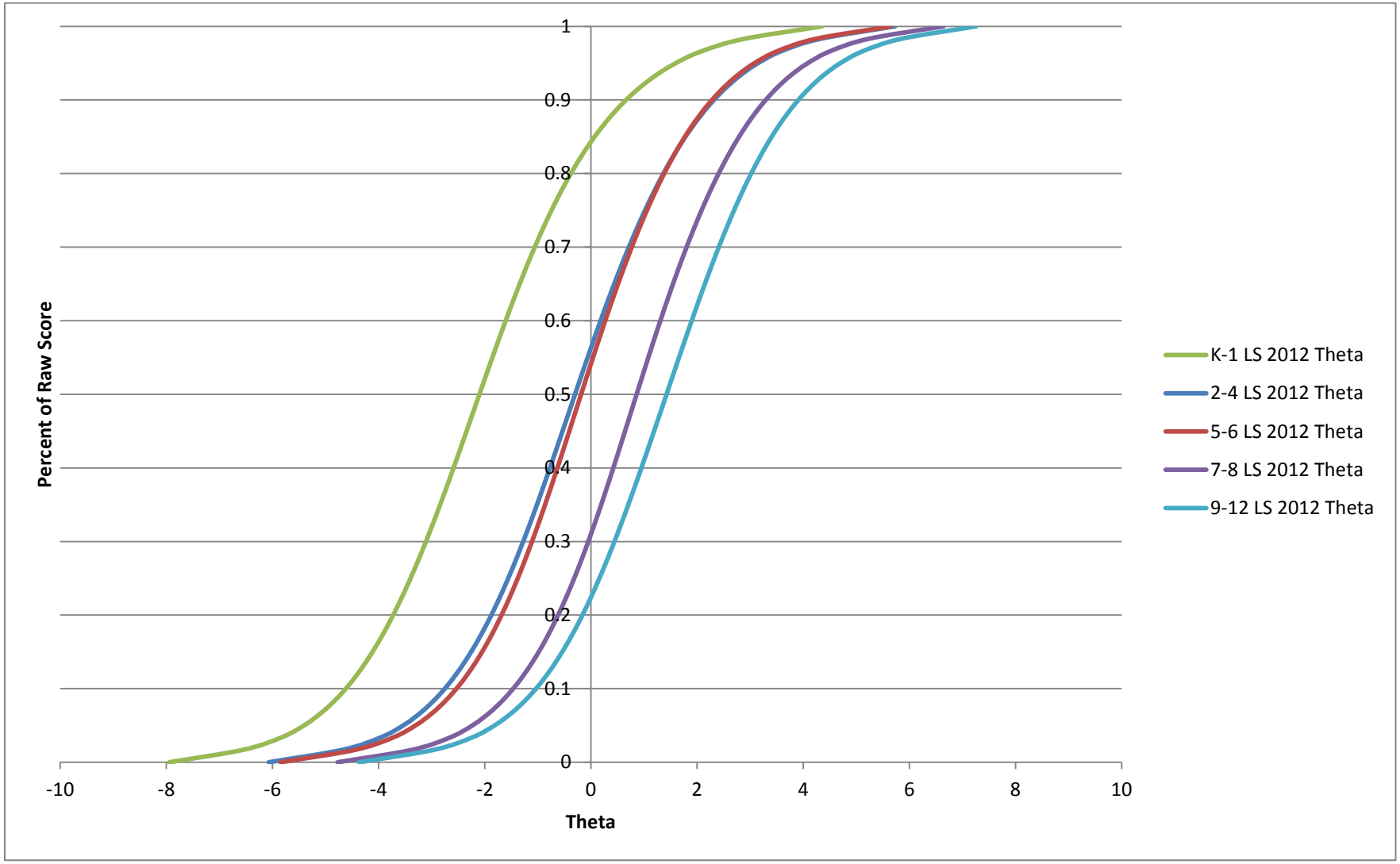


Figure 4—2012 TCC for Reading/Writing, Grades K–12

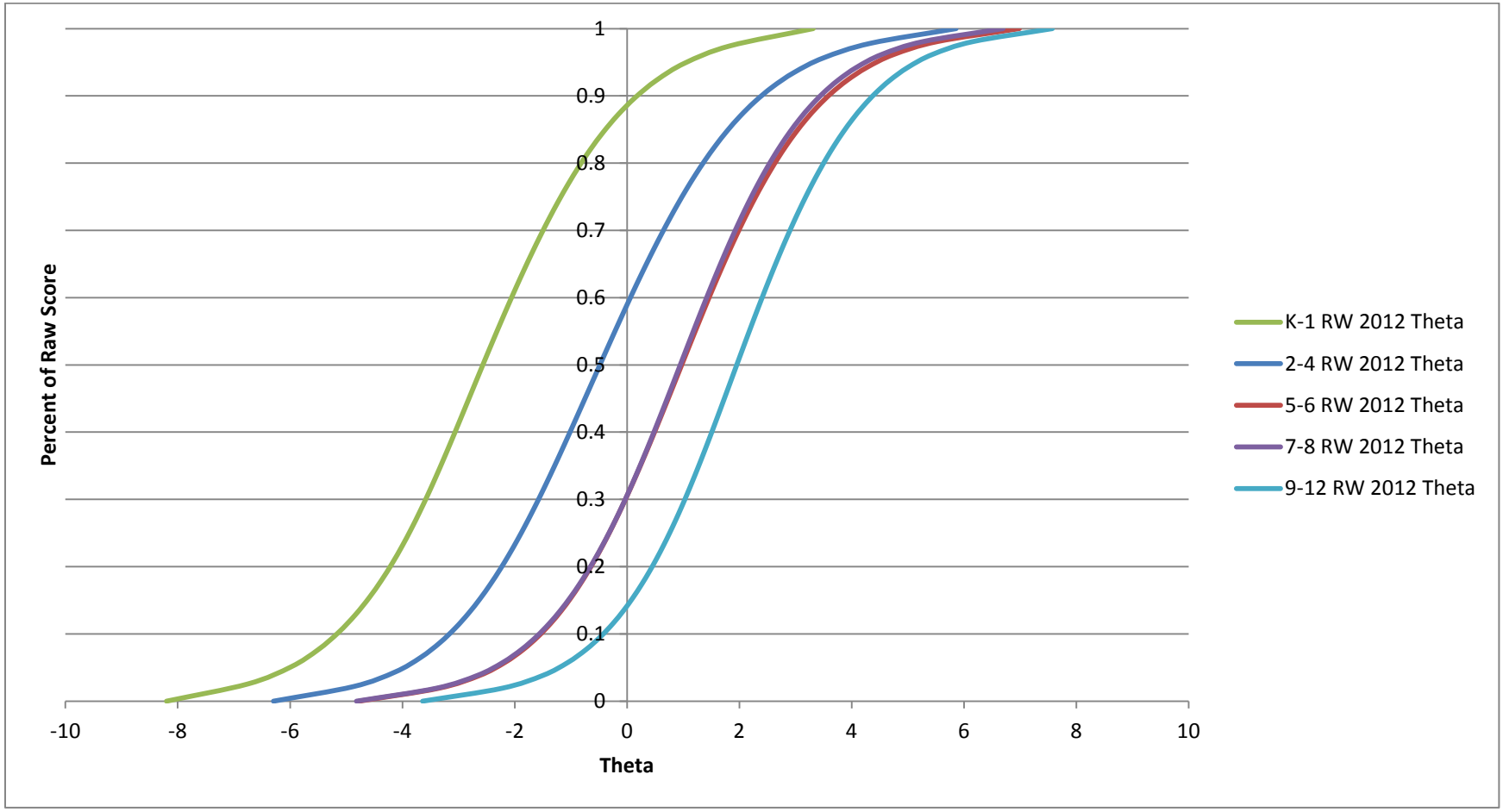


Figure 5—TCC for Listening/Speaking, Grade Band K–1

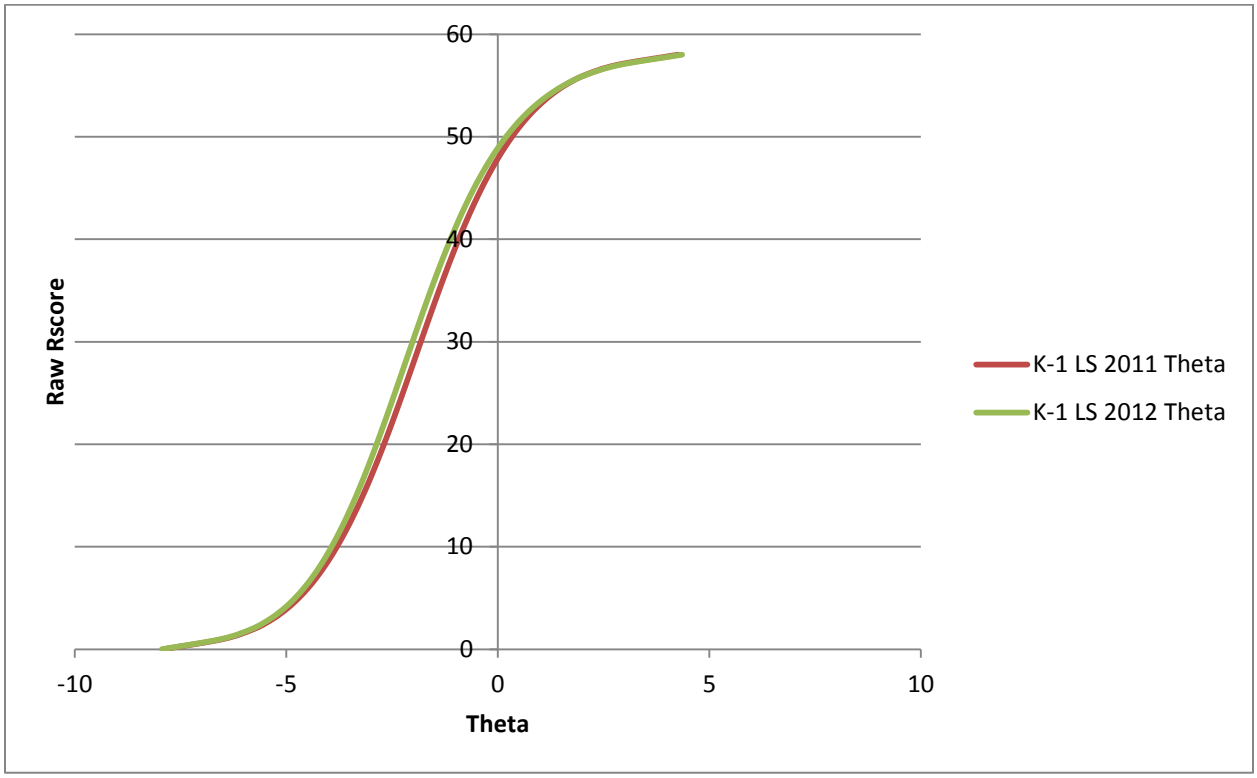


Figure 6—TCC for Listening/Speaking, Grade Band 2–4

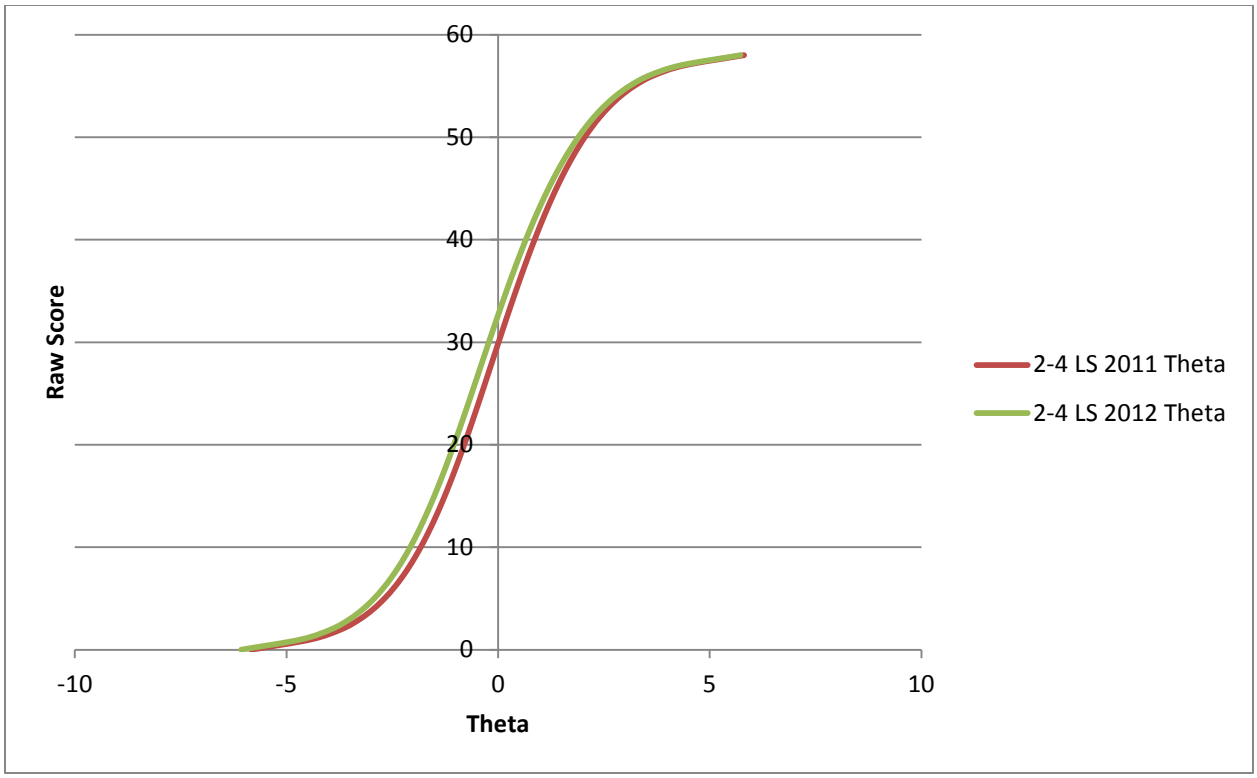


Figure 7—TCC for Listening/Speaking, Grade Band 5–6

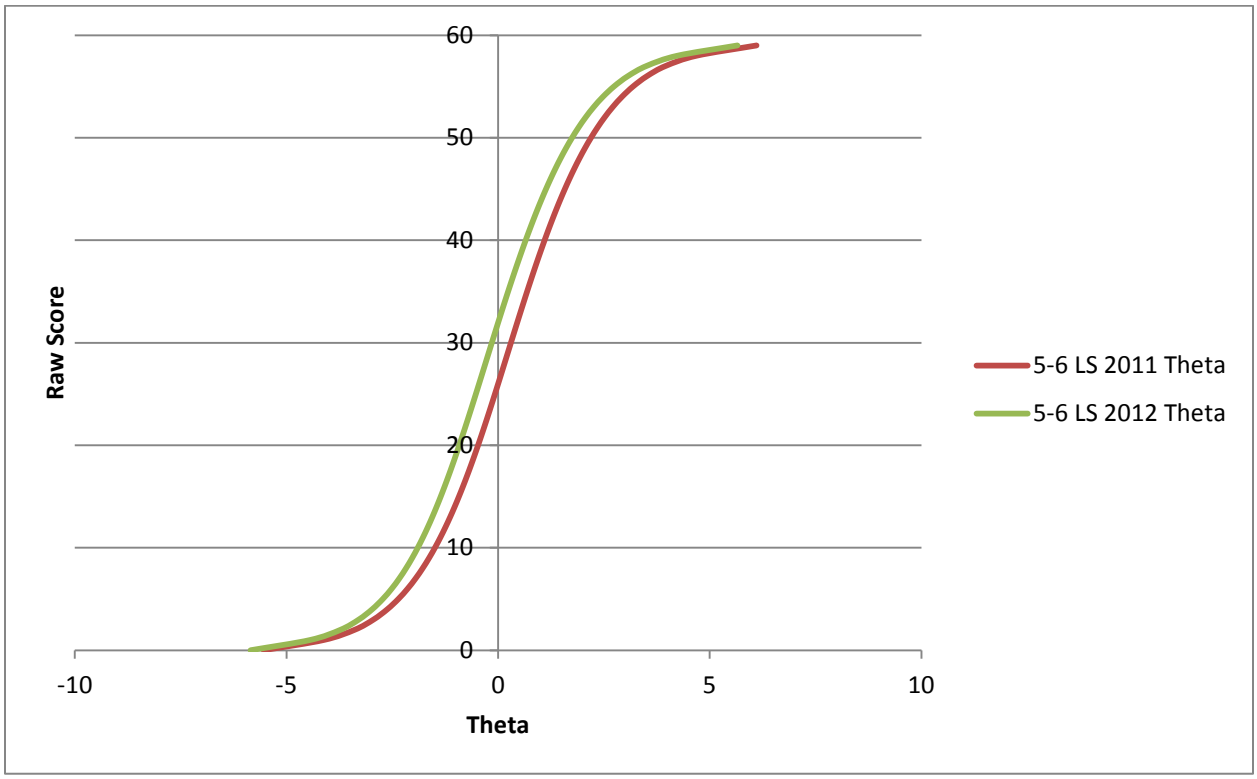


Figure 8—TCC for Listening/Speaking, Grade Band 7–8

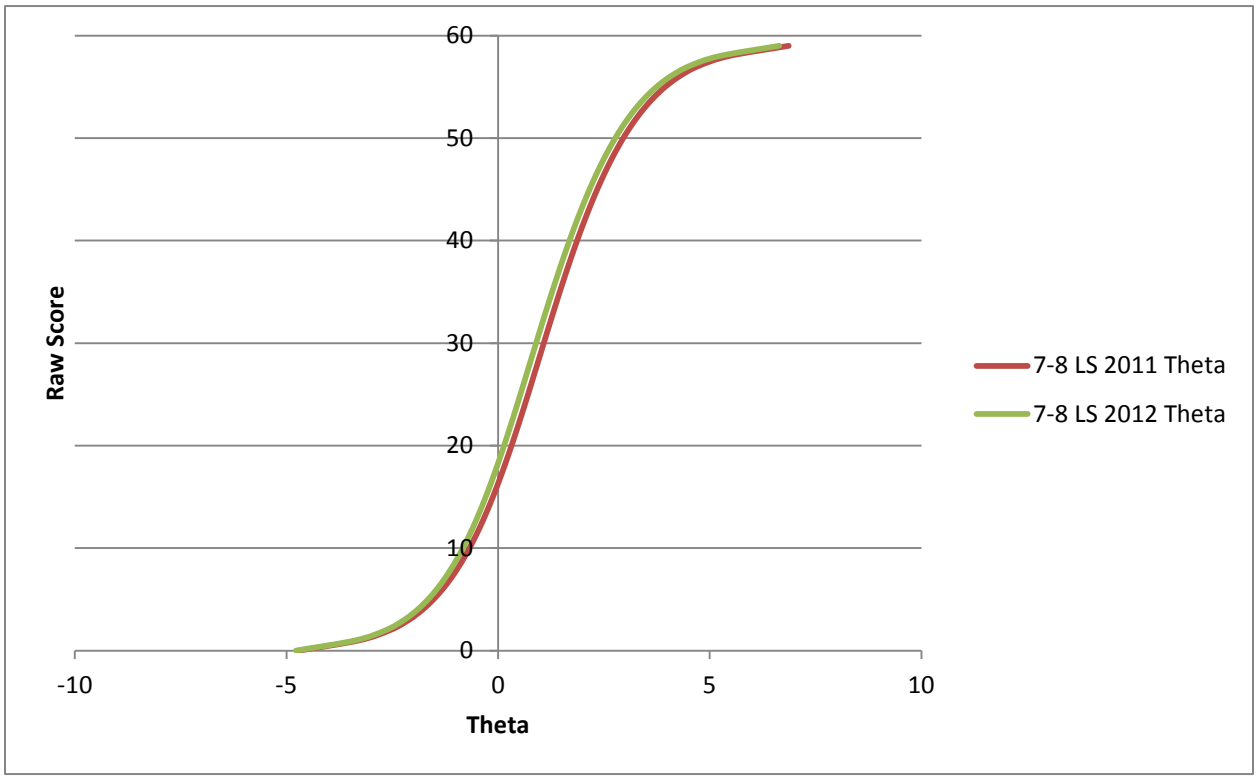


Figure 9—TCC for Listening/Speaking, Grade Band 9–12

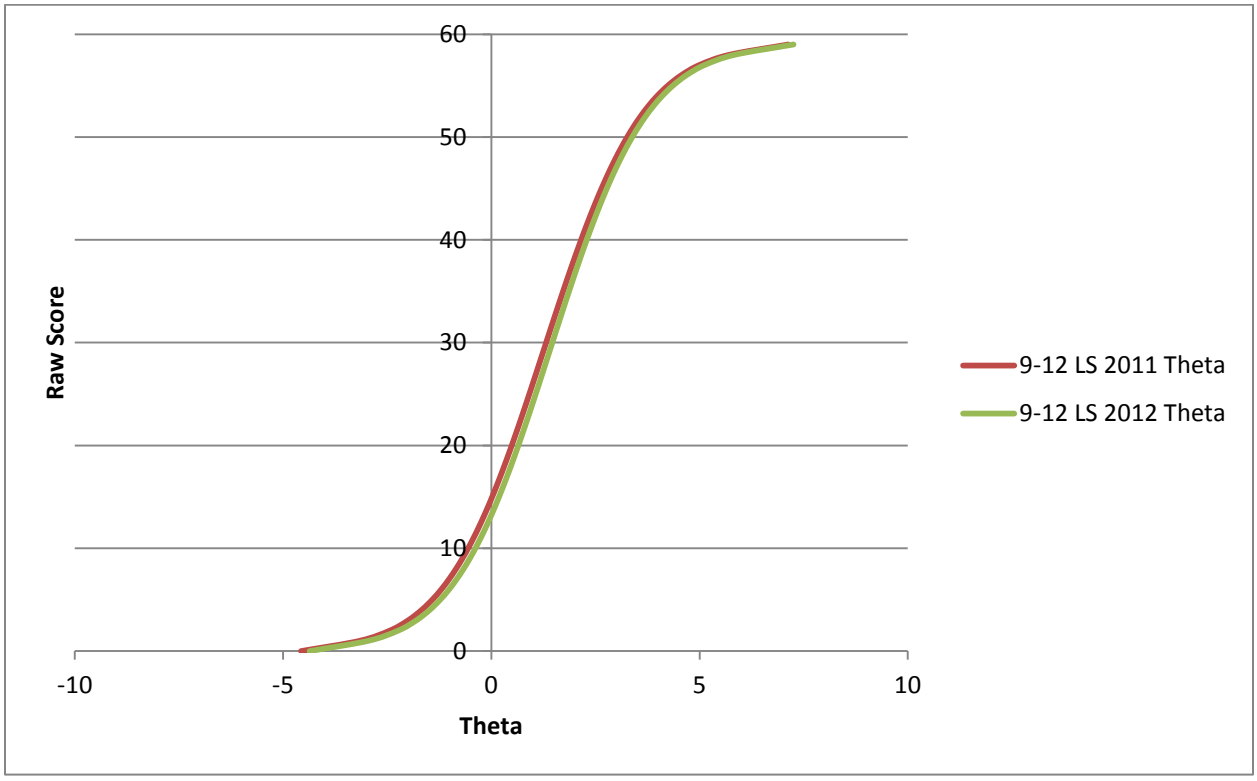


Figure 10—TCC for Reading/Writing, Grade Band K–1

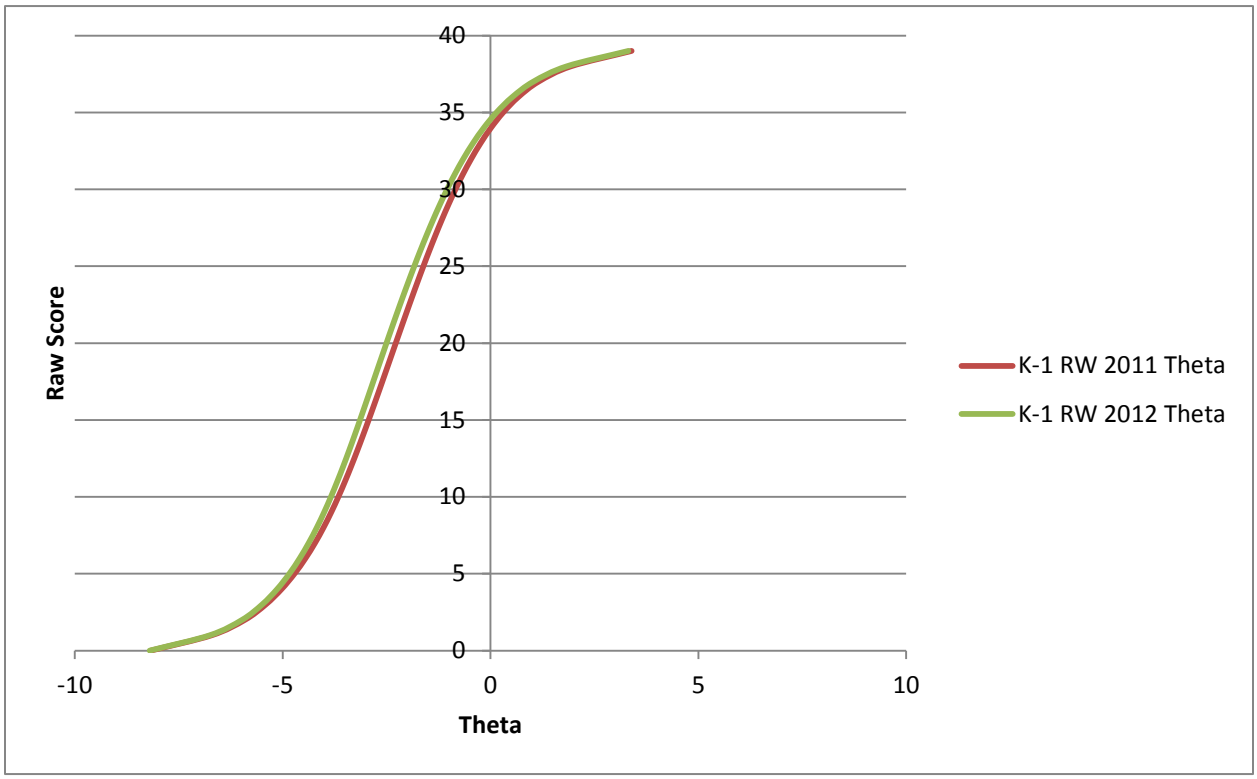


Figure 11—TCC for Reading/Writing, Grade Band 2–4

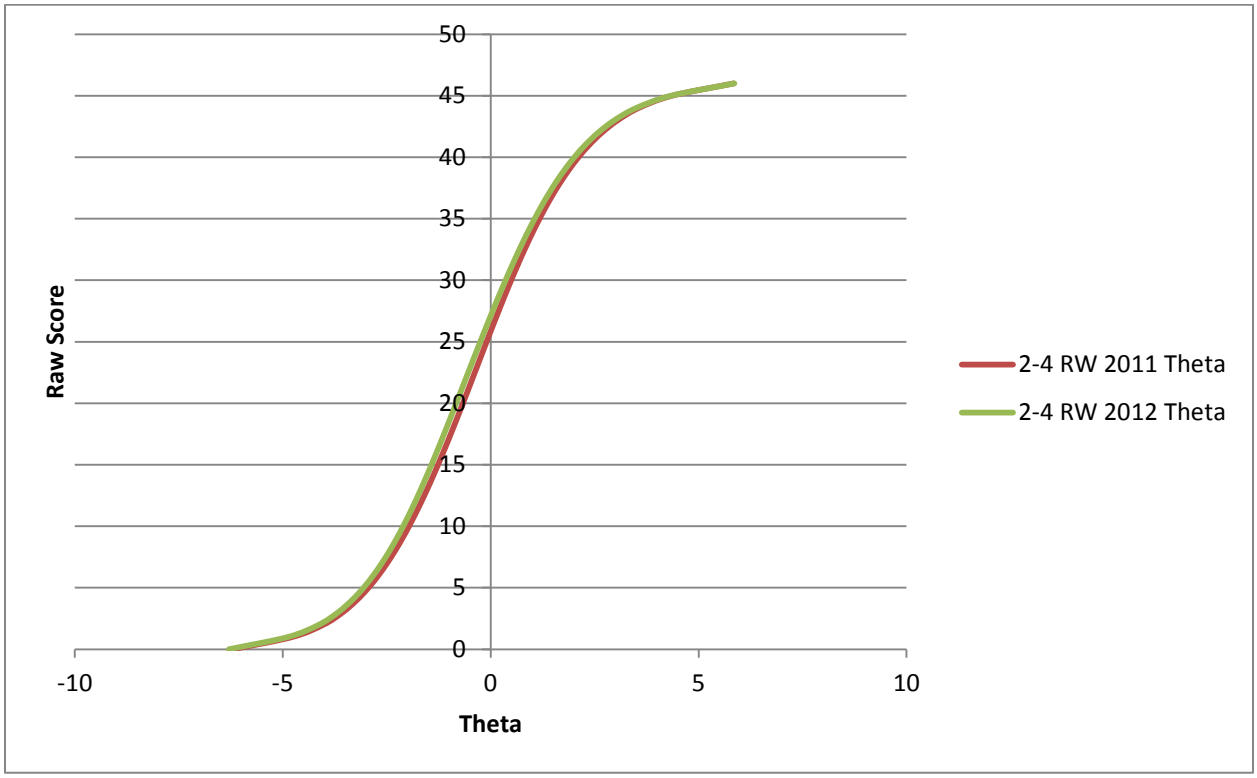


Figure 12—TCC for Reading/Writing, Grade Band 5–6

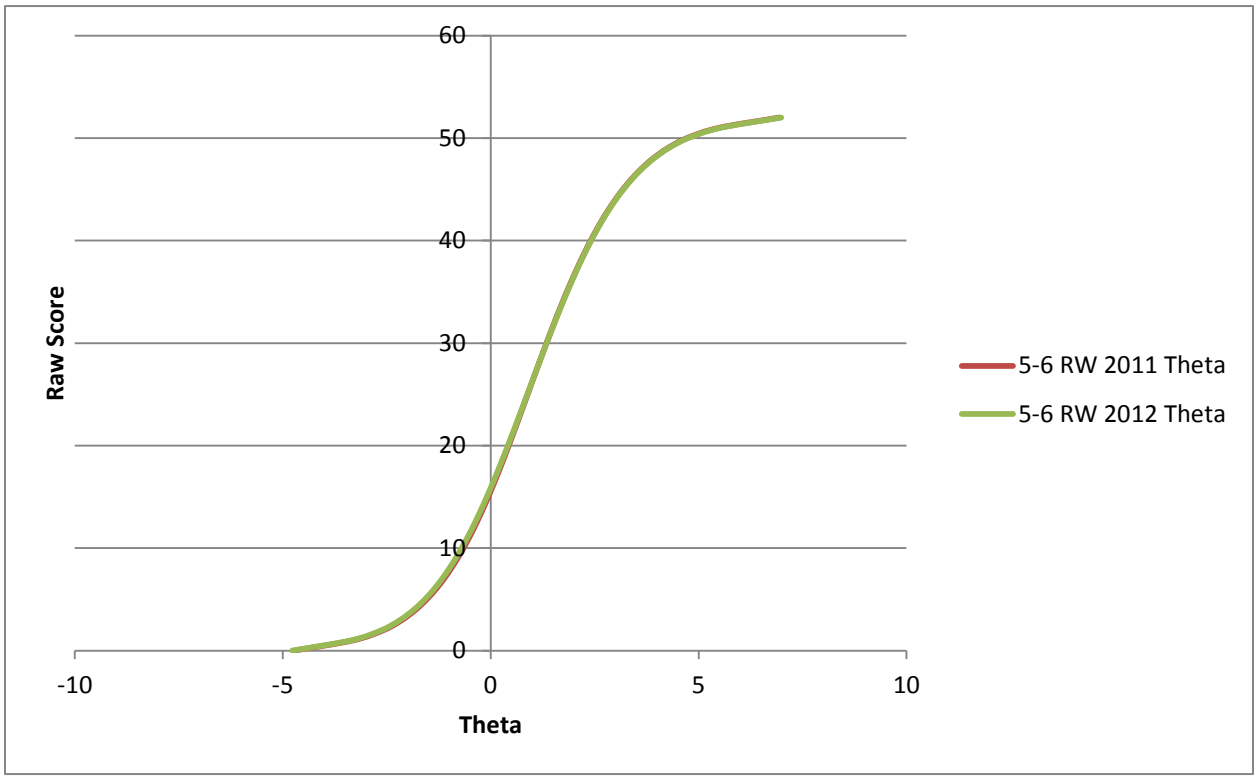


Figure 13—TCC for Reading/Writing, Grade Band 7–8

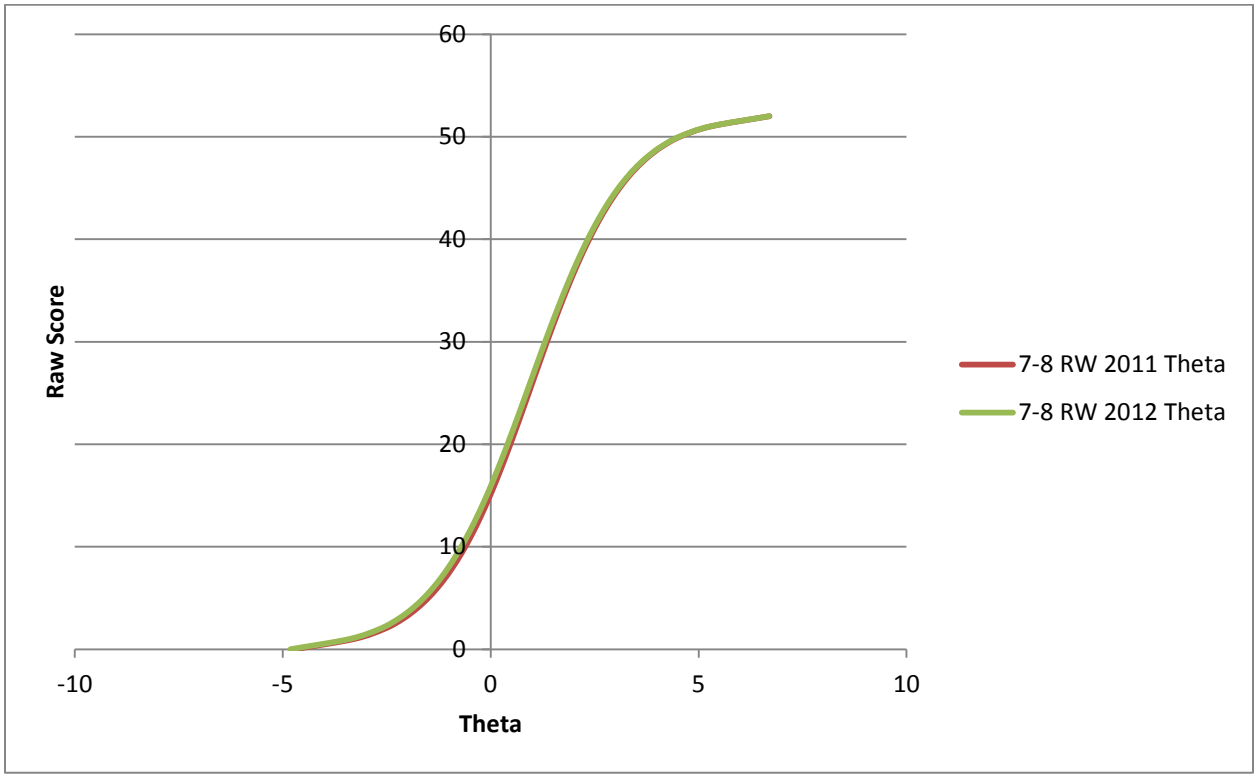
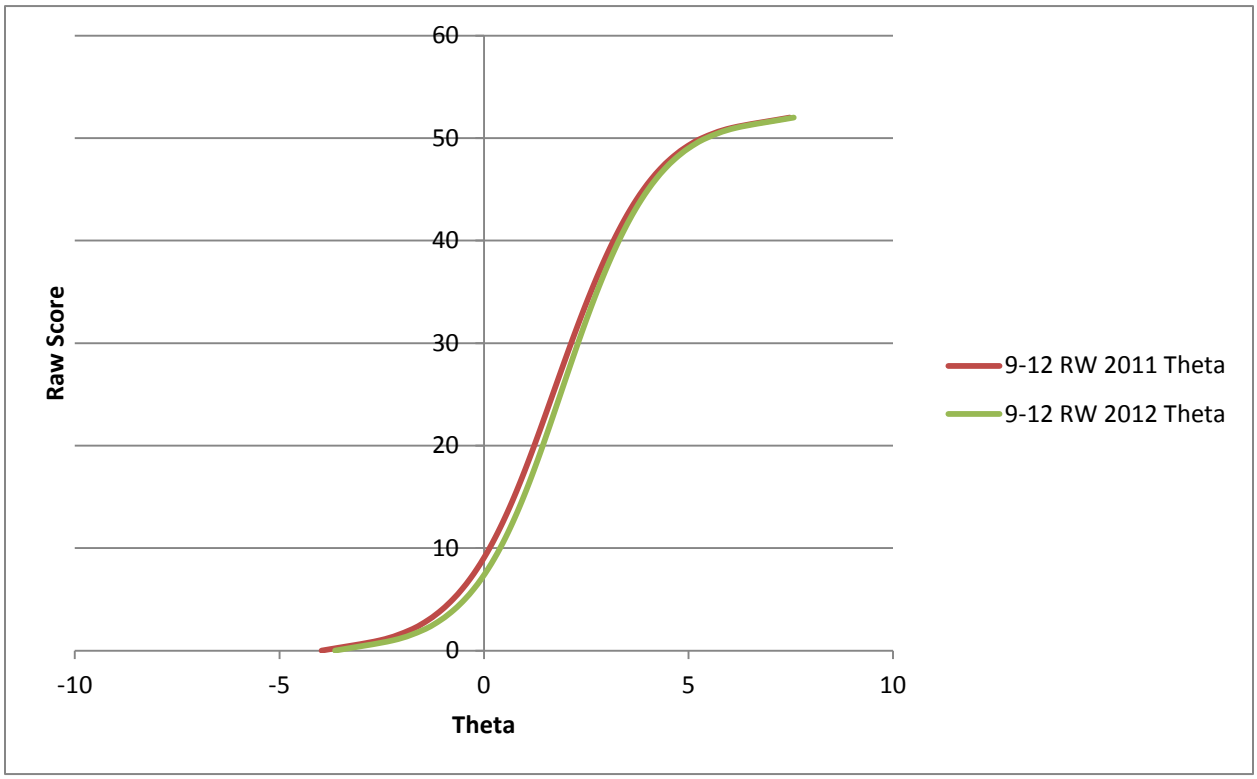


Figure 14—2011 TCC for Reading/Writing, Grade Band 9–12



7.5 Scaling of the NYSESLAT

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' grades 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:

$$\text{Scale Score} = 35*(\text{theta}) + 600 \tag{13}$$

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

Raw Score-to-Scale Score Tables

The final reporting scales were used to produce raw score-to-scale score conversion tables for the Listening/Speaking and Reading/Writing combinations. Appendix C provides the raw-to-scale score conversion tables for the Listening/Speaking and Reading/Writing combinations by grade band.

7.6 Rasch Information

Appendix D contains the results of the operational items for the 2012 NYSESLAT. The following IRT item parameters are presented for each item grouped by the Listening/Speaking and Reading/Writing combinations:

- N-count (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 39 presents the grade band, 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. As expected, the average Rasch difficulty was generally the lowest for grade band K—1, and the highest for grade band 9—12. Appendix E includes item information at different cut points at each grade.

Table 39—Average Rasch Difficulty by Grade Band and Modality

Grade Band	Modality	Number of Items	Max. Points	Average Rasch Difficulty
K–1	Listening	24	24	-2.56
	Speaking	16	34	-1.73
	Reading	15	15	-2.58
	Writing	15	24	-2.73
	Listening and Speaking	40	58	-2.23
	Reading and Writing	30	39	-2.65
	Total	70	97	-2.41
2–4	Listening	24	24	-0.10
	Speaking	16	34	-0.41
	Reading	24	24	-0.44
	Writing	16	22	-0.69
	Listening and Speaking	40	58	-0.22
	Reading and Writing	40	46	-0.54
	Total	80	104	-0.38
5–6	Listening	25	25	0.22
	Speaking	16	34	-0.46
	Reading	27	27	1.11
	Writing	19	25	0.67
	Listening and Speaking	41	59	-0.05
	Reading and Writing	46	52	0.93
	Total	87	111	0.47
7–8	Listening	25	25	1.21
	Speaking	16	34	0.61
	Reading	27	27	1.12
	Writing	19	25	0.56
	Listening and Speaking	41	59	0.98
	Reading and Writing	46	52	0.89
	Total	87	111	0.93
9–12	Listening	25	25	1.92
	Speaking	16	34	1.05
	Reading	27	27	2.08
	Writing	19	25	1.78
	Listening and Speaking	41	59	1.58
	Reading and Writing	46	52	1.95
	Total	87	111	1.78

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

CHAPTER 8: STANDARD SETTING

The previous vendor for the NYSESLAT carried out the standard setting study. This chapter has not been modified in order to maintain consistency with their representation of the standard setting activities.

The contractor for the NYSESLAT organized a performance standard setting meeting conducted in Albany, NY, from March 28–April 1, 2005. The purpose of this meeting was to provide preliminary recommendations on performance cut scores for the NYSESLAT. Each group had one psychometrics staff member from the contractor to facilitate the technical part of the standard setting. In addition, a content specialist from the vendor and a NYSED official were present to provide support during the standard setting sessions.

8.1 Standard Setting Model

Item mapping, a well-established method available for establishing performance standards, 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 the following:

- 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 (PLDs)

The item-mapping procedure orders items for each test into a booklet according to the difficulty of the items 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 also receive test blueprints and various other materials, including PLDs, scoring rubrics, test booklets, and anchor papers.

8.2 Committees of Panelists

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

NYSED recruited NYSESLAT panelists who were certified ESL, ELA, bilingual education, and bilingual special education teachers familiar with the content standards. The panelists possessed knowledge of working with ELL students and represented 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 per recommendations. These additional members tended to provide valuable insights from their area of expertise and helped strengthen the consequential validity argument during panel discussions.

Table 40—Panel Composition for Standard Setting Committees

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

8.3 Performance Levels and Cut Scores

The NYSESLAT has four performance levels that respond to three cut scores. The four performance levels are as follows:

- Beginning
- Intermediate
- Advanced
- Proficient

The three cut scores are as follows:

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

The item-mapping procedure was utilized in order to set the three cut points. The standard setting process is described below.

8.4 Standard Setting Process

Standard setting began with introductions from NYSED, the vendor, and panelists, followed by a presentation by the lead facilitator on the role of the panelists in the standard setting process, which was 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 with all five groups together.

After the orientation, panelists separated into specific breakout rooms according to their group assignments. An expert in the standard setting methodology facilitated each group/room. Assessment specialists rotated from group to group in order to provide content support. The panel members were further divided into three smaller table groups within their grade bands, each composed of five to seven members. These small groups worked independently but could collaborate with the other table groups in their grade band during the standard setting process. The following sequence of tasks was followed:

1. *Reviewing the Assessment*

   The first task was to review the assessment blueprint in order for 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 items were conducted. The panel members further defined the general PLDs 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.

2. *Experiencing the Assessment*

   Next, panel members experienced the assessment administered at the grade band 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.

3. *Scoring the Assessment*

   After the panelists finished taking the assessment, they were provided with an answer key so they could score their own assessments using the scoring rubrics and answer key provided. The scoring process allowed panelists to develop an understanding of the scoring of constructed responses. They were provided with exemplars of score points. A discussion session followed the scoring of the assessment.

4. *Reviewing the Student Performance Levels*

   Panelists reviewed the previously established definitions of performance levels, found in Appendices F.3–F.4 of the *2006 NYSESLAT Technical Report*. Then they discussed the performance levels 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 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 or 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 bands, the groups came together to discuss the process and results of their assigned grades among all grade bands. 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.

8.5 Summary Statistics for the Three Rounds of Ratings

Appendix F of the *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.

8.6 Post-Standard Setting Analyses

The median scores from the standard setting committees were used as the recommended cuts based on the total NYSESLAT score. After the standard setting meetings, the vendor 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 grades, K–12. Further analyses were performed to smooth the cuts 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 grade. 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 NYSED to help them make their final decisions on the cut points for the Listening/Speaking and Reading/Writing combinations.

8.7 Final Performance Level Cut Points

The final cut points adopted by NYSED for the Listening/Speaking and Reading/Writing combinations in raw score points, scale score, and theta metric (from standard setting) are presented in Tables 41 and 42. Three cut points correspond to four performance levels, and any score below the Intermediate cut point is the Beginning performance level. Questar applied the same scoring tables and cut points for the 2012 administration as reported in the 2009 technical report.

Table 41—Final Performance Level Cut Points for Listening/Speaking

	Raw Score Cut, 2011			Raw Score Cut, 2012			Scale Score			Theta		
Grade	**I**	**A**	**P**	**I**	**A**	**P**	**I**	**A**	**P**	**I**	**A**	**P**
K	22	42	51	24	44	52	511	573	616	-2.55	-0.76	0.47
1	25	44	55	27	46	55	520	582	649	-2.28	-0.52	1.39
2	13	32	48	14	34	49	543	602	661	-1.63	0.06	1.73
3	17	35	51	18	37	52	555	611	675	-1.27	0.31	2.14
4	20	38	52	22	40	53	567	621	684	-0.95	0.59	2.39
5	20	38	52	24	43	54	579	631	686	-0.59	0.89	2.46
6	23	41	54	27	45	55	587	639	696	-0.37	1.11	2.74
7	17	35	50	18	38	52	597	651	703	-0.08	1.45	2.95
8	19	38	50	20	41	52	605	661	703	0.15	1.74	2.95
9	22	41	50	20	40	50	623	678	715	0.66	2.23	3.28
10	24	43	51	22	42	51	628	686	721	0.80	2.44	3.46
11	26	45	53	25	45	52	635	694	728	1.00	2.68	3.66
12	28	48	53	26	47	52	640	704	728	1.14	2.96	3.66

Table 42—Final Performance Level Cut Points for Reading/Writing

	Raw Score Cut, 2011			Raw Score Cut, 2012			Scale Score			Theta		
Grade	**I**	**A**	**P**	**I**	**A**	**P**	**I**	**A**	**P**	**I**	**A**	**P**
K	25	32	35	26	33	36	537	582	609	-1.79	-0.52	0.25
1	29	35	38	30	36	38	559	609	655	-1.17	0.25	1.58
2	20	32	39	20	33	40	571	623	663	-0.82	0.64	1.81
3	22	33	41	23	34	41	580	630	679	-0.57	0.86	2.25
4	24	36	42	25	36	43	589	642	691	-0.33	1.19	2.61
5	23	35	44	22	34	44	620	660	704	0.56	1.73	2.98
6	24	37	45	23	36	45	623	668	710	0.66	1.94	3.15
7	25	38	45	25	39	46	629	674	710	0.83	2.13	3.15
8	26	40	45	26	40	46	632	682	710	0.92	2.34	3.15
9	22	38	44	18	36	43	643	698	726	1.21	2.80	3.60
10	23	39	45	19	37	44	646	702	732	1.30	2.91	3.78
11	24	41	46	20	39	45	649	711	739	1.39	3.16	3.97
12	24	42	46	21	40	45	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.

CHAPTER 9: SUMMARY OF THE OPERATIONAL TEST RESULTS

This chapter contains raw score and scale score summaries for each of the reporting scales (Listening/Speaking and Reading/Writing).

- Table 43 details raw score summary by grade.
- Table 44 contains the raw score summary by grade band.
- Subsequently, Table 45 includes the sample size (n-count), the mean, median, interquartile range (IQR), range, and the standard deviation (SD) of the scale score.
- Table 46 presents the summary of the scale scores by grade band.
- 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 band. The overall proficiency level is defined as the lower of the two proficiency levels that a student achieves—e.g., the Listening/Speaking level = 3, the Reading/Writing level = 4, and the Overall level = 3.
- Table 49 presents the exit rate by grade for the 2010, 2011, and 2012 administrations, and Table 50 presents the exit rate by grade band for the 2010, 2011, and 2012 administrations. The two tables include n-counts and percentages of ELL students obtaining exit status after scoring at the Proficient Level (level 4) on both the Listening/Speaking and Reading/Writing modality combinations. Similar analyses were conducted according to the following subgroups (see Appendices G, H, and I):

    - Gender
    - Ethnicity: American Indian or Alaska Native, Black or African American, Asian, Native Hawaiian/other Pacific Islander, Hispanic or Latino, White, and Multiracial (not of Hispanic origin)
    - Needs/Resource Group: NYC, Large City, High Need Urban/Suburban, High Need Rural, Average Need, Low Need, and Charter Schools
    - Five major language backgrounds: Arabic, Bengali, Chinese, English, and 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, and traumatic brain injury

Table 43—Raw Score Summary by Grade

Grade	Test	N-Count	Mean	Median	IQR	Range	SD
K	Listening and Speaking	26,266	45.23	48	12	58	10.64
	Reading and Writing	26,266	25.91	27	12	39	8.10
1	Listening and Speaking	25,676	50.62	53	7	58	9.04
	Reading and Writing	25,676	34.05	37	5	39	6.52
2	Listening and Speaking	21,159	46.91	50	9	58	10.04
	Reading and Writing	21,159	29.26	31	12	46	8.91
3	Listening and Speaking	19,024	49.12	53	7	58	10.18
	Reading and Writing	19,024	33.68	36	11	46	8.83
4	Listening and Speaking	17,853	50.30	54	7	58	9.96
	Reading and Writing	17,853	36.01	39	9	46	8.40
5	Listening and Speaking	15,459	49.17	53	8	58	11.22
	Reading and Writing	15,459	37.08	40	14	52	10.77
6	Listening and Speaking	13,164	48.16	53	11	59	12.16
	Reading and Writing	13,164	36.53	40	15	52	11.02
7	Listening and Speaking	12,609	45.70	50	13	59	12.37
	Reading and Writing	12,609	33.77	36	16	52	10.66
8	Listening and Speaking	12,506	45.92	51	14	59	12.42
	Reading and Writing	12,506	34.85	37	15	52	10.64
9	Listening and Speaking	14,288	41.42	47	23	59	14.89
	Reading and Writing	14,288	30.66	32	20	52	11.68
10	Listening and Speaking	13,289	42.74	46	19	58	13.04
	Reading and Writing	13,289	32.10	34	17	52	10.92
11	Listening and Speaking	9,775	45.28	48	15	59	11.18
	Reading and Writing	9,775	34.87	37	14	52	10.14
12	Listening and Speaking	7,352	46.91	50	13	59	10.58
	Reading and Writing	7,352	35.21	38	13	52	10.72

Table 44—Raw Score Summary by Grade Band

Grade Band	Test	N-Count	Mean	Median	IQR	Range	SD
K–1	Listening and Speaking	51,942	47.90	51	10	58	10.24
	Reading and Writing	51,942	29.94	32	13	39	8.41
2–4	Listening and Speaking	58,036	48.67	52	8	58	10.16
	Reading and Writing	58,036	32.79	35	13	46	9.18
5–6	Listening and Speaking	28,623	48.70	53	9	59	11.67
	Reading and Writing	28,623	36.83	40	14	52	10.89
7–8	Listening and Speaking	25,115	45.81	50	14	59	12.39
	Reading and Writing	25,115	34.31	37	16	52	10.67
9–12	Listening and Speaking	44,704	43.56	48	18	59	13.09
	Reading and Writing	44,704	32.76	35	17	52	11.13

Note: The data file used in Tables 44 and 45, above, is the final research file. The total n-counts for Grades K–12 are 208,420. 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 invalid codes).

Table 45—Scale Score Summary by Grade

Grade	Test	N-Count	Mean	Median	IQR	Range	SD
K	Listening and Speaking	26,266	596.75	595	66	430	55.34
	Reading and Writing	26,266	550.42	546	65	403	54.64
1	Listening and Speaking	25,676	636.63	630	71	430	61.76
	Reading and Writing	25,676	626.92	636	79	403	67.76
2	Listening and Speaking	21,159	662.67	666	50	413	48.21
	Reading and Writing	21,159	613.99	617	56	425	45.09
3	Listening and Speaking	19,024	680.13	688	54	413	54.41
	Reading and Writing	19,024	639.46	643	62	425	49.71
4	Listening and Speaking	17,853	691.41	697	64	413	57.72
	Reading and Writing	17,853	655.74	662	63	425	52.44
5	Listening and Speaking	15,459	675.93	679	55	353	54.05
	Reading and Writing	15,459	681.49	685	62	411	48.58
6	Listening and Speaking	13,164	672.66	679	67	403	57.91
	Reading and Writing	13,164	679.53	685	65	411	50.01
7	Listening and Speaking	12,609	692.24	697	60	399	51.72
	Reading and Writing	12,609	663.34	666	58	403	42.66
8	Listening and Speaking	12,506	694.86	702	69	399	54.20
	Reading and Writing	12,506	668.03	670	57	403	43.77
9	Listening and Speaking	14,288	698.24	705	90	407	61.30
	Reading and Writing	14,288	686.51	687	67	392	43.95
10	Listening and Speaking	13,289	702.07	701	74	357	55.84
	Reading and Writing	13,289	691.45	693	58	392	41.59
11	Listening and Speaking	9,775	710.89	709	67	407	50.64
	Reading and Writing	9,775	701.64	704	52	392	40.64
12	Listening and Speaking	7,352	718.04	718	66	407	50.61
	Reading and Writing	7,352	702.48	707	49	392	44.25

Table 46—Scale Score Summary by Grade Band

Grade Band	Test	N-Count	Mean	Median	IQR	Range	SD
K–1	Listening and Speaking	51,942	616.47	614	74	430	61.90
	Reading and Writing	51,942	588.23	577	105	403	72.41
2–4	Listening and Speaking	58,036	677.23	680	58	413	54.63
	Reading and Writing	58,036	635.18	637	70	425	51.93
5–6	Listening and Speaking	28,623	674.43	679	60	403	55.88
	Reading and Writing	28,623	680.59	685	62	411	49.25
7–8	Listening and Speaking	25,115	693.55	697	69	399	52.98
	Reading and Writing	25,115	665.67	670	60	403	43.28
9–12	Listening and Speaking	44,704	705.40	709	76	407	56.23
	Reading and Writing	44,704	693.91	697	59	392	43.11

Note: (1) Generally speaking, the mean should increase from one grade/grade band to the next higher grade/grade band in a similar manner, as shown in Tables 42 and 43 of this report, which depicts increases across the grades. However, because of artifacts of the population whereby some grades/grade bands may have a greater percentage of higher-scoring students than the next higher grade/grade band, the mean for the lower grade/grade band can be higher than the next higher grades/grade bands. (2) The statistics in Tables 46 and 47 are based on the final research file. The total n-counts for Grades K–12 are 208,420. 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 invalid codes).

Table 47—Percentage of Students in Each Proficiency Level by Grade

			Proficiency Level (Percent)			
Grade	**Test**	**N-Count**	**Beginning**	**Intermediate**	**Advanced**	**Proficient**
K	Listening and Speaking	26,266	5.65	26.45	35.47	32.43
	Reading and Writing	26,266	44.55	30.76	12.14	12.55
	Overall	26,266	44.91	33.02	13.85	8.23
1	Listening and Speaking	25,676	3.70	12.18	43.04	41.08
	Reading and Writing	25,676	17.20	23.26	19.82	39.72
	Overall	25,676	17.55	25.71	32.80	23.94
2	Listening and Speaking	21,159	2.42	6.59	32.97	58.01
	Reading and Writing	21,159	15.43	42.56	31.02	10.99
	Overall	21,159	15.51	42.51	31.62	10.36
3	Listening and Speaking	19,024	3.31	6.15	33.32	57.22
	Reading and Writing	19,024	12.72	24.42	39.98	22.89
	Overall	19,024	12.84	24.44	42.30	20.41
4	Listening and Speaking	17,853	3.71	6.80	29.46	60.03
	Reading and Writing	17,853	11.40	21.13	48.77	18.70
	Overall	17,853	11.61	21.20	50.18	17.02
5	Listening and Speaking	15,459	5.28	11.84	33.81	49.07
	Reading and Writing	15,459	12.07	17.47	35.09	35.38
	Overall	15,459	12.63	18.75	41.13	27.49
6	Listening and Speaking	13,164	8.58	14.78	37.36	39.28
	Reading and Writing	13,164	14.60	21.86	35.18	28.36
	Overall	13,164	15.64	23.52	41.42	19.42
7	Listening and Speaking	12,609	5.30	14.49	36.37	43.84
	Reading and Writing	12,609	21.70	37.29	29.52	11.49
	Overall	12,609	21.94	37.29	30.83	9.94
8	Listening and Speaking	12,506	5.82	18.47	30.15	45.55
	Reading and Writing	12,506	20.80	37.13	27.12	14.94
	Overall	12,506	21.16	37.43	28.30	13.11
9	Listening and Speaking	14,288	12.08	24.42	20.96	42.54
	Reading and Writing	14,288	18.23	40.65	22.88	18.24
	Overall	14,288	21.09	38.87	23.34	16.70
10	Listening and Speaking	13,289	8.51	29.87	24.19	37.43
	Reading and Writing	13,289	14.50	44.09	24.98	16.43
	Overall	13,289	16.68	44.44	24.97	13.91
11	Listening and Speaking	9,775	6.31	30.68	26.18	36.83
	Reading and Writing	9,775	9.85	45.86	26.82	17.46
	Overall	9,775	11.89	47.94	26.27	13.90
12	Listening and Speaking	7,352	5.47	30.81	20.78	42.94
	Reading and Writing	7,352	11.56	45.72	23.54	19.18
	Overall	7,352	12.64	48.52	23.46	15.38

Table 48—Percentage of Students in Each Proficiency Level by Grade Band

Grade Band	Test	N-Count	Proficiency Level (Percent)			
			Beginning	Intermediate	Advanced	Proficient
K–1	Listening and Speaking	51,942	4.69	19.39	39.21	36.70
	Reading and Writing	51,942	31.03	27.05	15.94	25.98
	Overall	51,942	31.38	29.41	23.22	16.00
2–4	Listening and Speaking	58,036	3.11	6.51	32.01	58.37
	Reading and Writing	58,036	13.30	30.02	39.42	17.27
	Overall	58,036	13.43	30.03	40.83	15.70
5–6	Listening and Speaking	28,623	6.80	13.19	35.44	44.57
	Reading and Writing	28,623	13.23	19.49	35.13	32.15
	Overall	28,623	14.02	20.94	41.26	23.77
7–8	Listening and Speaking	25,115	5.56	16.47	33.27	44.69
	Reading and Writing	25,115	21.25	37.21	28.33	13.21
	Overall	25,115	21.55	37.36	29.57	11.52
9–12	Listening and Speaking	44,704	8.67	28.46	23.03	39.84
	Reading and Writing	44,704	14.19	43.64	24.47	17.69
	Overall	44,704	16.38	44.09	24.49	15.04

Note: The statistics in Tables 48 and 49 are based on the final research file. The total n-counts for Grades K–12 are 208,420. 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 invalid codes).

Table 49—Exit Rate by Grade for the 2010, 2011, and 2012 Administrations

	2010		2011		2012	
Grade	N-Count	Percent	N-Count	Percent	N-Count	Percent
K	2,515	9.89	2,422	9.52	2,162	8.23
1	6,059	22.17	5,698	22.20	6,147	23.94
2	3,071	13.92	2,755	13.11	2,193	10.36
3	3,982	19.05	3,515	17.45	3,883	20.41
4	4,446	23.24	4,316	23.59	3,038	17.02
5	2,624	17.00	4,385	27.70	4,249	27.49
6	1,545	11.79	2,674	19.57	2,556	19.42
7	1,675	13.40	1,886	14.93	1,253	9.94
8	2,104	17.09	2,538	20.04	1,640	13.11
9	1,470	11.04	1,827	12.48	2,386	16.7
10	1,407	11.32	1,589	11.97	1,849	13.91
11	1,187	13.26	1,048	11.81	1,359	13.9
12	917	15.25	969	15.00	1,131	15.38
Total	33,002	15.8	35,622	16.80	33,846	16.24

Table 50—Exit Rate by Grade Band for the 2010, 2011, and 2012 Administrations

Grade Band	2010		2011		2012	
	N-Count	Percent	N-Count	Percent	N-Count	Percent
K–1	8,574	16.25	8,120	15.89	8,309	16.00
2–4	11,499	18.52	10,586	17.81	9,114	15.70
5–6	4,169	14.61	7,059	23.93	6,805	23.77
7–8	3,779	15.23	4,424	17.49	2,893	11.52
9–12	4,981	12.24	5,433	12.56	6,725	15.04
Total	33,002	15.80	35,622	17.00	33,846	16.24

APPENDIX A: ITEM MAP

A.1: 2012 NYSESLAT Operational Test Information for Grades K–1

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
Speaking						
1	SS	2	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Sentence Completion
2	SS	2	2	7	Present personal responses to published literature through words or pictures, referring to features of the text	Sentence Completion
3	SS	2	1	9	Convey information, using some organizational patterns and structures	Sentence Completion
4	SS	2	3	9	Apply learning strategies to explore a variety of materials	Sentence Completion
6	SS	2	1	9	Convey information, using some organizational patterns and structures	Sentence Completion
7	SS	4	2	8	Create personal stories, using appropriate vocabulary and elements of the literature students have read or heard	Storytelling
8	SS	2	5	3	Share cross-cultural experiences and ideas with others	Picture Description
9	SS	2	5	4	Interpret and demonstrate knowledge of nonverbal communication, and understand the contexts in which they are used appropriately	Picture Description
10	SS	2	2	5	Make predictions and inferences, and discuss the meaning of literary works to understand text presented orally and in written form	Picture Description
11	SS	2	3	3	Recognize personal point of view in self and others in discussing information	Picture Description
12	SS	2	4	2	Describe, read about, participate in, or recommend a favorite activity, book, song, or other interest to various audiences	Picture Description
13	SS	2	4	2	Describe, read about, participate in, or recommend a favorite activity, book, song, or other interest to various audiences	Social Interaction
14	SS	2	4	2	Describe, read about, participate in, or recommend a favorite activity, book, song, or other interest to various audiences	Social Interaction
15	SS	2	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction
17	SS	2	5	3	Share cross-cultural experiences and ideas with others	Social Interaction
18	SS	2	4	2	Describe, read about, participate in, or recommend a favorite activity, book, song, or other interest to various audiences	Social Interaction
Listening						
1	MC	1	1	12	Become familiar with some conventions of American English	First Sound
2	MC	1	1	12	Become familiar with some conventions of American English	First Sound
3	MC	1	1	4	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/Sentence Comprehension

A.1: 2012 NYSESLAT Operational Test Information for Grades K–1

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
4	MC	1	1	4	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/Sentence Comprehension
5	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word/Sentence Comprehension
6	MC	1	3	1	Form and express responses to ideas through reading, listening, viewing, discussing, and writing	Word/Sentence Comprehension
7	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word/Sentence Comprehension
8	MC	1	2	2	Use basic reading and listening strategies to make literary text comprehensible and meaningful	Word/Sentence Comprehension
9	MC	1	1	4	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/Sentence Comprehension
10	MC	1	2	2	Use basic reading and listening strategies to make literary text comprehensible and meaningful	Word/Sentence Comprehension
13	MC	1	1	4	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/Sentence Comprehension
14	MC	1	1	4	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/Sentence Comprehension
15	MC	1	2	2	Use basic reading and listening strategies to make literary text comprehensible and meaningful	Word/Sentence Comprehension
16	MC	1	1	4	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Word/Sentence Comprehension
17	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities	Comprehension of Conversational Language
18	MC	1	2	2	Use basic reading and listening strategies to make literary text comprehensible and meaningful	Comprehension of Conversational Language
19	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conversational Language
20	MC	1	1	4	Compare, contrast, and categorize to gain a deeper understanding of information and objects	Comprehension of Conversational Language
21	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conversational Language
22	MC	1	3	9	Apply learning strategies to explore a variety of materials	Comprehension of Conversational Language
23	MC	1	5	2	Demonstrate familiarity with some U.S. cultural referents at the local and national levels	Comprehension of Conversational Language
26	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities	Comprehension of Conversational Language

A.1: 2012 NYSESLAT Operational Test Information for Grades K–1

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
27	MC	1	5	2	Demonstrate familiarity with some U.S. cultural referents at the local and national levels	Comprehension of Conversational Language
28	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities	Comprehension of Conversational Language
Reading						
1	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
2	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
3	MC	1	1	3	Select information appropriate to the purpose of the investigation	Word Reading
4	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
5	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
6	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
7	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
8	MC	1	1	16	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Word Reading
12	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
13	MC	1	1	11	Express and develop ideas and understanding, using some elements of the “writing process”	Word Reading
14	MC	1	1	1	Identify and use basic reading and listening strategies to make text comprehensible and meaningful	Word Reading
15	MC	1	1	16	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Sentence Reading
16	MC	1	1	16	Apply learning strategies to acquire information and make oral and written texts comprehensible and meaningful	Sentence Reading
20	MC	1	2	2	Use basic reading and listening strategies to make literary text comprehensible and meaningful	Short Passage with Question
21	MC	1	3	5	Recognize how structural features affect readers’ and listeners’ understanding and appreciation of text	Short Passage with Question
Writing						
1	MC	1	1	12	Become familiar with some conventions of American English	Phonemic Understanding
2	MC	1	1	12	Become familiar with some conventions of American English	Phonemic Understanding

A.1: 2012 NYSESLAT Operational Test Information for Grades K–1

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
3	MC	1	1	12	Become familiar with some conventions of American English	Phonemic Understanding
4	MC	1	1	12	Become familiar with some conventions of American English	Phonemic Understanding
5	MC	1	1	12	Become familiar with some conventions of American English	Phonemic Understanding
6	MC	1	1	12	Become familiar with some conventions of American English	Phonemic Understanding
7	CR	2	1	12	Become familiar with some conventions of American English	Developmental Writing
8	CR	2	1	12	Become familiar with some conventions of American English	Developmental Writing
9	CR	2	1	12	Become familiar with some conventions of American English	Developmental Writing
10	CR	2	1	12	Become familiar with some conventions of American English	Developmental Writing
11	CR	2	1	12	Become familiar with some conventions of American English	Developmental Writing
12	CR	2	1	12	Become familiar with some conventions of American English	Developmental Writing
13	CR	2	1	12	Become familiar with some conventions of American English	Developmental Writing
14	CR	2	1	12	Become familiar with some conventions of American English	Developmental Writing
15	ER	4	3	1	Form and express responses to ideas through reading, listening, viewing, discussing, and writing	Extended Response

A.1: 2012 NYSESLAT Operational Test Information for Grades 2–4

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
Speaking						
1	SS	2	4	10	Demonstrate appropriate classroom behaviors (e.g., participating in small group and whole class discussions, being courteous, respecting the person and property of others)	Sentence Completion
2	SS	2	1	9	Convey information, using a variety of organizational patterns and structures	Sentence Completion
3	SS	2	1	9	Convey information, using a variety of organizational patterns and structures	Sentence Completion
4	SS	2	1	9	Convey information, using a variety of organizational patterns and structures	Sentence Completion
6	SS	2	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Sentence Completion
7	SE	4	2	8	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
8	SS	2	5	4	Interpret and demonstrate knowledge of nonverbal and oral communication features, and understand the contexts in which they are used appropriately	Picture Description
9	SS	2	1	9	Convey information, using a variety of organizational patterns and structures	Picture Description
10	SS	2	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Picture Description
11	SS	2	2	8	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	Picture Description
12	SS	2	1	6	Make and support inferences about information and ideas with reference to features in oral and written text.	Picture Description
13	SS	2	4	9	Use appropriate vocabulary, expressions, language, routines, and interaction styles for various audiences and formal and informal social or school situations	Social Interaction
14	SS	2	4	9	Use appropriate vocabulary, expressions, language, routines, and interaction styles for various audiences and formal and informal social or school situations	Social Interaction
15	SS	2	4	2	Describe, read about, participate in, or recommend a favorite activity, book, song, or other interest to various audiences	Social Interaction
16	SS	2	3	3	Recognize personal point of view in self and others in discussing, interpreting, and evaluating information	Social Interaction
18	SS	2	4	2	Describe, read about, participate in, or recommend a favorite activity, book, song, or other interest to various audiences	Social Interaction

Listening						
1	MC	1	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Word Sentence Comprehension
2	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word Sentence Comprehension
3	MC	1	1	3	Select information appropriate to the purpose of the investigation, and relate ideas from one written or spoken source to another	Word Sentence Comprehension
6	MC	1	1	4	Compare, contrast, and categorize, to gain a deeper understanding of information and objects	Word Sentence Comprehension
7	MC	1	1	4	Compare, contrast, and categorize, to gain a deeper understanding of information and objects	Word Sentence Comprehension
8	MC	1	2	1	Read, listen to, view, write about, and discuss a variety of literature of different genres	Word Sentence Comprehension
9	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Word Sentence Comprehension
10	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word Sentence Comprehension
11	MC	1	2	5	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
12	MC	1	1	16	Apply learning strategies to acquire information and make oral and written text comprehensible and meaningful	Word Sentence Comprehension
15	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conversational Language
16	MC	1	5	3	Recognize and share cross-cultural experiences, and ideas, and connect with those of others	Comprehension of Conversational Language
17	MC	1	2	5	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 of Conversational Language
18	MC	1	4	7	Follow oral and written direction to participate in classroom and social activities	Comprehension of Conversational Language
19	MC	1	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Comprehension of Conversational Language
20	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conversational Language
21	MC	1	4	7	Follow oral and written direction to participate in classroom and social activities	Comprehension of Conversational Language
22	MC	1	5	2	Demonstrate familiarity with a broad range of U.S. cultural and political referents through institutions, functions, and processes at the local and national levels	Comprehension of Conversational Language

23	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conversational Language
24	MC	1	4	7	Follow oral and written direction to participate in classroom and social activities	Task-based Listening
25	MC	1	4	7	Follow oral and written direction to participate in classroom and social activities	Task-based Listening
26	MC	1	4	7	Follow oral and written direction to participate in classroom and social activities	Task-based Listening
27	MC	1	4	7	Follow oral and written direction to participate in classroom and social activities	Task-based Listening
28	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Task-based Listening
Reading						
1	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word Sentence Reading
2	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word Sentence Reading
3	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Word Sentence Reading
7	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word Sentence Reading
8	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word Sentence Reading
9	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word Sentence Reading
10	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Word Sentence Reading
14	MC	1	2	12	Apply learning strategies to comprehend and make inferences about literature and produce literary responses	Comprehension
15	MC	1	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Comprehension
16	MC	1	2	4	Locate and identify key literary elements in texts and relate those elements to those in other works and to students' own experiences	Comprehension
17	MC	1	2	5	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
18	MC	1	3	5	Recognize and explain how structural features affect readers' and listeners' understanding and appreciation of text	Comprehension
19	MC	1	2	12	Apply learning strategies to comprehend and make inferences about literature and produce literary responses	Comprehension
20	MC	1	2	12	Apply learning strategies to comprehend and make inferences about literature and produce literary responses	Comprehension

21	MC	1	2	5	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
22	MC	1	2	4	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	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Comprehension
24	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities.	Comprehension
25	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
26	MC	1	2	4	Locate and identify key literary elements in texts and relate those elements to those in other works and to students' own experiences	Comprehension
27	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
28	MC	1	1	3	Select information appropriate to the purpose of the investigation, and relate ideas from one written or spoken source to another	Comprehension
29	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
30	MC	1	3	5	Recognize and explain how structural features affect readers' and listeners' understanding and appreciation of text	Comprehension
Writing						
1	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English	Phonemic Understanding
2	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Phonemic Understanding
3	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English	Phonemic Understanding
4	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English	Phonemic Understanding
5	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English	Mechanics & Structure
6	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English	Mechanics & Structure
7	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English	Mechanics & Structure
8	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English	Mechanics & Structure
9	MC	1	1	15	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 & Structure

10	MC	1	1	15	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 & Structure
11	MC	1	1	15	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 & Structure
12	MC	1	3	1	Form and express responses to a variety of literary, informational, and persuasive material through reading, listening, viewing, discussing, and writing; use details and evidence as support	Mechanics & Structure
13	CR	2	3	1	Form and express responses to a variety of literary, informational, and persuasive material through reading, listening, viewing, discussing, and writing; use details and evidence as support	Pre-Writing
14	CR	2	3	1	Form and express responses to a variety of literary, informational, and persuasive material through reading, listening, viewing, discussing, and writing; use details and evidence as support	Pre-Writing
15	CR	2	3	1	Form and express responses to a variety of literary, informational, and persuasive material through reading, listening, viewing, discussing, and writing; use details and evidence as support	Pre-Writing
16	ER	4	3	1	Form and express responses to a variety of literary, informational, and persuasive material through reading, listening, viewing, discussing, and writing; use details and evidence as support	Extended Response

A.1: 2012 NYSESLAT Operational Test Information for Grades 5–6

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
Speaking						
1	SS	2	2	8	Create stories, poems, songs, and plays, including those that reflect traditional and popular American culture, observing the conventions of the genre; create an effective voice, using a variety of writing styles appropriate to different audiences, purposes, and settings	Sentence Completion
2	SS	2	4	5	Explain actions, choices, and decisions in social and academic situations	Sentence Completion
3	SS	2	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Sentence Completion
4	SS	2	4	5	Explain actions, choices, and decisions in social and academic situations	Sentence Completion
5	SS	2	1	5	Formulate, ask, and respond to various question forms to obtain, clarify, and extend information and meaning	Sentence Completion
6	SE	4	2	8	Create stories, poems, songs, and plays, including those that reflect traditional and popular American culture, observing the conventions of the genre; create an effective voice, using a variety of writing styles appropriate to different audiences, purposes, and settings	Storytelling
7	SS	2	5	2	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 the students' native community	Picture Description
8	SS	2	4	5	Explain actions, choices, and decisions in social and academic situations	Picture Description
9	SS	2	3	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information in texts and presentations	Picture Description
10	SS	2	4	5	Explain actions, choices, and decisions in social and academic situations	Picture Description
11	SS	2	2	5	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	Picture Description
12	SS	2	4	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
13	SS	2	4	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
14	SS	2	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction

A.1: 2012 NYSESLAT Operational Test Information for Grades 5–6

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
15	SS	2	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction
16	SS	2	4	2	Describe, read about, participate in, or recommend a favorite activity, book, song, or other interest to various audiences	Social Interaction
Listening						
1	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word Sentence Comprehension
2	MC	1	1	4	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Word Sentence Comprehension
3	MC	1	3	5	Recognize, explain, and evaluate how structural features affect readers' and listeners' understanding and appreciation of text	Word Sentence Comprehension
4	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conversational Language
5	MC	1	1	3	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	1	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	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Comprehension of Conversational Language
8	MC	1	5	2	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 the students' native community	Comprehension of Conversational Language
9	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Comprehension of Conversational Language
10	MC	1	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Comprehension of Conversational Language
11	MC	1	4	8	Negotiate and manage interactions to accomplish social and classroom tasks	Comprehension of Conversational Language
12	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension of Conversational Language
13	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Comprehension of Conversational Language
14	MC	1	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Comprehension of Conversational Language

A.1: 2012 NYSESLAT Operational Test Information for Grades 5–6

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
15	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Comprehension of Conversational Language
16	MC	1	1	7	Present information clearly in a variety of oral and written forms for different audiences and purposes related to all academic content areas	Task-based Listening
17	MC	1	1	7	Present information clearly in a variety of oral and written forms for different audiences and purposes related to all academic content areas	Task-based Listening
18	MC	1	1	7	Present information clearly in a variety of oral and written forms for different audiences and purposes related to all academic content areas	Task-based Listening
19	MC	1	1	7	Present information clearly in a variety of oral and written forms for different audiences and purposes related to all academic content areas	Task-based Listening
20	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Task-based Listening
21	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Task-based Listening
22	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based Listening
23	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Task-based Listening
24	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based Listening
25	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Task-based Listening
Reading						
1	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/Sentence Reading
2	MC	1	1	16	Apply learning strategies to acquire information and make texts comprehensible and meaningful	Word/Sentence Reading
3	MC	1	1	4	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Word/Sentence Reading
4	MC	1	1	16	Apply learning strategies to acquire information and make texts comprehensible and meaningful	Comprehension
5	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension

6	MC	1	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Comprehension
7	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
8	MC	1	3	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information in texts and presentations	Comprehension
9	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
10	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension
11	MC	1	1	16	Apply learning strategies to acquire information and make texts comprehensible and meaningful	Comprehension
12	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
13	MC	1	3	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information in texts and presentations	Comprehension
14	MC	1	2	12	Apply learning strategies to comprehend and make inferences about literature and produce literary responses	Comprehension
15	MC	1	3	5	Recognize, explain, and evaluate how structural features affect readers' and listeners' understanding and appreciation of text	Comprehension
16	MC	1	2	4	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
17	MC	1	2	4	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
18	MC	1	2	5	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
19	MC	1	1	3	Select information appropriate to the purpose of the investigation, relate ideas from one written or spoken source to another, and exclude nonessential information	Comprehension
20	MC	1	1	4	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Comprehension
21	MC	1	1	3	Select information appropriate to the purpose of the investigation, relate ideas from one written or spoken source to another, and exclude nonessential information	Comprehension
22	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension

23	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
24	MC	1	2	4	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
25	MC	1	2	5	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	12	Apply learning strategies to comprehend and make inferences about literature and produce literary responses	Comprehension
27	MC	1	2	5	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
Writing						
1	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Phonemic Understanding
2	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Phonemic Understanding
3	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
4	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
5	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
6	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
7	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
8	MC	1	1	15	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 & Structure
9	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
10	MC	1	1	15	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 & Structure

11	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
12	MC	1	1	15	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 & Structure
13	MC	1	1	15	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 & Structure
14	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
15	MC	1	1	15	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 & Structure
16	CR	2	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Pre-Writing
17	CR	2	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Pre-Writing
18	CR	2	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Pre-Writing
19	ER	4	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Extended Response

A.1: 2012 NYSESLAT Operational Test Information for Grades 7–8

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
Speaking						
1	SS	2	2	8	Create stories, poems, songs, and plays, including those that reflect traditional and popular American culture, observing the conventions of the genre; create an effective voice, using a variety of writing styles appropriate to different audiences, purposes, and settings	Sentence Completion
2	SS	2	2	8	Create stories, poems, songs, and plays, including those that reflect traditional and popular American culture, observing the conventions of the genre; create an effective voice, using a variety of writing styles appropriate to different audiences, purposes, and settings	Sentence Completion
3	SS	2	4	5	Explain actions, choices, and decisions in social and academic situations	Sentence Completion
4	SS	2	2	8	Create stories, poems, songs, and plays, including those that reflect traditional and popular American culture, observing the conventions of the genre; create an effective voice, using a variety of writing styles appropriate to different audiences, purposes, and settings	Sentence Completion
5	SS	2	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Sentence Completion
6	SE	4	2	8	Create stories, poems, songs, and plays, including those that reflect traditional and popular American culture, observing the conventions of the genre; create an effective voice, using a variety of writing styles appropriate to different audiences, purposes, and settings	Storytelling
7	SS	2	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Picture Description
8	SS	2	3	9	Apply learning strategies to examine and interpret a variety of materials	Picture Description
9	SS	2	1	5	Formulate, ask, and respond to various question forms to obtain, clarify, and extend information and meaning	Picture Description
10	SS	2	3	9	Apply learning strategies to examine and interpret a variety of materials	Picture Description
11	SS	2	5	3	Recognize and share cross-cultural experiences and ideas, and connect with those of others	Picture Description
12	SS	2	4	9	Use appropriate vocabulary, expressions, language, routines, and interaction 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	2	Describe, read about, participate in, or recommend a favorite activity, book, song, or other interest to various audiences	Social Interaction

A.1: 2012 NYSESLAT Operational Test Information for Grades 7–8

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
14	SS	2	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction
15	SS	2	4	9	Use appropriate vocabulary, expressions, language, routines, and interaction styles for various audiences and formal and informal social or school situations, noticing how intention is realized through language	Social Interaction
16	SS	2	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction
Listening						
1	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word Sentence Comprehension
2	MC	1	3	9	Apply learning strategies to examine and interpret a variety of materials	Word Sentence Comprehension
3	MC	1	3	9	Apply learning strategies to examine and interpret a variety of materials	Word Sentence Comprehension
4	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Comprehension of Conventional Language
5	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Comprehension of Conventional Language
6	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Comprehension of Conventional Language
7	MC	1	2	5	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 of Conventional Language
8	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conventional Language
9	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conventional Language
10	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conventional Language
11	MC	1	5	1	Demonstrate an understanding of cultural and language patterns and norms in American English, including different regional and social varieties of English	Comprehension of Conventional Language
12	MC	1	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Comprehension of Conventional Language
13	MC	1	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Comprehension of Conventional Language
14	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Comprehension of Conventional Language
15	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conventional Language

A.1: 2012 NYSESLAT Operational Test Information for Grades 7–8

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
16	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Task-based Listening
17	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Task-based Listening
18	MC	1	4	7	Follow oral and written directions to participate in classroom and social activities, and provide directions to peers in selected interactions	Task-based Listening
19	MC	1	1	9	Convey and organize information, using facts, details, illustrative examples, and a variety of patterns and structures	Task-based Listening
20	MC	1	1	3	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 Listening
21	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based Listening
22	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Task-based Listening
23	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based Listening
24	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based Listening
25	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based Listening
Reading						
1	MC	1	3	9	Apply learning strategies to examine and interpret a variety of materials	Word/Sentence Reading
2	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Word/Sentence Reading
3	MC	1	3	9	Apply learning strategies to examine and interpret a variety of materials	Word/Sentence Reading
4	MC	1	2	12	Apply learning strategies to comprehend and make inferences about literature and produce literary responses	Comprehension
5	MC	1	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Comprehension
6	MC	1	2	4	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
7	MC	1	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Comprehension
8	MC	1	2	4	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

9	MC	1	3	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information in texts and presentations	Comprehension
10	MC	1	3	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information in texts and presentations	Comprehension
11	MC	1	1	10	Distinguish between fact and opinion, and relevant and irrelevant information, and exclude nonessential information in oral and written presentations	Comprehension
12	MC	1	3	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information in texts and presentations	Comprehension
13	MC	1	1	4	Compare, contrast, categorize, and synthesize to gain a deeper understanding of information and objects	Comprehension
14	MC	1	3	5	Recognize, explain, and evaluate how structural features affect readers' and listeners' understanding and appreciation of text	Comprehension
15	MC	1	1	16	Apply learning strategies to acquire information and make texts comprehensible and meaningful	Comprehension
16	MC	1	3	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information in texts and presentations	Comprehension
17	MC	1	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Comprehension
18	MC	1	2	1	Read, listen to, view, write about, and discuss texts and performances from a wide range of authors, subjects, and genres	Comprehension
19	MC	1	2	4	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
20	MC	1	2	12	Apply learning strategies to comprehend and make inferences about literature and produce literary responses	Comprehension
21	MC	1	2	5	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
22	MC	1	2	4	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
23	MC	1	2	12	Apply learning strategies to comprehend and make inferences about literature and produce literary responses	Comprehension
24	MC	1	2	5	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	12	Apply learning strategies to comprehend and make inferences about literature and produce literary responses	Comprehension
26	MC	1	2	12	Apply learning strategies to comprehend and make inferences about literature and produce literary responses	Comprehension
27	MC	1	2	2	Identify and use reading and listening strategies to make literary text comprehensible and meaningful	Comprehension
Writing						
1	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Phonemic Understanding
2	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Phonemic Understanding
3	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
4	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
5	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
6	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
7	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
8	MC	1	1	15	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 & Structure
9	MC	1	1	15	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 & Structure
10	MC	1	1	15	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 & Structure
11	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
12	MC	1	1	15	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 & Structure
13	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure

14	MC	1	1	15	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 & Structure
15	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
16	CR	2	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Pre-Writing
17	CR	2	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Pre-Writing
18	CR	2	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Pre-Writing
19	ER	4	3	1	Develop and present clear interpretations, analyses, and evaluations of issues, ideas, texts and experiences, supporting positions with well-developed arguments	Extended Response

A.1: 2012 NYSESLAT Operational Test Information for Grades 9–12

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
Speaking						
1	SS	2	1	12	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	2	8	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 audiences, purposes, and settings	Sentence Completion
3	SS	2	4	5	Explain actions, choices, and decisions in social and academic situations	Sentence Completion
4	SS	2	2	8	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 audiences, purposes, and settings	Sentence Completion
5	SS	2	3	9	Apply learning strategies to examine, interpret, analyze, synthesize, and evaluate a variety of materials	Sentence Completion
6	SE	4	2	8	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 audiences, purposes, and settings	Storytelling
7	SS	2	2	5	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	Picture Description
8	SS	2	2	8	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 audiences, purposes, and settings	Picture Description
9	SS	2	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Picture Description
10	SS	2	2	5	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	Picture Description
11	SS	2	2	8	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 audiences, purposes, and settings	Picture Description

A.1: 2012 NYSESLAT Operational Test Information for Grades 9–12

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
12	SS	2	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction
13	SS	2	4	9	Use appropriate vocabulary, expressions, language, routines, and interaction styles for various audiences and formal and informal social or school situations, noticing how intention is communicated in different ways through language in various contexts	Social Interaction
14	SS	2	4	8	Negotiate and manage interactions to accomplish social and classroom tasks	Social Interaction
15	SS	2	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Social Interaction
16	SS	2	4	5	Explain actions, choices, and decisions in social and academic situations	Social Interaction
Listening						
1	MC	1	3	9	Apply learning strategies to examine, interpret, analyze, synthesize, and evaluate a variety of materials	Word/Sentence Comprehension
2	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Word/Sentence Comprehension
3	MC	1	2	5	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	Word/Sentence Comprehension
4	MC	1	1	16	Apply learning strategies to acquire information and make texts comprehensible and meaningful	Comprehension of Conversational Language
5	MC	1	1	16	Apply learning strategies to acquire information and make texts comprehensible and meaningful	Comprehension of Conversational Language
6	MC	1	4	8	Negotiate and manage interactions to accomplish social and classroom tasks	Comprehension of Conversational Language
7	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension of Conversational Language
8	MC	1	5	1	Demonstrate an understanding of cultural and language patterns and norms in America English, including different regional and social varieties of English, and identify and interpret how these patterns and norms are used	Comprehension of Conversational Language
9	MC	1	5	1	Demonstrate an understanding of cultural and language patterns and norms in America English, including different regional and social varieties of English, and identify and interpret how these patterns and norms are used	Comprehension of Conversational Language
10	MC	1	4	7	Follow oral and written direction to participate in classroom activities, and provide directions to peers in selected interactions	Comprehension of Conversational Language
11	MC	1	4	7	Follow oral and written direction to participate in classroom activities, and provide directions to peers in selected interactions	Comprehension of Conversational Language

A.1: 2012 NYSESLAT Operational Test Information for Grades 9–12

Modality/ Position	Type	Score Point	Learning Standard	Performance Indicator	Performance Indicator Descriptor	Cluster
12	MC	1	4	6	Understand and use a variety of context-specific oral communication strategies in American English for a range of personal and academic purposes	Comprehension of Conversational Language
13	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension of Conversational Language
14	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Comprehension of Conversational Language
15	MC	1	4	7	Follow oral and written direction to participate in classroom activities, and provide directions to peers in selected interactions	Comprehension of Conversational Language
16	MC	1	4	7	Follow oral and written direction to participate in classroom activities, and provide directions to peers in selected interactions	Task-based Listening
17	MC	1	4	3	Request and provide information and assistance, orally or in writing, for personal, social, and academic purposes	Task-based Listening
18	MC	1	4	7	Follow oral and written direction to participate in classroom activities, and provide directions to peers in selected interactions	Task-based Listening
19	MC	1	4	7	Follow oral and written direction to participate in classroom activities, and provide directions to peers in selected interactions	Task-based Listening
20	MC	1	1	9	Convey and organize information, using facts, details, illustrative examples, and a variety of patterns and structures	Task-based Listening
21	MC	1	1	3	Select information appropriate to the purpose of the investigation with suitable supporting material	Task-based Listening
22	MC	1	4	7	Follow oral and written direction to participate in classroom activities, and provide directions to peers in selected interactions	Task-based Listening
23	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based Listening
24	MC	1	1	3	Select information appropriate to the purpose of the investigation with suitable supporting material	Task-based Listening
25	MC	1	1	1	Identify and use reading and listening strategies to make text comprehensible and meaningful	Task-based Listening
Reading						
1	MC	1	3	9	Apply learning strategies to examine, interpret, analyze, synthesize, and evaluate a variety of materials	Word/Sentence Reading
2	MC	1	1	4	Compare, contrast, categorize, and synthesize information and objects, and identify complexities and discrepancies in the information	Word/Sentence Reading
3	MC	1	3	9	Apply learning strategies to examine, interpret, analyze, synthesize, and evaluate a variety of materials	Word/Sentence Reading

4	MC	1	2	7	Compose and present personal and formal responses to and interpretations of published literary works and the work of peers, referring to details and features of text	Comprehension
5	MC	1	2	5	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
6	MC	1	2	10	Create, discuss, interpret, and respond to literary works, using appropriate and effective vocabulary, grammar, spelling, and punctuation in writing, and using appropriate vocabulary, grammar, and pronunciation in speaking	Comprehension
7	MC	1	2	4	Locate and identify a wide 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
8	MC	1	2	4	Locate and identify a wide 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
9	MC	1	5	5	Compare and contrast oral traditions, myths, folktales, and literature from different national and international regions and cultures, including the students' own, identifying similarities and differences and universal cultural themes, and exploring how language and literature transmit culture	Comprehension
10	MC	1	3	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information; make inferences about a writer's or speaker's point of view	Comprehension
11	MC	1	3	3	Recognize and communicate personal and multiple points of view within and among groups, in discussing, interpreting, and evaluating information; make inferences about a writer's or speaker's point of view	Comprehension
12	MC	1	3	5	Recognize, explain, evaluate, and analyze how structural features affect readers' and listeners' understanding and appreciation of text	Comprehension
13	MC	1	5	3	Recognize and share cross-cultural experiences and ideas, and connect with those of others	Comprehension
14	MC	1	1	16	Apply learning strategies to acquire information and make texts comprehensible and meaningful	Comprehension
15	MC	1	1	16	Apply learning strategies to acquire information and make texts comprehensible and meaningful	Comprehension
16	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Comprehension
17	MC	1	3	1	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	Comprehension

18	MC	1	3	2	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	3	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
20	MC	1	2	5	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
21	MC	1	2	5	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
22	MC	1	2	5	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	MC	1	2	7	Compose and present personal and formal responses to and interpretations of published literary works and the work of peers, referring to details and features of text	Comprehension
24	MC	1	3	5	Recognize, explain, evaluate, and analyze how structural features affect readers' and listeners' understanding and appreciation of text	Comprehension
25	MC	1	3	9	Apply learning strategies to examine, interpret, analyze, synthesize, and evaluate a variety of materials	Comprehension
26	MC	1	3	9	Apply learning strategies to examine, interpret, analyze, synthesize, and evaluate a variety of materials	Comprehension
27	MC	1	3	2	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
Writing						
1	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Phonemic Understanding
2	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Phonemic Understanding
3	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
4	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
5	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure

6	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
7	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
8	MC	1	1	15	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 & Structure
9	MC	1	1	15	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 & Structure
10	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
11	MC	1	1	12	Convey information and ideas through spoken and written language, using conventions and features of American English appropriate to audience and purpose	Mechanics & Structure
12	MC	1	1	15	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 & Structure
13	MC	1	1	15	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 & Structure
14	MC	1	1	6	Make and support inferences about information and ideas with reference to features in oral and written text	Mechanics & Structure
15	MC	1	1	15	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 & Structure
16	CR	2	3	1	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	Pre-Writing
17	CR	2	3	1	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	Pre-Writing
18	CR	2	3	1	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	Pre-Writing

19	ER	4	3	1	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: First Sound, 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 of 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 language proficiency.

Item Component	Description of Component
Stimulus	--
Stem	closed stem question, maximum 15 words: "Which picture shows a _____?" ; "Which girl is _____ [doing something]?" ; "Which [object] is _____ [adjective]?" ; etc.
Answer options	3 answer options consisting of 3 small individual graphics lettered A–C
Graphic	--

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, details, sequence, 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 band.

Item Component	Description of Component
Stimulus	simulated authentic conversation with one person speaking; 20–35 words (Grades K–1), 40–60 words (Grades 2–4), 50–80 words (Grades 5–8), 55–80 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	--

A.2: Item Specification (continued)

LISTENING: Synthesizing Information (Task-Based Listening), Grades 2–12
The test taker is required to perform real-world tasks through imagined manipulation of the graphics provided. 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 band.

Item Component	Description of Component
Advance organizer	1–2 sentence description (25–35 words) of an authentic context for the set of items; description appears in test booklet with a large overview graphic depicting the scene
Stimulus	simulated authentic conversation with one person speaking; 40–50 words (Grades 2–4), 50–90 words (Grades 5–8), 60–100 words (Grades 9–12), with distractor material embedded; there is one item per stimulus
Stem	closed stem question, maximum 15 words
Answer options	4 answer options consisting of either 4 small individual graphics lettered A–D
Graphic	large overview graphic depicting the scenario and containing elements of the items

READING: Word Reading, Grades K–4
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 language proficiency.

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

READING: Sentence Reading, Grades K–1
The test taker reads a question that includes a descriptive phrase, 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	--
Stem	closed stem question, maximum 10 words: “Which one _____?”; “Which picture shows _____?”; etc.
Answer options	3 answer options consisting of 3 small individual graphics
Graphic	--

A.2: Item Specification (continued)

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	--

READING: Word/Sentence Reading, Grades 2–12
The test taker reads a question that includes a key word or a descriptive phrase, and must choose which of three graphics correctly represents what is named or described. These items measure word-recognition skills, basic phrase, and sentence comprehension.

Item Component	Description of Component
Stimulus	--
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), 300 words (Grades 5–8), or 400 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 word represented by 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 language 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 language 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

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	4 answer options consisting of 1 sentence (or sentence fragment) each, maximum 20 words
Graphic	--

A.2: Item Specification (continued)

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 should 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 sentence with part of the sentence underlined; maximum 15 words
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 written English conventions.

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

WRITING: Letter Writing, Grades K–1
The test taker hears a letter name and writes the letter. These items measure understanding of written English conventions.

Item Component	Description of Component
Prompt	"__ is the first letter in the word ___. Write the letter __ in the box."
Graphic	--

WRITING: Word Writing, Grades K–1
The test taker sees a picture, hears the word depicted by the picture, and writes the word. These items measure understanding of written English conventions.

Item Component	Description of Component
Prompt	"This is a picture of a ____. Write the word ____ on the line."
Graphic	graphic depicting the dictated word

A.2: Item Specification (continued)

WRITING: Sentence Writing, Grades K–1
The test taker hears a dictated sentence, and writes the sentence. (The sentence may be declarative or interrogative.) The sentence is repeated three times. These items measure understanding of written English conventions.

Item Component	Description of Component
Prompt	"Listen to this sentence. _____ Now write the sentence on the lines. _____ (Pause.) _____"
Graphic	--

WRITING: Extended Response: Description and/or Narrative, Grades K–1
The test taker describes or tells a story about a picture. This item measures the student's ability to use vocabulary, grammar, features of extended discourse, and other writing conventions correctly and appropriately in order to describe a scene or tell a story.

Item Component	Description of Component
Prompt	graphic and directions: "Look at the picture. Write a story about what is happening in the picture. Write as much as you can."
Graphic	1 large graphic depicting a scene with multiple people in it

PRE-WRITING: Grades 2–12
The test taker briefly responds to three short questions or prompts 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: Exposition, Grades 2–12
The test taker compares and/or discusses activities of a particular type. This item measures the student's 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 . . . ." (2–4) or "Write an essay about . . . ." (5–12)
Graphic	2 graphics (photographs) depicting activities or situations related to the topic

A.2: Item Specification (continued)

SPEAKING: Sentence Completion, all grades
The test taker hears and reads the beginning of a sentence and must complete the sentence based on an accompanying graphic. These items measure the student's 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: Storytelling, 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 student's ability to use vocabulary, grammar, and discourse organizational features correctly and appropriately.

Item Component	Description of Component
Prompt	Graphics and Directions: "Look at the three pictures. Tell a story about what you see in the pictures. Tell what happened first, what happened next, and what happened last."
Graphic	3 sequenced graphics labeled First (1), Next (2), Last (3)

SPEAKING: Picture Description, all grades
The test taker looks at a photograph and orally responds to two questions (K–1) or a two-part question (2–12) about the picture. These items measure the student's ability to use vocabulary and grammar correctly and appropriately to describe everyday activities and situations.

Item Component	Description of Component
Prompt	a photograph and two oral questions (e.g., "What are the children doing? What might happen next?") or a two-part question (e.g., "What is the boy doing, and why?")
Graphic	photograph depicting a recognizable activity

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 fluency and the student's 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

Table B.1a: Item-Level Statistics—Listening and Speaking, Grades K–1

	Item #	Item Type	Max. Points	N-Count	% at 0/A	% at 1/B	% at 2/C	% at 3/D	% at 4	% Omit	P-value	Point Biserial
LISTENING	1	MC	1	51,942	0.04	0.02	0.93			0.00	0.93	0.38
	2	MC	1	51,942	0.03	0.94	0.01			0.01	0.94	0.38
	3	MC	1	51,942	0.02	0.03	0.95			0.00	0.95	0.46
	4	MC	1	51,942	0.03	0.95	0.01			0.00	0.95	0.39
	5	MC	1	51,942	0.91	0.07	0.02			0.00	0.91	0.49
	6	MC	1	51,942	0.13	0.83	0.03			0.00	0.83	0.48
	7	MC	1	51,942	0.03	0.92	0.04			0.00	0.92	0.54
	8	MC	1	51,942	0.95	0.02	0.02			0.00	0.95	0.47
	9	MC	1	51,942	0.05	0.07	0.88			0.00	0.88	0.43
	10	MC	1	51,942	0.01	0.95	0.04			0.00	0.95	0.41
	13	MC	1	51,942	0.84	0.13	0.02			0.00	0.84	0.48
	14	MC	1	51,942	0.81	0.05	0.14			0.00	0.81	0.35
	15	MC	1	51,942	0.06	0.03	0.90			0.00	0.90	0.44
	16	MC	1	51,942	0.07	0.86	0.07			0.00	0.86	0.36
	17	MC	1	51,942	0.03	0.03	0.94			0.00	0.94	0.47
	18	MC	1	51,942	0.03	0.03	0.94			0.00	0.94	0.50
	19	MC	1	51,942	0.87	0.09	0.04			0.01	0.87	0.39
	20	MC	1	51,942	0.92	0.04	0.03			0.01	0.92	0.38
	21	MC	1	51,942	0.05	0.05	0.90			0.01	0.90	0.38
	22	MC	1	51,942	0.85	0.04	0.10			0.00	0.85	0.42
	23	MC	1	51,942	0.09	0.22	0.69			0.00	0.69	0.50
	26	MC	1	51,942	0.74	0.12	0.13			0.01	0.74	0.36
	27	MC	1	51,942	0.09	0.17	0.74			0.01	0.74	0.46
	28	MC	1	51,942	0.89	0.04	0.06			0.00	0.89	0.33
SPEAKING	1	CR	2	51,942	0.05	0.24	0.71			0.00	0.83	0.67
	2	SS	2	51,942	0.06	0.21	0.73			0.00	0.83	0.70
	3	CR	2	51,942	0.07	0.50	0.43			0.00	0.68	0.63
	4	SS	2	51,942	0.07	0.26	0.67			0.00	0.80	0.72
	6	CR	2	51,942	0.04	0.54	0.42			0.00	0.69	0.58
	7	CR	4	51,942	0.05	0.09	0.27	0.36	0.22	0.00	0.66	0.80
	8	CR	2	51,942	0.04	0.21	0.75			0.00	0.86	0.72
	9	SS	2	51,942	0.06	0.33	0.61			0.00	0.77	0.74
	10	CR	2	51,942	0.07	0.29	0.64			0.00	0.78	0.75
	11	CR	2	51,942	0.05	0.26	0.69			0.00	0.82	0.74
	12	SS	2	51,942	0.06	0.36	0.58			0.00	0.76	0.73
	13	CR	2	51,942	0.05	0.11	0.84			0.00	0.89	0.64
	14	CR	2	51,942	0.15	0.12	0.74			0.00	0.80	0.72
	15	SS	2	51,942	0.06	0.11	0.84			0.00	0.89	0.71
	17	SS	2	51,942	0.08	0.21	0.72			0.00	0.82	0.72

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=208,420) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score > 900).

Table B.1b: Item-Level Statistics—Reading and Writing, Grades K–1

	Item #	Item Type	Max. Points	N-Count	% at 0/A	% at 1/B	% at 2/C	% at 3/D	% at 4	% Omit	P-value	Point Biserial
READING	1	MC	1	51,942	0.08	0.86	0.05			0.01	0.86	0.54
	2	MC	1	51,942	0.12	0.81	0.07			0.00	0.81	0.55
	3	MC	1	51,942	0.14	0.16	0.69			0.01	0.69	0.53
	4	MC	1	51,942	0.90	0.05	0.05			0.01	0.90	0.52
	5	MC	1	51,942	0.08	0.05	0.86			0.00	0.86	0.59
	6	MC	1	51,942	0.08	0.07	0.85			0.01	0.85	0.57
	7	MC	1	51,942	0.81	0.08	0.11			0.01	0.81	0.56
	8	MC	1	51,942	0.10	0.29	0.61			0.00	0.61	0.58
	12	MC	1	51,942	0.85	0.08	0.06			0.01	0.85	0.52
	13	MC	1	51,942	0.12	0.82	0.05			0.01	0.82	0.59
	14	MC	1	51,942	0.17	0.12	0.70			0.01	0.70	0.61
	15	MC	1	51,942	0.24	0.67	0.08			0.01	0.67	0.53
	16	MC	1	51,942	0.82	0.11	0.06			0.01	0.82	0.51
	20	MC	1	51,942	0.15	0.74	0.11			0.01	0.74	0.46
	21	MC	1	51,942	0.12	0.78	0.09			0.01	0.78	0.39
WRITING	1	MC	1	51,942	0.03	0.95	0.02			0.00	0.95	0.43
	2	MC	1	51,942	0.90	0.05	0.05			0.00	0.90	0.47
	3	MC	1	51,942	0.13	0.81	0.06			0.00	0.81	0.47
	4	MC	1	51,942	0.06	0.86	0.08			0.00	0.86	0.39
	5	MC	1	51,942	0.88	0.04	0.07			0.00	0.88	0.38
	6	MC	1	51,942	0.04	0.02	0.94			0.00	0.94	0.45
	7	CR	2	51,942	0.03	0.29	0.68			0.00	0.82	0.56
	8	CR	2	51,942	0.03	0.33	0.64			0.00	0.81	0.54
	9	CR	2	51,942	0.07	0.11	0.82			0.00	0.88	0.54
	10	CR	2	51,942	0.08	0.03	0.89			0.00	0.90	0.49
	11	CR	2	51,942	0.20	0.18	0.62			0.00	0.71	0.77
	12	CR	2	51,942	0.29	0.34	0.37			0.00	0.54	0.75
	13	CR	2	51,942	0.22	0.27	0.52			0.00	0.65	0.83
	14	CR	2	51,942	0.24	0.28	0.48			0.00	0.62	0.84
	15	CR	2	51,942	0.32	0.37	0.31			0.00	0.49	0.80

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=208,420) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score > 900).

Table B.2a: Item-Level Statistics—Listening and Speaking, Grades 2–4

	Item #	Item Type	Max. Points	N-Count	% at 0/A	% at 1/B	% at 2/C	% at 3/D	% at 4	% Omit	P-value	Point Biserial
LISTENING	1	MC	1	58,036	0.01	0.02	0.96			0.00	0.96	0.48
	2	MC	1	58,036	0.03	0.95	0.02			0.00	0.95	0.53
	3	MC	1	58,036	0.07	0.03	0.90			0.00	0.90	0.52
	6	MC	1	58,036	0.04	0.03	0.93			0.00	0.93	0.55
	7	MC	1	58,036	0.03	0.92	0.05			0.00	0.92	0.54
	8	MC	1	58,036	0.02	0.96	0.03			0.00	0.96	0.46
	9	MC	1	58,036	0.05	0.38	0.56			0.00	0.56	0.31
	10	MC	1	58,036	0.07	0.03	0.90			0.00	0.90	0.37
	11	MC	1	58,036	0.89	0.06	0.05			0.00	0.89	0.43
	12	MC	1	58,036	0.14	0.03	0.83			0.00	0.83	0.33
	15	MC	1	58,036	0.08	0.04	0.04	0.84		0.00	0.84	0.54
	16	MC	1	58,036	0.17	0.04	0.72	0.07		0.00	0.72	0.33
	17	MC	1	58,036	0.06	0.23	0.59	0.11		0.00	0.59	0.42
	18	MC	1	58,036	0.04	0.83	0.08	0.04		0.00	0.83	0.54
	19	MC	1	58,036	0.08	0.57	0.21	0.14		0.00	0.57	0.32
	20	MC	1	58,036	0.11	0.18	0.67	0.03		0.00	0.67	0.36
	21	MC	1	58,036	0.04	0.09	0.06	0.81		0.00	0.81	0.54
	22	MC	1	58,036	0.12	0.10	0.69	0.08		0.00	0.69	0.37
	23	MC	1	58,036	0.17	0.07	0.13	0.62		0.01	0.62	0.35
	24	MC	1	58,036	0.03	0.01	0.02	0.93		0.00	0.93	0.51
	25	MC	1	58,036	0.03	0.90	0.04	0.03		0.00	0.90	0.45
	26	MC	1	58,036	0.11	0.72	0.03	0.14		0.00	0.72	0.36
	27	MC	1	58,036	0.01	0.06	0.89	0.03		0.00	0.89	0.34
	28	MC	1	58,036	0.02	0.91	0.02	0.04		0.01	0.91	0.36
SPEAKING	1	CR	2	58,036	0.05	0.20	0.74			0.00	0.85	0.75
	2	SS	2	58,036	0.04	0.15	0.81			0.00	0.89	0.76
	3	SS	2	58,036	0.05	0.20	0.76			0.00	0.86	0.76
	4	CR	2	58,036	0.03	0.22	0.75			0.00	0.86	0.68
	6	CR	2	58,036	0.04	0.21	0.75			0.00	0.85	0.73
	7	CR	4	58,036	0.03	0.03	0.14	0.38	0.42	0.00	0.78	0.81
	8	CR	2	58,036	0.05	0.22	0.73			0.00	0.84	0.75
	9	SS	2	58,036	0.04	0.23	0.73			0.00	0.84	0.75
	10	CR	2	58,036	0.04	0.16	0.81			0.00	0.89	0.77
	11	CR	2	58,036	0.04	0.15	0.81			0.00	0.88	0.76
	12	CR	2	58,036	0.05	0.20	0.76			0.00	0.85	0.75
	13	SS	2	58,036	0.03	0.18	0.79			0.00	0.88	0.66
	14	CR	2	58,036	0.05	0.16	0.79			0.00	0.87	0.72
	15	CR	2	58,036	0.03	0.08	0.89			0.00	0.93	0.64
	16	SS	2	58,036	0.07	0.12	0.81			0.00	0.87	0.75
	18	CR	2	58,036	0.07	0.10	0.83			0.00	0.88	0.75

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=208,420) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score > 900).

Table B.2b: Item-Level Statistics—Reading and Writing, Grades 2–4

	Item #	Item Type	Max. Points	N-Count	% at 0/A	% at 1/B	% at 2/C	% at 3/D	% at 4	% Omit	P-value	Point Biserial
READING	1	MC	1	58,036	0.04	0.04	0.92			0.00	0.92	0.51
	2	MC	1	58,036	0.94	0.03	0.02			0.00	0.94	0.48
	3	MC	1	58,036	0.06	0.06	0.87			0.00	0.87	0.52
	7	MC	1	58,036	0.05	0.89	0.05			0.00	0.89	0.50
	8	MC	1	58,036	0.89	0.07	0.04			0.00	0.89	0.43
	9	MC	1	58,036	0.06	0.88	0.05			0.00	0.88	0.44
	10	MC	1	58,036	0.04	0.90	0.06			0.00	0.90	0.41
	14	MC	1	58,036	0.07	0.08	0.05	0.79		0.00	0.79	0.51
	15	MC	1	58,036	0.58	0.12	0.19	0.11		0.01	0.58	0.40
	16	MC	1	58,036	0.08	0.11	0.11	0.69		0.01	0.69	0.60
	17	MC	1	58,036	0.11	0.54	0.19	0.15		0.01	0.54	0.51
	18	MC	1	58,036	0.29	0.18	0.05	0.48		0.01	0.48	0.48
	19	MC	1	58,036	0.09	0.15	0.57	0.17		0.01	0.57	0.47
	20	MC	1	58,036	0.71	0.08	0.08	0.12		0.01	0.71	0.57
	21	MC	1	58,036	0.12	0.04	0.07	0.77		0.01	0.77	0.60
	22	MC	1	58,036	0.13	0.12	0.64	0.09		0.01	0.64	0.57
	23	MC	1	58,036	0.28	0.07	0.10	0.55		0.01	0.55	0.54
	24	MC	1	58,036	0.07	0.20	0.63	0.09		0.01	0.63	0.48
	25	MC	1	58,036	0.66	0.07	0.07	0.19		0.01	0.66	0.50
	26	MC	1	58,036	0.04	0.88	0.03	0.04		0.01	0.88	0.52
	27	MC	1	58,036	0.22	0.23	0.11	0.44		0.01	0.44	0.40
	28	MC	1	58,036	0.12	0.63	0.12	0.12		0.01	0.63	0.54
	29	MC	1	58,036	0.77	0.06	0.11	0.06		0.01	0.77	0.34
	30	MC	1	58,036	0.17	0.29	0.09	0.44		0.01	0.44	0.25
WRITING	1	MC	1	58,036	0.02	0.93	0.05	0.00		0.00	0.93	0.46
	2	MC	1	58,036	0.04	0.02	0.93	0.00		0.00	0.93	0.47
	3	MC	1	58,036	0.06	0.06	0.88	0.00		0.00	0.88	0.47
	4	MC	1	58,036	0.07	0.84	0.09	0.00		0.00	0.84	0.50
	5	MC	1	58,036	0.04	0.02	0.03	0.90		0.00	0.90	0.53
	6	MC	1	58,036	0.09	0.73	0.07	0.10		0.00	0.73	0.58
	7	MC	1	58,036	0.07	0.09	0.22	0.61		0.00	0.61	0.47
	8	MC	1	58,036	0.04	0.13	0.77	0.06		0.01	0.77	0.53
	9	MC	1	58,036	0.10	0.78	0.05	0.07		0.01	0.78	0.52
	10	MC	1	58,036	0.17	0.09	0.10	0.63		0.01	0.63	0.48
	11	MC	1	58,036	0.13	0.11	0.08	0.67		0.01	0.67	0.51
	12	MC	1	58,036	0.65	0.13	0.06	0.16		0.01	0.65	0.47
	13	CR	2	58,036	0.08	0.08	0.84			0.00	0.88	0.64
	14	CR	2	58,036	0.15	0.27	0.57			0.00	0.71	0.66
	15	CR	2	58,036	0.11	0.47	0.42			0.00	0.66	0.62
	16	CR	4	58,036	0.09	0.26	0.38	0.21	0.06	0.00	0.47	0.74

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=208,420) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score > 900).

Table B.3a: Item-Level Statistics—Listening and Speaking, Grades 5–6

	Item #	Item Type	Max. Points	N-Count	% at 0/A	% at 1/B	% at 2/C	% at 3/D	% at 4	% Omit	P-value	Point Biserial
LISTENING	1	MC	1	28,623	0.04	0.91	0.05			0.00	0.91	0.56
	2	MC	1	28,623	0.04	0.92	0.04			0.00	0.92	0.49
	3	MC	1	28,623	0.68	0.21	0.11			0.00	0.68	0.31
	4	MC	1	28,623	0.05	0.06	0.06	0.82		0.00	0.82	0.54
	5	MC	1	28,623	0.05	0.05	0.85	0.05		0.00	0.85	0.57
	6	MC	1	28,623	0.02	0.04	0.89	0.05		0.00	0.89	0.42
	7	MC	1	28,623	0.04	0.03	0.06	0.88		0.00	0.88	0.62
	8	MC	1	28,623	0.14	0.05	0.76	0.05		0.00	0.76	0.36
	9	MC	1	28,623	0.04	0.78	0.09	0.09		0.00	0.78	0.40
	10	MC	1	28,623	0.06	0.09	0.04	0.80		0.00	0.80	0.50
	11	MC	1	28,623	0.03	0.12	0.80	0.04		0.00	0.80	0.45
	12	MC	1	28,623	0.06	0.07	0.80	0.06		0.00	0.80	0.52
	13	MC	1	28,623	0.17	0.05	0.74	0.04		0.00	0.74	0.37
	14	MC	1	28,623	0.78	0.11	0.07	0.04		0.00	0.78	0.34
	15	MC	1	28,623	0.09	0.09	0.10	0.72		0.00	0.72	0.46
	16	MC	1	28,623	0.03	0.11	0.07	0.79		0.00	0.79	0.53
	17	MC	1	28,623	0.57	0.11	0.10	0.22		0.00	0.57	0.40
	18	MC	1	28,623	0.03	0.07	0.87	0.02		0.00	0.87	0.63
	19	MC	1	28,623	0.06	0.78	0.07	0.08		0.00	0.78	0.52
	20	MC	1	28,623	0.81	0.06	0.07	0.06		0.00	0.81	0.49
	21	MC	1	28,623	0.15	0.52	0.27	0.06		0.00	0.52	0.28
	22	MC	1	28,623	0.10	0.07	0.80	0.03		0.00	0.80	0.59
	23	MC	1	28,623	0.04	0.14	0.03	0.77		0.00	0.77	0.57
	24	MC	1	28,623	0.82	0.04	0.10	0.04		0.00	0.82	0.49
	25	MC	1	28,623	0.06	0.79	0.06	0.09		0.00	0.79	0.52
SPEAKING	1	CR	2	28,623	0.03	0.17	0.80			0.00	0.89	0.77
	2	CR	2	28,623	0.04	0.16	0.79			0.00	0.88	0.78
	3	CR	2	28,623	0.06	0.19	0.75			0.00	0.85	0.78
	4	CR	2	28,623	0.04	0.27	0.69			0.00	0.83	0.68
	5	SS	2	28,623	0.08	0.32	0.60			0.00	0.76	0.73
	6	SE	4	28,623	0.03	0.04	0.13	0.32	0.48	0.00	0.79	0.84
	7	CR	2	28,623	0.06	0.15	0.79			0.00	0.87	0.81
	8	SS	2	28,623	0.07	0.25	0.68			0.00	0.81	0.78
	9	CR	2	28,623	0.04	0.15	0.81			0.00	0.88	0.81
	10	SS	2	28,623	0.03	0.15	0.82			0.00	0.89	0.78
	11	CR	2	28,623	0.05	0.18	0.78			0.00	0.87	0.80
	12	CR	2	28,623	0.04	0.08	0.88			0.00	0.92	0.67
	13	SS	2	28,623	0.06	0.09	0.85			0.00	0.90	0.75
	14	CR	2	28,623	0.05	0.18	0.77			0.00	0.86	0.74
	15	CR	2	28,623	0.08	0.11	0.81			0.00	0.87	0.76
	16	CR	2	28,623	0.05	0.13	0.82			0.00	0.89	0.73

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=208,420) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score > 900).

Table B.3b: Item-Level Statistics—Reading and Writing, Grades 5–6

	Item #	Item Type	Max. Points	N-Count	% at 0/A	% at 1/B	% at 2/C	% at 3/D	% at 4	% Omit	P-value	Point Biserial
READING	1	MC	1	28,623	0.03	0.04	0.93			0.00	0.93	0.49
	2	MC	1	28,623	0.04	0.93	0.03			0.00	0.93	0.50
	3	MC	1	28,623	0.07	0.11	0.82			0.00	0.82	0.40
	4	MC	1	28,623	0.83	0.06	0.07	0.04		0.00	0.83	0.45
	5	MC	1	28,623	0.89	0.03	0.04	0.03		0.00	0.89	0.56
	6	MC	1	28,623	0.04	0.05	0.85	0.05		0.00	0.85	0.54
	7	MC	1	28,623	0.14	0.15	0.08	0.63		0.00	0.63	0.40
	8	MC	1	28,623	0.16	0.07	0.71	0.06		0.00	0.71	0.51
	9	MC	1	28,623	0.30	0.05	0.59	0.05		0.00	0.59	0.50
	10	MC	1	28,623	0.12	0.13	0.09	0.66		0.00	0.66	0.49
	11	MC	1	28,623	0.23	0.06	0.61	0.10		0.00	0.61	0.52
	12	MC	1	28,623	0.63	0.17	0.11	0.08		0.00	0.63	0.49
	13	MC	1	28,623	0.11	0.60	0.21	0.07		0.00	0.60	0.48
	14	MC	1	28,623	0.72	0.09	0.10	0.09		0.00	0.72	0.54
	15	MC	1	28,623	0.18	0.10	0.65	0.07		0.00	0.65	0.51
	16	MC	1	28,623	0.14	0.52	0.15	0.19		0.00	0.52	0.47
	17	MC	1	28,623	0.08	0.09	0.10	0.73		0.00	0.73	0.57
	18	MC	1	28,623	0.51	0.14	0.25	0.10		0.00	0.51	0.35
	19	MC	1	28,623	0.10	0.06	0.78	0.05		0.00	0.78	0.59
	20	MC	1	28,623	0.06	0.13	0.14	0.67		0.00	0.67	0.59
	21	MC	1	28,623	0.54	0.10	0.08	0.27		0.00	0.54	0.43
	22	MC	1	28,623	0.11	0.62	0.15	0.12		0.00	0.62	0.56
	23	MC	1	28,623	0.10	0.09	0.72	0.09		0.00	0.72	0.61
	24	MC	1	28,623	0.30	0.08	0.47	0.14		0.00	0.47	0.36
	25	MC	1	28,623	0.70	0.09	0.07	0.14		0.00	0.70	0.55
	26	MC	1	28,623	0.13	0.07	0.71	0.09		0.00	0.71	0.57
	27	MC	1	28,623	0.08	0.11	0.10	0.71		0.00	0.71	0.58
WRITING	1	MC	1	28,623	0.03	0.03	0.93	0.00		0.00	0.93	0.42
	2	MC	1	28,623	0.18	0.77	0.04	0.00		0.00	0.77	0.27
	3	MC	1	28,623	0.04	0.03	0.04	0.89		0.00	0.89	0.54
	4	MC	1	28,623	0.80	0.04	0.13	0.03		0.00	0.80	0.55
	5	MC	1	28,623	0.06	0.04	0.85	0.04		0.00	0.85	0.51
	6	MC	1	28,623	0.03	0.11	0.10	0.76		0.00	0.76	0.53
	7	MC	1	28,623	0.75	0.08	0.06	0.11		0.00	0.75	0.49
	8	MC	1	28,623	0.06	0.03	0.11	0.80		0.00	0.80	0.55
	9	MC	1	28,623	0.89	0.04	0.03	0.04		0.00	0.89	0.57
	10	MC	1	28,623	0.78	0.06	0.05	0.10		0.00	0.78	0.49
	11	MC	1	28,623	0.75	0.11	0.04	0.10		0.00	0.75	0.52
	12	MC	1	28,623	0.03	0.05	0.80	0.12		0.00	0.80	0.44
	13	MC	1	28,623	0.07	0.65	0.06	0.21		0.00	0.65	0.42
	14	MC	1	28,623	0.07	0.25	0.58	0.10		0.00	0.58	0.39
	15	MC	1	28,623	0.09	0.07	0.57	0.27		0.00	0.57	0.37
	16	CR	2	28,623	0.06	0.08	0.86			0.00	0.90	0.58
	17	CR	2	28,623	0.14	0.26	0.60			0.00	0.73	0.68
	18	CR	2	28,623	0.18	0.37	0.45			0.00	0.63	0.64
	19	CR	4	28,623	0.11	0.18	0.37	0.26	0.08	0.00	0.51	0.77

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=208,420) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score > 900).

Table B.4a: Item-Level Statistics—Listening and Speaking, Grades 7–8

	Item #	Item Type	Max. Points	N-Count	% at 0/A	% at 1/B	% at 2/C	% at 3/D	% at 4	% Omit	P-value	Point Biserial
LISTENING	1	MC	1	25,115	0.92	0.02	0.05			0.00	0.92	0.29
	2	MC	1	25,115	0.89	0.07	0.03			0.00	0.89	0.40
	3	MC	1	25,115	0.02	0.84	0.14			0.00	0.84	0.55
	4	MC	1	25,115	0.05	0.06	0.19	0.69		0.00	0.69	0.33
	5	MC	1	25,115	0.05	0.16	0.60	0.19		0.00	0.60	0.36
	6	MC	1	25,115	0.11	0.04	0.78	0.06		0.00	0.78	0.53
	7	MC	1	25,115	0.15	0.17	0.14	0.54		0.01	0.54	0.31
	8	MC	1	25,115	0.10	0.08	0.69	0.12		0.00	0.69	0.35
	9	MC	1	25,115	0.06	0.07	0.81	0.05		0.00	0.81	0.46
	10	MC	1	25,115	0.08	0.71	0.17	0.05		0.00	0.71	0.36
	11	MC	1	25,115	0.17	0.12	0.58	0.13		0.00	0.58	0.41
	12	MC	1	25,115	0.16	0.08	0.15	0.61		0.01	0.61	0.30
	13	MC	1	25,115	0.59	0.11	0.07	0.22		0.00	0.59	0.35
	14	MC	1	25,115	0.47	0.13	0.18	0.22		0.00	0.47	0.38
	15	MC	1	25,115	0.04	0.11	0.77	0.07		0.01	0.77	0.48
	16	MC	1	25,115	0.06	0.02	0.86	0.05		0.00	0.86	0.45
	17	MC	1	25,115	0.04	0.04	0.09	0.83		0.00	0.83	0.42
	18	MC	1	25,115	0.02	0.07	0.14	0.77		0.00	0.77	0.40
	19	MC	1	25,115	0.80	0.11	0.05	0.04		0.00	0.80	0.44
	20	MC	1	25,115	0.82	0.04	0.03	0.10		0.00	0.82	0.29
	21	MC	1	25,115	0.08	0.14	0.12	0.65		0.00	0.65	0.42
	22	MC	1	25,115	0.08	0.09	0.62	0.20		0.00	0.62	0.46
	23	MC	1	25,115	0.07	0.05	0.04	0.83		0.00	0.83	0.37
	24	MC	1	25,115	0.19	0.08	0.06	0.66		0.00	0.66	0.48
	25	MC	1	25,115	0.08	0.06	0.79	0.06		0.01	0.79	0.53
SPEAKING	1	CR	2	25,115	0.05	0.22	0.73			0.00	0.84	0.77
	2	CR	2	25,115	0.06	0.25	0.69			0.00	0.81	0.78
	3	SS	2	25,115	0.06	0.25	0.69			0.00	0.81	0.76
	4	CR	2	25,115	0.08	0.22	0.71			0.00	0.82	0.79
	5	CR	2	25,115	0.12	0.25	0.63			0.00	0.76	0.78
	6	CR	4	25,115	0.06	0.06	0.16	0.28	0.44	0.00	0.74	0.87
	7	SS	2	25,115	0.06	0.25	0.70			0.00	0.82	0.77
	8	CR	2	25,115	0.05	0.19	0.76			0.00	0.85	0.78
	9	SS	2	25,115	0.06	0.20	0.73			0.00	0.83	0.79
	10	CR	2	25,115	0.06	0.22	0.72			0.00	0.83	0.79
	11	SS	2	25,115	0.08	0.30	0.62			0.00	0.77	0.77
	12	CR	2	25,115	0.09	0.13	0.78			0.00	0.85	0.77
	13	CR	2	25,115	0.06	0.18	0.75			0.00	0.85	0.75
	14	SS	2	25,115	0.09	0.18	0.72			0.00	0.82	0.80
	15	CR	2	25,115	0.07	0.12	0.81			0.00	0.87	0.76
	16	CR	2	25,115	0.09	0.17	0.74			0.00	0.83	0.80

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=208,420) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score > 900).

Table B.4b: Item-Level Statistics—Reading and Writing, Grades 7–8

	Item #	Item Type	Max. Points	N-Count	% at 0/A	% at 1/B	% at 2/C	% at 3/D	% at 4	% Omit	P-value	Point Biserial
READING	1	MC	1	25,115	0.02	0.92	0.06			0.00	0.92	0.33
	2	MC	1	25,115	0.04	0.10	0.86			0.00	0.86	0.40
	3	MC	1	25,115	0.91	0.06	0.03			0.00	0.91	0.36
	4	MC	1	25,115	0.15	0.11	0.57	0.16		0.00	0.57	0.46
	5	MC	1	25,115	0.11	0.77	0.05	0.07		0.00	0.77	0.54
	6	MC	1	25,115	0.11	0.05	0.06	0.77		0.00	0.77	0.59
	7	MC	1	25,115	0.08	0.07	0.77	0.08		0.00	0.77	0.56
	8	MC	1	25,115	0.09	0.14	0.64	0.14		0.00	0.64	0.55
	9	MC	1	25,115	0.10	0.10	0.09	0.71		0.00	0.71	0.54
	10	MC	1	25,115	0.54	0.13	0.14	0.19		0.00	0.54	0.41
	11	MC	1	25,115	0.17	0.12	0.63	0.08		0.00	0.63	0.48
	12	MC	1	25,115	0.06	0.62	0.12	0.20		0.00	0.62	0.43
	13	MC	1	25,115	0.06	0.13	0.10	0.70		0.00	0.70	0.48
	14	MC	1	25,115	0.06	0.10	0.65	0.18		0.00	0.65	0.51
	15	MC	1	25,115	0.11	0.09	0.55	0.25		0.00	0.55	0.44
	16	MC	1	25,115	0.42	0.07	0.08	0.43		0.00	0.43	0.33
	17	MC	1	25,115	0.52	0.28	0.13	0.07		0.00	0.52	0.35
	18	MC	1	25,115	0.62	0.08	0.16	0.14		0.00	0.62	0.49
	19	MC	1	25,115	0.55	0.14	0.16	0.14		0.00	0.55	0.49
	20	MC	1	25,115	0.11	0.57	0.24	0.07		0.00	0.57	0.38
	21	MC	1	25,115	0.14	0.53	0.25	0.08		0.00	0.53	0.50
	22	MC	1	25,115	0.20	0.09	0.25	0.46		0.00	0.46	0.37
	23	NS	1	25,115	0.65	0.17	0.10	0.08		0.01	0.65	0.50
	24	MC	1	25,115	0.09	0.07	0.58	0.26		0.01	0.58	0.41
	25	MC	1	25,115	0.13	0.20	0.10	0.57		0.01	0.57	0.50
	26	MC	1	25,115	0.52	0.19	0.09	0.19		0.01	0.52	0.37
	27	MC	1	25,115	0.15	0.45	0.21	0.18		0.01	0.45	0.30
WRITING	1	MC	1	25,115	0.03	0.91	0.06	0.00		0.00	0.91	0.43
	2	MC	1	25,115	0.89	0.06	0.04	0.00		0.00	0.89	0.34
	3	MC	1	25,115	0.04	0.88	0.05	0.03		0.00	0.88	0.53
	4	MC	1	25,115	0.07	0.03	0.04	0.86		0.00	0.86	0.45
	5	MC	1	25,115	0.79	0.07	0.10	0.04		0.00	0.79	0.52
	6	MC	1	25,115	0.13	0.04	0.05	0.77		0.00	0.77	0.53
	7	MC	1	25,115	0.54	0.08	0.05	0.33		0.00	0.54	0.38
	8	MC	1	25,115	0.05	0.05	0.77	0.12		0.00	0.77	0.52
	9	MC	1	25,115	0.07	0.11	0.04	0.77		0.00	0.77	0.60
	10	MC	1	25,115	0.65	0.10	0.11	0.14		0.00	0.65	0.49
	11	MC	1	25,115	0.47	0.22	0.10	0.20		0.00	0.47	0.26
	12	MC	1	25,115	0.05	0.15	0.75	0.05		0.00	0.75	0.36
	13	MC	1	25,115	0.66	0.05	0.13	0.15		0.00	0.66	0.45
	14	MC	1	25,115	0.17	0.05	0.70	0.07		0.00	0.70	0.42
	15	MC	1	25,115	0.02	0.16	0.23	0.59		0.00	0.59	0.41
	16	CR	2	25,115	0.12	0.21	0.67			0.00	0.78	0.69
	17	CR	2	25,115	0.19	0.35	0.46			0.00	0.63	0.61
	18	CR	2	25,115	0.15	0.27	0.58			0.00	0.72	0.67
	19	ER	4	25,115	0.12	0.19	0.34	0.25	0.09	0.00	0.50	0.77

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=208,420) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score > 900).

Table B.5a: Item-Level Statistics—Listening and Speaking, Grades 9–12

	Item #	Item Type	Max. Points	N-Count	% at 0/A	% at 1/B	% at 2/C	% at 3/D	% at 4	% Omit	P-value	Point Biserial
LISTENING	1	MC	1	44,704	0.85	0.11	0.04			0.01	0.85	0.41
	2	MC	1	44,704	0.76	0.12	0.11			0.01	0.76	0.51
	3	MC	1	44,704	0.03	0.89	0.07			0.01	0.89	0.51
	4	MC	1	44,704	0.10	0.07	0.78	0.05		0.01	0.78	0.46
	5	MC	1	44,704	0.61	0.11	0.10	0.17		0.01	0.61	0.41
	6	MC	1	44,704	0.15	0.15	0.64	0.06		0.01	0.64	0.48
	7	MC	1	44,704	0.08	0.20	0.22	0.50		0.01	0.50	0.27
	8	MC	1	44,704	0.70	0.10	0.16	0.04		0.01	0.70	0.49
	9	MC	1	44,704	0.04	0.12	0.08	0.75		0.01	0.75	0.48
	10	MC	1	44,704	0.05	0.79	0.04	0.11		0.01	0.79	0.34
	11	MC	1	44,704	0.80	0.08	0.06	0.05		0.01	0.80	0.43
	12	MC	1	44,704	0.08	0.51	0.22	0.18		0.01	0.51	0.40
	13	MC	1	44,704	0.21	0.49	0.23	0.06		0.01	0.49	0.44
	14	MC	1	44,704	0.11	0.74	0.07	0.07		0.01	0.74	0.44
	15	MC	1	44,704	0.09	0.17	0.15	0.59		0.01	0.59	0.49
	16	MC	1	44,704	0.50	0.21	0.17	0.11		0.01	0.50	0.43
	17	MC	1	44,704	0.17	0.08	0.64	0.11		0.01	0.64	0.46
	18	MC	1	44,704	0.11	0.10	0.28	0.50		0.01	0.50	0.53
	19	MC	1	44,704	0.56	0.19	0.15	0.09		0.01	0.56	0.37
	20	MC	1	44,704	0.05	0.26	0.09	0.59		0.01	0.59	0.40
	21	MC	1	44,704	0.80	0.12	0.03	0.04		0.01	0.80	0.51
	22	MC	1	44,704	0.10	0.13	0.12	0.63		0.01	0.63	0.53
	23	MC	1	44,704	0.12	0.65	0.15	0.08		0.01	0.65	0.47
	24	MC	1	44,704	0.10	0.10	0.73	0.06		0.01	0.73	0.46
	25	MC	1	44,704	0.70	0.10	0.10	0.09		0.01	0.70	0.43
SPEAKING	1	CR	2	44,704	0.09	0.30	0.61			0.00	0.76	0.70
	2	CR	2	44,704	0.07	0.26	0.68			0.00	0.81	0.74
	3	SS	2	44,704	0.07	0.19	0.73			0.00	0.83	0.75
	4	CR	2	44,704	0.07	0.30	0.63			0.00	0.78	0.74
	5	SS	2	44,704	0.09	0.17	0.74			0.00	0.83	0.73
	6	CR	4	44,704	0.04	0.06	0.18	0.33	0.39	0.00	0.75	0.84
	7	CR	2	44,704	0.07	0.29	0.64			0.00	0.79	0.77
	8	CR	2	44,704	0.06	0.23	0.72			0.00	0.83	0.74
	9	SS	2	44,704	0.05	0.26	0.68			0.00	0.82	0.74
	10	CR	2	44,704	0.05	0.31	0.64			0.00	0.79	0.76
	11	CR	2	44,704	0.05	0.23	0.72			0.00	0.84	0.75
	12	CR	2	44,704	0.17	0.17	0.66			0.00	0.75	0.78
	13	CR	2	44,704	0.07	0.16	0.78			0.00	0.85	0.72
	14	SS	2	44,704	0.11	0.17	0.73			0.00	0.81	0.74
	15	CR	2	44,704	0.27	0.17	0.57			0.00	0.65	0.76
	16	CR	2	44,704	0.07	0.21	0.72			0.00	0.82	0.75

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=208,420) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score > 900).

Table B.5b: Item-Level Statistics—Reading and Writing, Grades 9–12

	Item #	Item Type	Max. Points	N-Count	% at 0/A	% at 1/B	% at 2/C	% at 3/D	% at 4	% Omit	P-value	Point Biserial
READING	1	MC	1	44,704	0.03	0.09	0.88			0.00	0.88	0.41
	2	MC	1	44,704	0.11	0.86	0.02			0.00	0.86	0.44
	3	MC	1	44,704	0.11	0.81	0.08			0.01	0.81	0.40
	4	MC	1	44,704	0.20	0.65	0.06	0.07		0.01	0.65	0.40
	5	MC	1	44,704	0.07	0.07	0.78	0.07		0.01	0.78	0.52
	6	MC	1	44,704	0.08	0.78	0.07	0.05		0.01	0.78	0.55
	7	MC	1	44,704	0.30	0.14	0.42	0.12		0.01	0.42	0.35
	8	MC	1	44,704	0.47	0.20	0.18	0.15		0.01	0.47	0.35
	9	MC	1	44,704	0.14	0.49	0.09	0.27		0.01	0.49	0.39
	10	MC	1	44,704	0.69	0.08	0.11	0.11		0.01	0.69	0.44
	11	MC	1	44,704	0.17	0.07	0.34	0.41		0.01	0.41	0.34
	12	MC	1	44,704	0.18	0.09	0.24	0.47		0.01	0.47	0.40
	13	MC	1	44,704	0.64	0.15	0.11	0.09		0.01	0.64	0.56
	14	MC	1	44,704	0.76	0.10	0.05	0.08		0.01	0.76	0.50
	15	MC	1	44,704	0.25	0.07	0.07	0.60		0.01	0.60	0.57
	16	MC	1	44,704	0.07	0.41	0.13	0.38		0.01	0.38	0.37
	17	MC	1	44,704	0.24	0.59	0.09	0.07		0.01	0.59	0.40
	18	MC	1	44,704	0.04	0.12	0.07	0.77		0.01	0.77	0.46
	19	MC	1	44,704	0.43	0.23	0.15	0.17		0.01	0.43	0.43
	20	MC	1	44,704	0.16	0.23	0.13	0.47		0.01	0.47	0.42
	21	MC	1	44,704	0.12	0.21	0.19	0.47		0.01	0.47	0.44
	22	MC	1	44,704	0.56	0.20	0.11	0.11		0.01	0.56	0.35
	23	MC	1	44,704	0.24	0.08	0.50	0.17		0.01	0.50	0.32
	24	MC	1	44,704	0.09	0.57	0.22	0.10		0.01	0.57	0.35
	25	MC	1	44,704	0.55	0.18	0.17	0.09		0.01	0.55	0.48
	26	MC	1	44,704	0.08	0.15	0.64	0.11		0.01	0.64	0.49
	27	MC	1	44,704	0.11	0.09	0.09	0.69		0.01	0.69	0.54
WRITING	1	MC	1	44,704	0.07	0.90	0.03	0.00		0.01	0.90	0.38
	2	MC	1	44,704	0.07	0.86	0.07	0.00		0.01	0.86	0.43
	3	MC	1	44,704	0.75	0.08	0.06	0.10		0.01	0.75	0.53
	4	MC	1	44,704	0.18	0.16	0.19	0.46		0.01	0.46	0.44
	5	MC	1	44,704	0.15	0.63	0.13	0.08		0.01	0.63	0.50
	6	MC	1	44,704	0.07	0.12	0.54	0.26		0.01	0.54	0.44
	7	MC	1	44,704	0.65	0.13	0.11	0.11		0.01	0.65	0.48
	8	MC	1	44,704	0.02	0.23	0.71	0.03		0.01	0.71	0.45
	9	MC	1	44,704	0.52	0.21	0.16	0.09		0.01	0.52	0.52
	10	MC	1	44,704	0.13	0.03	0.78	0.06		0.01	0.78	0.49
	11	MC	1	44,704	0.04	0.15	0.58	0.22		0.01	0.58	0.40
	12	MC	1	44,704	0.13	0.08	0.05	0.73		0.01	0.73	0.58
	13	MC	1	44,704	0.24	0.18	0.44	0.14		0.01	0.44	0.39
	14	MC	1	44,704	0.05	0.10	0.68	0.16		0.01	0.68	0.58
	15	MC	1	44,704	0.11	0.10	0.15	0.63		0.01	0.63	0.53
	16	CR	2	44,704	0.14	0.20	0.66			0.00	0.76	0.66
	17	CR	2	44,704	0.16	0.15	0.69			0.00	0.77	0.69
	18	CR	2	44,704	0.19	0.28	0.53			0.00	0.67	0.67
	19	CR	4	44,704	0.15	0.14	0.28	0.28	0.16	0.00	0.54	0.77

Note: The data file used in Tables B.1–B.5 were based on the final research file (N=208,420) after cleaning (removal of duplicate cases, off-level cases, and all records with Strand Raw Score > 900).

APPENDIX C: RAW SCORE-TO-SCALE SCORE CONVERSIONS

Table C.1a: Raw Score-to-Scale Score Conversions—Listening/Speaking, Grades K–1

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
0	-7.94	2.01	322	70
1	-6.53	1.02	371	36
2	-5.81	0.73	397	26
3	-5.37	0.60	412	21
4	-5.05	0.53	423	19
5	-4.79	0.48	432	17
6	-4.58	0.45	440	16
7	-4.39	0.42	446	15
8	-4.22	0.40	452	14
9	-4.07	0.38	458	13
10	-3.93	0.37	462	13
11	-3.80	0.36	467	12
12	-3.68	0.34	471	12
13	-3.56	0.34	475	12
14	-3.45	0.33	479	11
15	-3.35	0.32	483	11
16	-3.25	0.32	486	11
17	-3.15	0.31	490	11
18	-3.05	0.31	493	11
19	-2.96	0.30	496	11
20	-2.87	0.30	500	10
21	-2.78	0.30	503	10
22	-2.69	0.29	506	10
23	-2.61	0.29	509	10
24	-2.52	0.29	512	10
25	-2.44	0.29	515	10
26	-2.35	0.29	518	10
27	-2.27	0.29	521	10
28	-2.19	0.29	524	10
29	-2.10	0.29	526	10
30	-2.02	0.29	529	10
31	-1.93	0.29	532	10
32	-1.85	0.29	535	10
33	-1.76	0.29	538	10
34	-1.68	0.30	541	10
35	-1.59	0.30	544	10
36	-1.50	0.30	548	11
37	-1.41	0.30	551	11
38	-1.31	0.31	554	11
39	-1.22	0.31	557	11
40	-1.12	0.32	561	11
41	-1.02	0.32	564	11
42	-0.91	0.33	568	12

C.1a (continued): Raw Score-to-Scale Score Conversions—Listening/Speaking, Grades K–1

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
43	-0.80	0.34	572	12
44	-0.68	0.35	576	12
45	-0.56	0.35	580	12
46	-0.43	0.37	585	13
47	-0.29	0.38	590	13
48	-0.14	0.39	595	14
49	0.02	0.41	601	14
50	0.19	0.43	607	15
51	0.39	0.46	614	16
52	0.61	0.49	621	17
53	0.87	0.53	630	18
54	1.18	0.58	641	20
55	1.55	0.66	654	23
56	2.06	0.78	672	27
57	2.87	1.06	700	37
58	4.35	2.03	752	71

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

Table C.1b: Raw Score-to-Scale Score Conversions—Reading/Writing, Grades K–1

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
0	-8.20	2.02	313	71
1	-6.75	1.04	364	36
2	-5.98	0.76	391	27
3	-5.49	0.64	408	22
4	-5.13	0.57	420	20
5	-4.84	0.52	431	18
6	-4.59	0.48	439	17
7	-4.37	0.45	447	16
8	-4.18	0.43	454	15
9	-4.00	0.41	460	14
10	-3.84	0.40	466	14
11	-3.69	0.39	471	13
12	-3.54	0.38	476	13
13	-3.40	0.37	481	13
14	-3.27	0.36	486	13
15	-3.14	0.36	490	13
16	-3.01	0.36	495	13
17	-2.88	0.36	499	12
18	-2.76	0.36	503	12
19	-2.63	0.36	508	12
20	-2.50	0.36	512	13
21	-2.37	0.36	517	13
22	-2.24	0.36	521	13
23	-2.11	0.37	526	13
24	-1.98	0.37	531	13
25	-1.84	0.38	536	13
26	-1.69	0.38	541	13
27	-1.54	0.39	546	14
28	-1.39	0.40	552	14
29	-1.22	0.41	557	14
30	-1.05	0.43	563	15
31	-0.86	0.44	570	16
32	-0.65	0.46	577	16
33	-0.42	0.49	585	17
34	-0.16	0.53	594	19
35	0.14	0.58	605	20
36	0.52	0.65	618	23
37	1.03	0.78	636	27
38	1.83	1.06	664	37
39	3.31	2.03	716	71

Table C.2a: Raw Score-to-Scale Score Conversions—Listening/Speaking, Grades 2–4

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
0	-6.07	2.01	387	70
1	-4.66	1.02	437	36
2	-3.94	0.73	462	26
3	-3.50	0.60	477	21
4	-3.18	0.53	489	19
5	-2.93	0.48	498	17
6	-2.71	0.44	505	16
7	-2.53	0.42	511	15
8	-2.36	0.40	517	14
9	-2.21	0.38	522	13
10	-2.08	0.36	527	13
11	-1.95	0.35	532	12
12	-1.83	0.34	536	12
13	-1.72	0.33	540	12
14	-1.61	0.32	544	11
15	-1.51	0.32	547	11
16	-1.41	0.31	551	11
17	-1.31	0.31	554	11
18	-1.22	0.30	557	11
19	-1.13	0.30	560	10
20	-1.04	0.30	563	10
21	-0.96	0.29	566	10
22	-0.87	0.29	569	10
23	-0.79	0.29	572	10
24	-0.71	0.29	575	10
25	-0.62	0.29	578	10
26	-0.54	0.29	581	10
27	-0.46	0.28	584	10
28	-0.38	0.28	587	10
29	-0.30	0.28	590	10
30	-0.22	0.29	592	10
31	-0.14	0.29	595	10
32	-0.05	0.29	598	10
33	0.03	0.29	601	10
34	0.11	0.29	604	10
35	0.20	0.29	607	10
36	0.29	0.30	610	10
37	0.37	0.30	613	10
38	0.47	0.30	616	11
39	0.56	0.31	620	11
40	0.66	0.31	623	11
41	0.75	0.32	626	11
42	0.86	0.32	630	11
43	0.96	0.33	634	12
44	1.07	0.34	638	12
45	1.19	0.35	642	12

Table C.2a (continued): Raw Score-to-Scale Score Conversions—Listening/Speaking, Grades 2–4

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
46	1.31	0.36	646	12
47	1.44	0.37	651	13
48	1.58	0.38	655	13
49	1.73	0.39	661	14
50	1.90	0.41	666	14
51	2.07	0.43	673	15
52	2.27	0.46	680	16
53	2.50	0.50	688	17
54	2.77	0.55	697	19
55	3.11	0.62	709	22
56	3.56	0.74	725	26
57	4.30	1.02	751	36
58	5.73	2.01	800	70

Table C.2b: Raw Score-to-Scale Score Conversions—Reading/Writing, Grades 2–4

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
0	-6.30	2.01	380	70
1	-4.87	1.02	430	36
2	-4.13	0.74	456	26
3	-3.67	0.62	472	22
4	-3.33	0.55	483	19
5	-3.06	0.50	493	18
6	-2.82	0.47	501	16
7	-2.62	0.44	508	15
8	-2.43	0.42	515	15
9	-2.26	0.40	521	14
10	-2.10	0.39	526	14
11	-1.96	0.38	532	13
12	-1.82	0.37	536	13
13	-1.68	0.36	541	13
14	-1.55	0.36	546	13
15	-1.42	0.35	550	12
16	-1.30	0.35	554	12
17	-1.18	0.35	559	12
18	-1.06	0.34	563	12
19	-0.95	0.34	567	12
20	-0.83	0.34	571	12
21	-0.71	0.34	575	12
22	-0.60	0.34	579	12
23	-0.48	0.34	583	12
24	-0.37	0.34	587	12
25	-0.25	0.34	591	12
26	-0.13	0.34	595	12
27	-0.01	0.35	600	12
28	0.11	0.35	604	12
29	0.23	0.35	608	12
30	0.36	0.36	613	13
31	0.49	0.36	617	13
32	0.62	0.37	622	13
33	0.76	0.38	627	13
34	0.90	0.38	632	13
35	1.06	0.40	637	14
36	1.22	0.41	643	14
37	1.39	0.42	649	15
38	1.58	0.44	655	15
39	1.78	0.46	662	16
40	2.01	0.49	670	17
41	2.27	0.53	680	19
42	2.59	0.59	690	21
43	2.97	0.66	704	23
44	3.50	0.79	722	28
45	4.34	1.08	752	38
46	5.86	2.05	805	72

Table C.3a: Raw Score-to-Scale Score Conversions—Listening/Speaking, Grades 5–6

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
0	-5.85	2.01	395	70
1	-4.44	1.01	445	35
2	-3.72	0.73	470	25
3	-3.29	0.60	485	21
4	-2.97	0.53	496	18
5	-2.72	0.48	505	17
6	-2.52	0.44	512	15
7	-2.33	0.41	518	14
8	-2.17	0.39	524	14
9	-2.03	0.37	529	13
10	-1.90	0.36	534	12
11	-1.78	0.34	538	12
12	-1.66	0.33	542	12
13	-1.55	0.32	546	11
14	-1.45	0.32	549	11
15	-1.36	0.31	553	11
16	-1.26	0.30	556	11
17	-1.17	0.30	559	10
18	-1.08	0.29	562	10
19	-1.00	0.29	565	10
20	-0.92	0.29	568	10
21	-0.84	0.28	571	10
22	-0.76	0.28	574	10
23	-0.68	0.28	576	10
24	-0.60	0.28	579	10
25	-0.53	0.28	582	10
26	-0.45	0.28	584	10
27	-0.37	0.27	587	10
28	-0.30	0.27	590	10
29	-0.22	0.27	592	10
30	-0.15	0.27	595	10
31	-0.07	0.28	597	10
32	0.00	0.28	600	10
33	0.08	0.28	603	10
34	0.16	0.28	606	10
35	0.24	0.28	608	10
36	0.32	0.28	611	10
37	0.40	0.29	614	10
38	0.48	0.29	617	10
39	0.57	0.29	620	10
40	0.65	0.30	623	10
41	0.74	0.30	626	11
42	0.83	0.31	629	11

Table C.3a (continued): Raw Score-to-Scale Score Conversions—Listening/Speaking, Grades 5–6

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
43	0.93	0.31	633	11
44	1.03	0.32	636	11
45	1.13	0.33	640	11
46	1.24	0.33	643	12
47	1.36	0.34	647	12
48	1.48	0.35	652	12
49	1.61	0.37	656	13
50	1.75	0.38	661	13
51	1.90	0.40	667	14
52	2.07	0.42	672	15
53	2.26	0.45	679	16
54	2.48	0.49	687	17
55	2.74	0.54	696	19
56	3.06	0.61	707	21
57	3.51	0.73	723	26
58	4.24	1.02	748	36
59	5.65	2.01	798	70

Table C.3b: Raw Score-to-Scale Score Conversions—Reading/Writing, Grades 5–6

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
0	-4.77	2.01	433	70
1	-3.35	1.02	483	36
2	-2.62	0.74	508	26
3	-2.17	0.61	524	21
4	-1.85	0.54	535	19
5	-1.58	0.49	545	17
6	-1.36	0.45	552	16
7	-1.17	0.43	559	15
8	-1.00	0.40	565	14
9	-0.84	0.39	571	14
10	-0.70	0.37	576	13
11	-0.56	0.36	580	13
12	-0.44	0.35	585	12
13	-0.32	0.34	589	12
14	-0.20	0.34	593	12
15	-0.09	0.33	597	12
16	0.01	0.32	600	11
17	0.12	0.32	604	11
18	0.22	0.32	608	11
19	0.32	0.31	611	11
20	0.41	0.31	614	11
21	0.51	0.31	618	11
22	0.60	0.31	621	11
23	0.70	0.30	624	11
24	0.79	0.30	628	11
25	0.88	0.30	631	11
26	0.97	0.30	634	11
27	1.07	0.30	637	11
28	1.16	0.30	641	11
29	1.25	0.31	644	11
30	1.35	0.31	647	11
31	1.44	0.31	650	11
32	1.54	0.31	654	11
33	1.64	0.32	657	11
34	1.74	0.32	661	11
35	1.84	0.32	664	11
36	1.95	0.33	668	12
37	2.06	0.33	672	12
38	2.17	0.34	676	12
39	2.29	0.35	680	12
40	2.42	0.36	685	13
41	2.55	0.37	689	13
42	2.69	0.38	694	13
43	2.84	0.40	699	14
44	3.01	0.42	705	15
45	3.19	0.44	712	15
46	3.40	0.47	719	16
47	3.63	0.51	727	18
48	3.92	0.56	737	20
49	4.27	0.63	749	22
50	4.75	0.76	766	27
51	5.53	1.04	793	37
52	6.99	2.03	844	71

Table C.4a: Raw Score-to-Scale Score Conversions—Listening/Speaking, Grades 7–8

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
0	-4.77	2.01	433	70
1	-3.37	1.01	482	35
2	-2.65	0.73	507	25
3	-2.22	0.60	522	21
4	-1.90	0.53	533	18
5	-1.65	0.48	542	17
6	-1.45	0.44	549	15
7	-1.27	0.41	556	14
8	-1.11	0.39	561	14
9	-0.96	0.37	566	13
10	-0.83	0.36	571	12
11	-0.71	0.34	575	12
12	-0.60	0.33	579	12
13	-0.49	0.32	583	11
14	-0.39	0.31	586	11
15	-0.29	0.31	590	11
16	-0.20	0.30	593	11
17	-0.11	0.30	596	10
18	-0.02	0.29	599	10
19	0.06	0.29	602	10
20	0.14	0.28	605	10
21	0.22	0.28	608	10
22	0.30	0.28	610	10
23	0.38	0.28	613	10
24	0.45	0.27	616	10
25	0.53	0.27	618	10
26	0.60	0.27	621	10
27	0.68	0.27	624	10
28	0.75	0.27	626	10
29	0.82	0.27	629	10
30	0.90	0.27	631	10
31	0.97	0.27	634	10
32	1.05	0.27	637	10
33	1.12	0.28	639	10
34	1.20	0.28	642	10
35	1.28	0.28	645	10
36	1.35	0.28	647	10
37	1.44	0.28	650	10
38	1.52	0.29	653	10
39	1.60	0.29	656	10
40	1.69	0.30	659	10
41	1.77	0.30	662	10
42	1.87	0.30	665	11
43	1.96	0.31	669	11
44	2.06	0.32	672	11
45	2.16	0.32	676	11

Table C.4a (continued): Raw Score-to-Scale Score Conversions—Listening/Speaking, Grades 7–8

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
46	2.27	0.33	679	12
47	2.38	0.34	683	12
48	2.50	0.35	688	12
49	2.63	0.36	692	13
50	2.77	0.38	697	13
51	2.92	0.40	702	14
52	3.09	0.42	708	15
53	3.27	0.45	715	16
54	3.49	0.48	722	17
55	3.75	0.53	731	19
56	4.07	0.61	742	21
57	4.50	0.73	758	26
58	5.23	1.02	783	36
59	6.64	2.01	832	70

Table C.4b: Raw Score-to-Scale Score Conversions—Reading/Writing, Grades 7–8

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
0	-4.82	2.01	431	70
1	-3.40	1.02	481	36
2	-2.66	0.74	507	26
3	-2.21	0.61	523	22
4	-1.88	0.54	534	19
5	-1.61	0.49	544	17
6	-1.39	0.46	552	16
7	-1.19	0.43	558	15
8	-1.01	0.41	565	14
9	-0.85	0.39	570	14
10	-0.71	0.38	575	13
11	-0.57	0.36	580	13
12	-0.44	0.35	584	12
13	-0.32	0.34	589	12
14	-0.21	0.34	593	12
15	-0.10	0.33	597	12
16	0.01	0.32	600	11
17	0.11	0.32	604	11
18	0.21	0.31	607	11
19	0.31	0.31	611	11
20	0.40	0.31	614	11
21	0.50	0.30	617	11
22	0.59	0.30	621	11
23	0.68	0.30	624	11
24	0.77	0.30	627	10
25	0.86	0.30	630	10
26	0.95	0.30	633	10
27	1.04	0.30	636	10
28	1.13	0.30	639	10
29	1.22	0.30	643	11
30	1.31	0.30	646	11
31	1.40	0.30	649	11
32	1.49	0.31	652	11
33	1.59	0.31	656	11
34	1.68	0.31	659	11
35	1.78	0.32	662	11
36	1.88	0.32	666	11
37	1.99	0.33	670	11
38	2.10	0.33	673	12
39	2.21	0.34	677	12
40	2.33	0.35	682	12
41	2.46	0.36	686	13
42	2.60	0.37	691	13
43	2.74	0.39	696	14
44	2.90	0.41	701	14
45	3.07	0.43	708	15
46	3.27	0.46	714	16
47	3.49	0.49	722	17
48	3.75	0.54	731	19
49	4.08	0.61	743	21
50	4.53	0.74	759	26
51	5.27	1.02	785	36
52	6.69	2.01	834	70

Table C.5a: Raw Score-to-Scale Score Conversions—Listening/Speaking, Grades 9–12

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
0	-4.37	2.01	447	70
1	-2.96	1.01	497	36
2	-2.23	0.73	522	25
3	-1.80	0.60	537	21
4	-1.48	0.53	548	19
5	-1.23	0.48	557	17
6	-1.02	0.44	564	16
7	-0.83	0.42	571	15
8	-0.67	0.39	577	14
9	-0.52	0.38	582	13
10	-0.38	0.36	587	13
11	-0.26	0.35	591	12
12	-0.14	0.34	595	12
13	-0.03	0.33	599	12
14	0.08	0.32	603	11
15	0.18	0.32	606	11
16	0.28	0.31	610	11
17	0.38	0.31	613	11
18	0.47	0.30	616	11
19	0.56	0.30	620	10
20	0.65	0.29	623	10
21	0.73	0.29	626	10
22	0.82	0.29	629	10
23	0.90	0.29	632	10
24	0.98	0.29	634	10
25	1.06	0.28	637	10
26	1.15	0.28	640	10
27	1.23	0.28	643	10
28	1.31	0.28	646	10
29	1.38	0.28	648	10
30	1.46	0.28	651	10
31	1.54	0.28	654	10
32	1.62	0.28	657	10
33	1.70	0.28	660	10
34	1.78	0.28	662	10
35	1.87	0.29	665	10
36	1.95	0.29	668	10
37	2.03	0.29	671	10
38	2.12	0.29	674	10
39	2.20	0.30	677	10
40	2.29	0.30	680	10
41	2.38	0.30	683	11
42	2.47	0.31	687	11
43	2.57	0.31	690	11
44	2.67	0.32	693	11
45	2.77	0.33	697	11
46	2.88	0.33	701	12

Table C.5a (continued): Raw Score-to-Scale Score Conversions—Listening/Speaking, Grades 9–12

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
47	3.00	0.34	705	12
48	3.12	0.35	709	12
49	3.25	0.37	714	13
50	3.38	0.38	718	13
51	3.53	0.40	724	14
52	3.70	0.42	730	15
53	3.89	0.45	736	16
54	4.10	0.48	744	17
55	4.36	0.53	753	19
56	4.68	0.61	764	21
57	5.12	0.73	779	26
58	5.84	1.02	805	36
59	7.26	2.01	854	70

Table C.5b: Raw Score-to-Scale Score Conversions—Reading/Writing, Grades 9–12

Raw Score	Measure	SE (RS)	Scale Score	SE (SS)
0	-3.64	2.01	473	70
1	-2.22	1.02	522	36
2	-1.49	0.73	548	26
3	-1.05	0.61	563	21
4	-0.73	0.53	575	19
5	-0.47	0.49	584	17
6	-0.25	0.45	591	16
7	-0.06	0.42	598	15
8	0.10	0.40	604	14
9	0.26	0.38	609	13
10	0.39	0.37	614	13
11	0.52	0.35	618	12
12	0.64	0.34	623	12
13	0.76	0.33	627	12
14	0.87	0.32	630	11
15	0.97	0.32	634	11
16	1.07	0.31	637	11
17	1.16	0.31	641	11
18	1.26	0.30	644	11
19	1.35	0.30	647	11
20	1.44	0.30	650	10
21	1.53	0.30	653	10
22	1.61	0.29	656	10
23	1.70	0.29	659	10
24	1.78	0.29	662	10
25	1.87	0.29	665	10
26	1.95	0.29	668	10
27	2.04	0.29	671	10
28	2.12	0.29	674	10
29	2.21	0.29	677	10
30	2.30	0.30	680	10
31	2.39	0.30	684	10
32	2.48	0.30	687	11
33	2.57	0.31	690	11
34	2.66	0.31	693	11
35	2.76	0.31	697	11
36	2.86	0.32	700	11
37	2.96	0.32	704	11
38	3.07	0.33	707	12
39	3.18	0.34	711	12
40	3.30	0.35	715	12
41	3.42	0.36	720	12
42	3.55	0.37	724	13
43	3.69	0.38	729	13
44	3.85	0.40	735	14
45	4.02	0.42	741	15
46	4.20	0.45	747	16
47	4.42	0.48	755	17
48	4.68	0.53	764	19
49	5.00	0.61	775	21
50	5.44	0.73	790	26
51	6.16	1.02	816	36
52	7.57	2.01	865	70

APPENDIX D: IRT STATISTICS

Table D.1a: IRT Statistics—Listening and Speaking, Grades K–1

	Item #	N-Count	Rasch Difficulty	SE	MNSQ INFIT	MNSQ OUTFIT	Misfit
LISTENING	1	51,942	-3.19	0.02	1.14	1.28	--
	2	51,942	-3.52	0.02	1.09	1.15	--
	3	51,942	-3.55	0.02	0.97	0.82	--
	4	51,942	-3.47	0.02	0.86	0.90	--
	5	51,942	-2.87	0.02	1.02	0.82	--
	6	51,942	-1.90	0.01	1.08	1.04	--
	7	51,942	-3.04	0.02	0.93	0.68	--
	8	51,942	-3.77	0.02	0.95	0.70	--
	9	51,942	-2.35	0.02	1.14	1.32	--
	10	51,942	-3.65	0.02	1.03	1.26	--
	13	51,942	-1.96	0.01	1.07	1.03	--
	14	51,942	-1.68	0.01	1.28	1.35	--
	15	51,942	-2.76	0.02	1.09	1.14	--
	16	51,942	-2.33	0.02	1.33	2.09	--
	17	51,942	-3.40	0.02	0.97	0.90	--
	18	51,942	-3.26	0.02	0.78	0.54	--
	19	51,942	-2.24	0.02	1.20	1.41	--
	20	51,942	-2.54	0.02	0.96	1.01	--
	21	51,942	-2.63	0.02	1.18	1.48	--
	22	51,942	-1.98	0.01	1.12	1.09	--
	23	51,942	-0.58	0.01	1.03	1.15	--
	26	51,942	-1.11	0.01	1.30	1.73	--
	27	51,942	-1.08	0.01	1.12	1.30	--
	28	51,942	-2.59	0.02	1.26	1.90	--
SPEAKING	1	51,942	-2.03	0.01	0.94	0.98	--
	2	51,942	-1.89	0.01	0.90	0.85	--
	3	51,942	-1.15	0.01	0.96	0.95	--
	4	51,942	-1.68	0.01	0.82	0.80	--
	6	51,942	-1.47	0.01	1.03	1.04	--
	7	51,942	-0.65	0.01	0.79	0.79	--
	8	51,942	-2.34	0.01	0.81	0.80	--
	9	51,942	-1.62	0.01	0.78	0.78	--
	10	51,942	-1.61	0.01	0.77	0.71	--
	11	51,942	-2.04	0.01	0.78	0.72	--
	12	51,942	-1.43	0.01	0.75	0.73	--
	13	51,942	-2.42	0.01	1.01	1.15	--
	14	51,942	-1.26	0.01	0.91	0.81	--
	15	51,942	-2.30	0.01	0.85	0.80	--
	17	51,942	-1.70	0.01	0.86	0.84	--
	18	51,942	-2.09	0.01	0.85	0.79	--

Note: "√" indicates that the item was flagged as misfit, and "--" indicates that the item was not flagged.

Table D.1b: IRT Statistics—Reading and Writing, Grades K–1

	Item #	N-Count	Rasch Difficulty	SE	MNSQ INFIT	MNSQ OUTFIT	Misfit
READING	1	51,942	-3.28	0.01	0.93	0.75	--
	2	51,942	-2.73	0.01	1.00	0.78	--
	3	51,942	-1.76	0.01	1.13	1.16	--
	4	51,942	-3.71	0.02	0.90	0.61	--
	5	51,942	-3.31	0.01	0.84	0.60	--
	6	51,942	-2.92	0.01	0.83	0.73	--
	7	51,942	-2.73	0.01	0.99	0.79	--
	8	51,942	-1.13	0.01	1.04	1.08	--
	12	51,942	-3.14	0.01	0.99	0.85	--
	13	51,942	-2.87	0.01	0.90	0.95	--
	14	51,942	-1.88	0.01	0.96	0.88	--
	15	51,942	-1.66	0.01	1.14	1.12	--
	16	51,942	-2.86	0.01	1.05	0.88	--
	20	51,942	-2.15	0.01	1.24	1.26	--
	21	51,942	-2.56	0.01	1.32	1.72	--
WRITING	1	51,942	-4.69	0.02	0.87	0.64	--
	2	51,942	-3.74	0.02	0.97	0.88	--
	3	51,942	-2.76	0.01	1.10	1.21	--
	4	51,942	-3.02	0.01	1.07	2.16	--
	5	51,942	-3.40	0.02	1.07	1.49	--
	6	51,942	-4.42	0.02	0.86	0.97	--
	7	51,942	-3.34	0.01	1.17	1.37	--
	8	51,942	-3.33	0.01	1.24	1.47	--
	9	51,942	-3.23	0.01	1.13	2.20	--
	10	51,942	-3.27	0.01	1.19	3.52	--
	11	51,942	-1.93	0.01	0.94	0.79	--
	12	51,942	-0.71	0.01	0.94	0.90	--
	13	51,942	-1.48	0.01	0.73	0.73	--
	14	51,942	-1.27	0.01	0.65	0.62	--
	15	51,942	-0.31	0.01	0.71	0.67	--

Table D.2a: IRT Statistics—Listening and Speaking, Grades 2–4

	Item #	N-Count	Rasch Difficulty	SE	MNSQ INFIT	MNSQ OUTFIT	Misfit
LISTENING	1	58,036	-2.11	0.02	0.94	0.67	--
	2	58,036	-1.75	0.02	0.93	0.57	--
	3	58,036	-0.77	0.02	1.02	0.79	--
	6	58,036	-1.24	0.02	0.93	0.68	--
	7	58,036	-1.08	0.02	0.97	0.74	--
	8	58,036	-1.99	0.02	0.97	0.89	--
	9	58,036	1.75	0.01	1.21	1.67	--
	10	58,036	-0.74	0.02	1.23	1.37	--
	11	58,036	-0.57	0.02	1.15	1.15	--
	12	58,036	0.06	0.01	1.30	1.45	--
	15	58,036	-0.07	0.01	0.98	0.82	--
	16	58,036	1.04	0.01	1.23	1.47	--
	17	58,036	1.75	0.01	1.05	1.30	--
	18	58,036	0.04	0.01	0.98	0.93	--
	19	58,036	1.91	0.01	1.18	1.68	--
	20	58,036	1.30	0.01	1.17	1.41	--
	21	58,036	0.26	0.01	0.96	0.87	--
	22	58,036	1.15	0.01	1.20	1.39	--
	23	58,036	1.57	0.01	1.18	1.51	--
	24	58,036	-1.32	0.02	0.95	0.82	--
	25	58,036	-1.00	0.02	1.24	1.41	--
	26	58,036	0.97	0.01	1.23	1.43	--
	27	58,036	-0.60	0.02	1.28	1.43	--
	28	58,036	-0.92	0.02	1.22	1.58	--
SPEAKING	1	58,036	-0.18	0.01	0.76	0.70	--
	2	58,036	-0.66	0.01	0.73	0.67	--
	3	58,036	-0.29	0.01	0.75	0.75	--
	4	58,036	-0.68	0.01	0.86	0.86	--
	6	58,036	-0.38	0.01	0.79	0.76	--
	7	58,036	0.26	0.01	0.79	0.82	--
	8	58,036	-0.29	0.01	0.76	0.71	--
	9	58,036	-0.51	0.01	0.84	0.81	--
	10	58,036	-0.66	0.01	0.70	0.62	--
	11	58,036	-0.60	0.01	0.73	0.65	--
	12	58,036	-0.30	0.01	0.77	0.77	--
	13	58,036	-0.61	0.01	0.94	0.95	--
	14	58,036	-0.36	0.01	0.83	0.79	--
	15	58,036	-1.07	0.01	0.95	1.02	--
	16	58,036	-0.07	0.01	0.77	0.68	--
	18	58,036	-0.15	0.01	0.80	0.68	--

Table D.2b: IRT Statistics—Reading and Writing, Grades 2–4

	Item #	N-Count	Rasch Difficulty	SE	MNSQ INFIT	MNSQ OUTFIT	Misfit
READING	1	58,036	-2.23	0.02	0.85	0.55	--
	2	58,036	-2.72	0.02	0.83	0.39	--
	3	58,036	-1.64	0.01	0.90	0.72	--
	7	58,036	-1.87	0.02	0.90	0.82	--
	8	58,036	-1.77	0.01	0.99	1.23	--
	9	58,036	-1.74	0.01	0.98	1.05	--
	10	58,036	-1.97	0.02	1.10	0.85	--
	14	58,036	-0.82	0.01	1.00	0.96	--
	15	58,036	0.58	0.01	1.19	1.31	--
	16	58,036	-0.09	0.01	0.89	0.79	--
	17	58,036	0.78	0.01	0.97	1.03	--
	18	58,036	1.15	0.01	1.00	1.16	--
	19	58,036	0.61	0.01	1.07	1.16	--
	20	58,036	-0.24	0.01	0.93	0.85	--
	21	58,036	-0.66	0.01	0.88	0.73	--
	22	58,036	0.21	0.01	0.93	0.88	--
	23	58,036	0.75	0.01	0.93	0.96	--
	24	58,036	0.27	0.01	1.07	1.14	--
	25	58,036	0.06	0.01	1.05	1.14	--
	26	58,036	-1.69	0.01	0.89	0.74	--
	27	58,036	1.36	0.01	1.10	1.36	--
	28	58,036	0.38	0.01	0.95	0.94	--
	29	58,036	-0.67	0.01	1.24	1.65	--
	30	58,036	1.40	0.01	1.34	1.85	--
WRITING	1	58,036	-2.44	0.02	0.89	0.61	--
	2	58,036	-2.44	0.02	0.84	0.49	--
	3	58,036	-1.75	0.01	0.97	0.89	--
	4	58,036	-1.25	0.01	0.98	0.81	--
	5	58,036	-2.02	0.02	0.82	0.67	--
	6	58,036	-0.23	0.01	0.88	0.82	--
	7	58,036	0.37	0.01	1.08	1.14	--
	8	58,036	-0.66	0.01	0.98	0.96	--
	9	58,036	-0.76	0.01	0.98	1.01	--
	10	58,036	0.26	0.01	1.08	1.13	--
	11	58,036	0.00	0.01	1.03	1.14	--
	12	58,036	0.24	0.01	1.06	1.12	--
	13	58,036	-1.33	0.01	0.84	0.91	--
	14	58,036	-0.19	0.01	1.03	1.07	--
	15	58,036	-0.11	0.01	1.00	1.00	--
	16	58,036	1.22	0.01	0.89	0.88	--

Table D.3a: IRT Statistics—Listening and Speaking, Grades 5–6

	Item #	N-Count	Rasch Difficulty	SE	MNSQ INFIT	MNSQ OUTFIT	Misfit
LISTENING	1	28,623	-1.07	0.02	0.97	0.62	--
	2	28,623	-1.31	0.03	1.06	0.69	--
	3	28,623	1.04	0.02	1.41	1.73	--
	4	28,623	-0.01	0.02	1.02	0.94	--
	5	28,623	-0.36	0.02	1.02	0.90	--
	6	28,623	-0.64	0.02	1.14	1.21	--
	7	28,623	-0.68	0.02	0.89	0.63	--
	8	28,623	0.57	0.02	1.32	1.57	--
	9	28,623	0.36	0.02	1.25	1.41	--
	10	28,623	0.26	0.02	1.08	1.12	--
	11	28,623	0.21	0.02	1.16	1.19	--
	12	28,623	0.17	0.02	1.06	1.12	--
	13	28,623	0.73	0.02	1.29	1.53	--
	14	28,623	0.35	0.02	1.38	1.68	--
	15	28,623	0.85	0.02	1.14	1.25	--
	16	28,623	0.35	0.02	1.03	1.01	--
	17	28,623	1.84	0.01	1.12	1.39	--
	18	28,623	-0.55	0.02	0.88	0.60	--
	19	28,623	0.36	0.02	1.04	0.98	--
	20	28,623	-0.10	0.02	1.23	1.32	--
	21	28,623	2.04	0.01	1.33	1.79	--
	22	28,623	0.23	0.02	0.94	0.78	--
	23	28,623	0.44	0.02	0.98	0.95	--
	24	28,623	0.06	0.02	1.11	1.07	--
	25	28,623	0.28	0.02	1.05	1.04	--
SPEAKING	1	28,623	-0.93	0.02	0.72	0.60	--
	2	28,623	-0.67	0.02	0.71	0.61	--
	3	28,623	-0.34	0.01	0.74	0.66	--
	4	28,623	-0.44	0.01	0.91	0.90	--
	5	28,623	0.49	0.01	0.79	0.82	--
	6	28,623	0.08	0.01	0.74	0.75	--
	7	28,623	-0.45	0.01	0.66	0.55	--
	8	28,623	0.20	0.01	0.67	0.63	--
	9	28,623	-0.81	0.02	0.64	0.51	--
	10	28,623	-1.02	0.02	0.67	0.55	--
	11	28,623	-0.59	0.01	0.68	0.58	--
	12	28,623	-1.10	0.02	0.92	0.84	--
	13	28,623	-0.43	0.02	0.72	0.69	--
	14	28,623	-0.42	0.01	0.84	0.80	--
	15	28,623	-0.25	0.01	0.83	0.76	--
	16	28,623	-0.67	0.02	0.84	0.78	--

Table D.3b: IRT Statistics—Reading and Writing, Grades 5–6

	Item #	N-Count	Rasch Difficulty	SE	MNSQ INFIT	MNSQ OUTFIT	Misfit
READING	1	28,623	-1.11	0.03	0.84	0.45	--
	2	28,623	-0.99	0.02	0.83	0.45	--
	3	28,623	0.23	0.02	1.15	1.24	--
	4	28,623	0.19	0.02	1.04	1.03	--
	5	28,623	-0.43	0.02	0.82	0.58	--
	6	28,623	-0.05	0.02	0.89	0.67	--
	7	28,623	1.63	0.01	1.18	1.25	--
	8	28,623	1.09	0.02	1.01	0.97	--
	9	28,623	1.83	0.01	1.02	1.06	--
	10	28,623	1.42	0.01	1.04	1.05	--
	11	28,623	1.73	0.01	0.99	0.98	--
	12	28,623	1.59	0.01	1.03	1.03	--
	13	28,623	1.80	0.01	1.05	1.08	--
	14	28,623	1.07	0.02	0.96	0.96	--
	15	28,623	1.47	0.01	1.01	0.98	--
	16	28,623	2.24	0.01	1.02	1.11	--
	17	28,623	0.95	0.02	0.91	0.86	--
	18	28,623	2.29	0.01	1.21	1.36	--
	19	28,623	0.58	0.02	0.90	0.73	--
	20	28,623	1.42	0.01	0.88	0.80	--
	21	28,623	2.12	0.01	1.09	1.19	--
	22	28,623	1.70	0.01	0.91	0.88	--
	23	28,623	1.10	0.02	0.86	0.74	--
	24	28,623	2.52	0.01	1.17	1.38	--
	25	28,623	1.20	0.02	0.95	0.90	--
	26	28,623	1.13	0.02	0.92	0.88	--
	27	28,623	1.14	0.02	0.91	0.79	--
WRITING	1	28,623	-1.13	0.03	0.92	0.66	--
	2	28,623	0.68	0.02	1.32	1.53	--
	3	28,623	-0.47	0.02	0.82	0.62	--
	4	28,623	0.43	0.02	0.94	0.89	--
	5	28,623	-0.02	0.02	0.91	0.98	--
	6	28,623	0.79	0.02	0.97	0.97	--
	7	28,623	0.83	0.02	1.03	1.09	--
	8	28,623	0.45	0.02	0.91	0.89	--
	9	28,623	-0.35	0.02	0.76	0.59	--
	10	28,623	0.61	0.02	1.01	0.97	--
	11	28,623	0.54	0.02	1.10	1.30	--
	12	28,623	0.42	0.02	1.07	1.10	--
	13	28,623	1.51	0.01	1.15	1.24	--
	14	28,623	1.82	0.01	1.20	1.30	--
	15	28,623	1.95	0.01	1.21	1.36	--
	16	28,623	-0.22	0.01	0.87	1.02	--
	17	28,623	1.03	0.01	0.94	0.95	--
	18	28,623	1.56	0.01	1.02	1.01	--
	19	28,623	2.33	0.01	0.85	0.84	--

Table D.4a: IRT Statistics—Listening and Speaking, Grades 7–8

	Item #	N-Count	Rasch Difficulty	SE	MNSQ INFIT	MNSQ OUTFIT	Misfit
LISTENING	1	25,115	-0.66	0.03	1.20	1.55	--
	2	25,115	-0.49	0.03	1.34	1.29	--
	3	25,115	0.42	0.02	0.97	0.67	--
	4	25,115	1.59	0.02	1.32	1.55	--
	5	25,115	2.18	0.02	1.21	1.46	--
	6	25,115	0.92	0.02	1.01	0.93	--
	7	25,115	2.51	0.02	1.28	1.59	--
	8	25,115	1.35	0.02	1.37	1.49	--
	9	25,115	0.45	0.02	1.21	1.20	--
	10	25,115	1.51	0.02	1.27	1.42	--
	11	25,115	2.32	0.02	1.11	1.29	--
	12	25,115	2.14	0.02	1.33	1.55	--
	13	25,115	2.23	0.02	1.25	1.44	--
	14	25,115	2.93	0.01	1.11	1.46	--
	15	25,115	1.01	0.02	1.08	1.08	--
	16	25,115	0.19	0.02	1.08	1.28	--
	17	25,115	0.54	0.02	1.15	1.21	--
	18	25,115	1.07	0.02	1.20	1.26	--
	19	25,115	1.00	0.02	1.06	1.06	--
	20	25,115	0.46	0.02	1.42	1.72	--
	21	25,115	1.77	0.02	1.17	1.26	--
	22	25,115	1.67	0.02	1.22	1.31	--
	23	25,115	0.48	0.02	1.23	1.38	--
	24	25,115	1.81	0.02	1.07	1.07	--
	25	25,115	0.89	0.02	1.01	0.91	--
SPEAKING	1	25,115	0.21	0.01	0.68	0.58	--
	2	25,115	0.57	0.01	0.69	0.64	--
	3	25,115	0.61	0.01	0.70	0.66	--
	4	25,115	0.64	0.01	0.67	0.62	--
	5	25,115	1.15	0.01	0.74	0.70	--
	6	25,115	1.25	0.01	0.66	0.68	--
	7	25,115	0.53	0.01	0.68	0.64	--
	8	25,115	0.28	0.01	0.68	0.60	--
	9	25,115	0.45	0.01	0.66	0.58	--
	10	25,115	0.46	0.01	0.66	0.59	--
	11	25,115	0.88	0.01	0.74	0.70	--
	12	25,115	0.59	0.01	0.72	0.59	--
	13	25,115	0.39	0.01	0.75	0.69	--
	14	25,115	0.89	0.01	0.61	0.54	--
	15	25,115	0.26	0.02	0.69	0.55	--
	16	25,115	0.69	0.01	0.65	0.56	--

Table D.4b: IRT Statistics—Reading and Writing, Grades 7–8

	Item #	N-Count	Rasch Difficulty	SE	MNSQ INFIT	MNSQ OUTFIT	Misfit
READING	1	25,115	-1.20	0.03	0.99	0.92	--
	2	25,115	-0.22	0.02	0.88	0.97	--
	3	25,115	-0.97	0.02	0.98	0.88	--
	4	25,115	1.50	0.01	1.02	1.02	--
	5	25,115	0.35	0.02	0.90	0.77	--
	6	25,115	0.31	0.02	0.83	0.72	--
	7	25,115	0.32	0.02	0.87	0.73	--
	8	25,115	1.16	0.01	0.91	0.86	--
	9	25,115	0.73	0.02	0.91	0.85	--
	10	25,115	1.68	0.01	1.08	1.12	--
	11	25,115	1.20	0.01	1.00	0.98	--
	12	25,115	1.24	0.01	1.06	1.07	--
	13	25,115	0.78	0.02	0.99	0.94	--
	14	25,115	1.07	0.02	0.96	0.94	--
	15	25,115	1.65	0.01	1.04	1.07	--
	16	25,115	2.26	0.01	1.14	1.29	--
	17	25,115	1.76	0.01	1.15	1.25	--
	18	25,115	1.30	0.01	0.97	0.93	--
	19	25,115	1.55	0.01	0.97	0.96	--
	20	25,115	1.66	0.01	1.12	1.16	--
	21	25,115	1.86	0.01	0.95	0.97	--
	22	25,115	2.23	0.01	1.10	1.25	--
	23	25,115	1.09	0.02	0.97	0.95	--
	24	25,115	1.48	0.01	1.08	1.11	--
	25	25,115	1.52	0.01	0.97	0.95	--
	26	25,115	1.79	0.01	1.14	1.19	--
	27	25,115	2.16	0.01	1.19	1.34	--
WRITING	1	25,115	-1.02	0.02	0.91	0.64	--
	2	25,115	-0.60	0.02	0.92	0.98	--
	3	25,115	-0.61	0.02	0.83	0.57	--
	4	25,115	-0.77	0.02	1.14	1.10	--
	5	25,115	0.19	0.02	0.91	0.83	--
	6	25,115	0.31	0.02	0.91	0.82	--
	7	25,115	1.69	0.01	1.12	1.19	--
	8	25,115	0.33	0.02	0.92	0.88	--
	9	25,115	0.31	0.02	0.82	0.66	--
	10	25,115	1.08	0.02	0.98	0.99	--
	11	25,115	1.76	0.01	1.29	1.42	--
	12	25,115	0.43	0.02	1.10	1.21	--
	13	25,115	1.13	0.02	1.01	1.01	--
	14	25,115	0.76	0.02	1.06	1.11	--
	15	25,115	1.42	0.01	1.09	1.15	--
	16	25,115	0.43	0.01	0.79	0.74	--
	17	25,115	1.19	0.01	1.01	1.02	--
	18	25,115	0.76	0.01	0.87	0.88	--
	19	25,115	1.92	0.01	0.82	0.82	--

Table D.5a: IRT Statistics—Listening and Speaking, Grades 9–12

	Item #	N-Count	Rasch Difficulty	SE	MNSQ INFIT	MNSQ OUTFIT	Misfit
LISTENING	1	44,704	0.61	0.02	1.15	1.06	--
	2	44,704	0.93	0.01	1.32	1.27	--
	3	44,704	0.12	0.02	0.94	0.63	--
	4	44,704	1.33	0.01	1.09	1.05	--
	5	44,704	2.06	0.01	1.32	1.52	--
	6	44,704	2.23	0.01	1.10	1.18	--
	7	44,704	3.07	0.01	1.40	1.91	--
	8	44,704	1.85	0.01	1.08	1.09	--
	9	44,704	1.46	0.01	1.09	1.12	--
	10	44,704	0.79	0.01	1.46	2.32	--
	11	44,704	0.73	0.01	1.29	1.65	--
	12	44,704	2.99	0.01	1.19	1.43	--
	13	44,704	3.13	0.01	1.09	1.40	--
	14	44,704	1.52	0.01	1.15	1.29	--
	15	44,704	2.55	0.01	1.06	1.16	--
	16	44,704	3.03	0.01	1.14	1.37	--
	17	44,704	2.24	0.01	1.14	1.24	--
	18	44,704	3.04	0.01	0.96	1.10	--
	19	44,704	2.73	0.01	1.28	1.54	--
	20	44,704	2.73	0.01	1.21	1.39	--
	21	44,704	0.93	0.01	1.08	1.03	--
	22	44,704	2.26	0.01	1.01	1.02	--
	23	44,704	2.19	0.01	1.12	1.20	--
	24	44,704	1.63	0.01	1.14	1.22	--
	25	44,704	1.79	0.01	1.18	1.24	--
SPEAKING	1	44,704	1.26	0.01	0.88	0.84	--
	2	44,704	0.89	0.01	0.76	0.69	--
	3	44,704	0.75	0.01	0.78	0.70	--
	4	44,704	1.08	0.01	0.79	0.74	--
	5	44,704	0.94	0.01	0.77	0.75	--
	6	44,704	1.37	0.01	0.73	0.75	--
	7	44,704	1.01	0.01	0.70	0.65	--
	8	44,704	0.69	0.01	0.73	0.66	--
	9	44,704	0.80	0.01	0.72	0.64	--
	10	44,704	0.86	0.01	0.72	0.66	--
	11	44,704	0.55	0.01	0.69	0.60	--
	12	44,704	1.66	0.01	0.75	0.64	--
	13	44,704	0.68	0.01	0.75	0.66	--
	14	44,704	1.19	0.01	0.77	0.70	--
	15	44,704	2.28	0.01	0.84	0.71	--
	16	44,704	0.86	0.01	0.73	0.64	--

Table D.5b: IRT Statistics—Reading and Writing, Grades 9–12

	Item #	N-Count	Rasch Difficulty	SE	MNSQ INFIT	MNSQ OUTFIT	Misfit
READING	1	44,704	0.23	0.02	0.95	0.83	--
	2	44,704	0.43	0.01	0.91	0.77	--
	3	44,704	0.89	0.01	1.03	0.98	--
	4	44,704	1.88	0.01	1.09	1.07	--
	5	44,704	1.09	0.01	0.90	0.79	--
	6	44,704	1.05	0.01	0.86	0.75	--
	7	44,704	3.12	0.01	1.13	1.24	--
	8	44,704	2.87	0.01	1.14	1.23	--
	9	44,704	2.58	0.01	1.10	1.15	--
	10	44,704	1.75	0.01	1.01	1.01	--
	11	44,704	2.99	0.01	1.13	1.21	--
	12	44,704	2.92	0.01	1.08	1.13	--
	13	44,704	1.99	0.01	0.90	0.84	--
	14	44,704	1.23	0.01	0.93	0.90	--
	15	44,704	2.16	0.01	0.88	0.85	--
	16	44,704	3.31	0.01	1.09	1.20	--
	17	44,704	2.23	0.01	1.10	1.11	--
	18	44,704	1.13	0.01	0.97	0.93	--
	19	44,704	3.08	0.01	1.02	1.09	--
	20	44,704	2.88	0.01	1.06	1.09	--
	21	44,704	2.88	0.01	1.01	1.03	--
	22	44,704	2.38	0.01	1.16	1.22	--
	23	44,704	2.70	0.01	1.18	1.26	--
	24	44,704	2.32	0.01	1.16	1.26	--
	25	44,704	2.46	0.01	0.99	0.98	--
	26	44,704	1.94	0.01	0.98	0.94	--
	27	44,704	1.66	0.01	0.91	0.83	--
WRITING	1	44,704	0.01	0.02	0.96	0.86	--
	2	44,704	0.34	0.02	1.03	0.93	--
	3	44,704	1.31	0.01	0.90	0.82	--
	4	44,704	2.80	0.01	1.03	1.07	--
	5	44,704	1.99	0.01	0.97	0.95	--
	6	44,704	2.52	0.01	1.04	1.07	--
	7	44,704	1.91	0.01	0.99	0.98	--
	8	44,704	1.56	0.01	1.01	1.00	--
	9	44,704	2.59	0.01	0.93	0.92	--
	10	44,704	1.04	0.01	0.94	0.89	--
	11	44,704	2.22	0.01	1.10	1.13	--
	12	44,704	1.42	0.01	0.85	0.73	--
	13	44,704	3.05	0.01	1.08	1.15	--
	14	44,704	1.83	0.01	0.84	0.75	--
	15	44,704	1.99	0.01	0.93	0.90	--
	16	44,704	1.40	0.01	0.85	0.86	--
	17	44,704	1.43	0.01	0.79	0.76	--
	18	44,704	1.83	0.01	0.89	0.90	--
	19	44,704	2.55	0.01	0.88	0.88	--

APPENDIX E: ITEM INFORMATION AT DIFFERENT CUT POINTS[3]

Table E.1a: Item Information by Cut Point—Listening and Speaking, Grade K

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	-3.19	0.23	0.07	0.02
	2	-3.52	0.20	0.06	0.02
	3	-3.55	0.20	0.05	0.02
	4	-3.47	0.20	0.06	0.02
	5	-2.87	0.24	0.10	0.03
	6	-1.90	0.23	0.18	0.08
	7	-3.04	0.24	0.08	0.03
	8	-3.77	0.18	0.04	0.01
	9	-2.35	0.25	0.14	0.05
	10	-3.65	0.19	0.05	0.02
	11	-1.96	0.23	0.18	0.07
	12	-1.68	0.21	0.20	0.09
	13	-2.76	0.25	0.11	0.04
	14	-2.33	0.25	0.14	0.05
	15	-3.40	0.21	0.06	0.02
	16	-3.26	0.22	0.07	0.02
	17	-2.24	0.24	0.15	0.06
	18	-2.54	0.25	0.12	0.04
	19	-2.63	0.25	0.12	0.04
	20	-1.98	0.23	0.18	0.07
	21	-0.58	0.11	0.25	0.19
	22	-1.11	0.15	0.24	0.14
	23	-1.08	0.15	0.24	0.14
	24	-2.59	0.25	0.12	0.04
SPEAKING	1	-2.03	0.23	0.17	0.07
	2	-1.89	0.22	0.18	0.08
	3	-1.15	0.16	0.24	0.14
	4	-1.68	0.21	0.20	0.09
	5	-1.47	0.19	0.22	0.11
	6	-0.65	0.11	0.25	0.19
	7	-2.34	0.25	0.14	0.05
	8	-1.62	0.20	0.21	0.10
	9	-1.61	0.20	0.21	0.10
	10	-2.04	0.23	0.17	0.07
	11	-1.43	0.19	0.22	0.11
	12	-2.42	0.25	0.13	0.05
	13	-1.26	0.17	0.23	0.13
	14	-2.30	0.25	0.15	0.06
	15	-1.70	0.21	0.20	0.09
	16	-2.09	0.24	0.17	0.07

[3] The statistics in Tables E.1–E.13 were based on the equating sample. Invalid cases were removed after cleaning procedures.

Table E.1b: Item Information by Cut Point—Reading and Writing, Grade K

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	-3.28	0.15	0.06	0.03
	2	-2.73	0.20	0.09	0.05
	3	-1.76	0.25	0.17	0.10
	4	-3.71	0.11	0.04	0.02
	5	-3.31	0.15	0.05	0.03
	6	-2.92	0.18	0.08	0.04
	7	-2.73	0.20	0.09	0.05
	8	-1.13	0.22	0.23	0.16
	9	-3.14	0.16	0.06	0.03
	10	-2.87	0.19	0.08	0.04
	11	-1.88	0.25	0.16	0.10
	12	-1.66	0.25	0.18	0.11
	13	-2.86	0.19	0.08	0.04
	14	-2.15	0.24	0.14	0.08
	15	-2.56	0.22	0.10	0.05
WRITING	1	-4.69	0.05	0.01	0.01
	2	-3.74	0.11	0.04	0.02
	3	-2.76	0.20	0.09	0.04
	4	-3.02	0.17	0.07	0.04
	5	-3.40	0.14	0.05	0.02
	6	-4.42	0.06	0.02	0.01
	7	-3.34	0.14	0.05	0.03
	8	-3.33	0.15	0.05	0.03
	9	-3.23	0.16	0.06	0.03
	10	-3.27	0.15	0.06	0.03
	11	-1.93	0.25	0.16	0.09
	12	-0.71	0.19	0.25	0.20
	13	-1.48	0.24	0.20	0.13
	14	-1.27	0.23	0.22	0.15
	15	-0.31	0.15	0.25	0.23

Table E.2a: Item Information by Cut Point—Listening and Speaking, Grade 1

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	-3.19	0.20	0.06	0.01
	2	-3.52	0.17	0.05	0.01
	3	-3.55	0.17	0.04	0.01
	4	-3.47	0.18	0.05	0.01
	5	-2.87	0.23	0.08	0.01
	6	-1.90	0.24	0.16	0.03
	7	-3.04	0.22	0.07	0.01
	8	-3.77	0.15	0.04	0.01
	9	-2.35	0.25	0.12	0.02
	10	-3.65	0.16	0.04	0.01
	11	-1.96	0.24	0.15	0.03
	12	-1.68	0.23	0.18	0.04
	13	-2.76	0.24	0.09	0.02
	14	-2.33	0.25	0.12	0.02
	15	-3.40	0.19	0.05	0.01
	16	-3.26	0.20	0.06	0.01
	17	-2.24	0.25	0.13	0.03
	18	-2.54	0.25	0.10	0.02
	19	-2.63	0.24	0.10	0.02
	20	-1.98	0.24	0.15	0.03
	21	-0.58	0.13	0.25	0.11
	22	-1.11	0.18	0.23	0.07
	23	-1.08	0.18	0.23	0.07
	24	-2.59	0.24	0.10	0.02
SPEAKING	1	-2.03	0.25	0.15	0.03
	2	-1.89	0.24	0.16	0.03
	3	-1.15	0.18	0.23	0.07
	4	-1.68	0.23	0.18	0.04
	5	-1.47	0.21	0.20	0.05
	6	-0.65	0.14	0.25	0.10
	7	-2.34	0.25	0.12	0.02
	8	-1.62	0.22	0.19	0.04
	9	-1.61	0.22	0.19	0.05
	10	-2.04	0.25	0.15	0.03
	11	-1.43	0.21	0.20	0.05
	12	-2.42	0.25	0.11	0.02
	13	-1.26	0.20	0.22	0.06
	14	-2.30	0.25	0.12	0.02
	15	-1.70	0.23	0.18	0.04
	16	-2.09	0.25	0.14	0.03

Table E.2b: Item Information by Cut Point—Reading and Writing, Grade 1

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	-3.28	0.10	0.03	0.01
	2	-2.73	0.14	0.05	0.01
	3	-1.76	0.23	0.10	0.03
	4	-3.71	0.07	0.02	0.00
	5	-3.31	0.09	0.03	0.01
	6	-2.92	0.13	0.04	0.01
	7	-2.73	0.14	0.05	0.01
	8	-1.13	0.25	0.16	0.06
	9	-3.14	0.11	0.03	0.01
	10	-2.87	0.13	0.04	0.01
	11	-1.88	0.22	0.10	0.03
	12	-1.66	0.24	0.11	0.04
	13	-2.86	0.13	0.04	0.01
	14	-2.15	0.20	0.08	0.02
	15	-2.56	0.16	0.05	0.02
WRITING	1	-4.69	0.03	0.01	0.00
	2	-3.74	0.07	0.02	0.00
	3	-2.76	0.14	0.04	0.01
	4	-3.02	0.12	0.04	0.01
	5	-3.40	0.09	0.02	0.01
	6	-4.42	0.04	0.01	0.00
	7	-3.34	0.09	0.03	0.01
	8	-3.33	0.09	0.03	0.01
	9	-3.23	0.10	0.03	0.01
	10	-3.27	0.10	0.03	0.01
	11	-1.93	0.22	0.09	0.03
	12	-0.71	0.24	0.20	0.08
	13	-1.48	0.24	0.13	0.04
	14	-1.27	0.25	0.15	0.05
	15	-0.31	0.21	0.23	0.11

Table E.3a: Item Information by Cut Point—Listening and Speaking, Grade 2

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	-2.11	0.24	0.09	0.02
	2	-1.75	0.25	0.12	0.03
	3	-0.77	0.21	0.21	0.07
	4	-1.24	0.24	0.17	0.05
	5	-1.08	0.23	0.18	0.05
	6	-1.99	0.24	0.10	0.02
	7	1.75	0.03	0.13	0.25
	8	-0.74	0.21	0.21	0.07
	9	-0.57	0.19	0.23	0.08
	10	0.06	0.13	0.25	0.13
	11	-0.07	0.14	0.25	0.12
	12	1.04	0.06	0.20	0.22
	13	1.75	0.03	0.13	0.25
	14	0.04	0.13	0.25	0.13
	15	1.91	0.03	0.12	0.25
	16	1.30	0.05	0.17	0.24
	17	0.26	0.11	0.25	0.15
	18	1.15	0.06	0.19	0.23
	19	1.57	0.04	0.15	0.25
	20	-1.32	0.24	0.16	0.04
	21	-1.00	0.23	0.19	0.06
	22	0.97	0.06	0.20	0.22
	23	-0.60	0.19	0.22	0.08
	24	-0.92	0.22	0.20	0.06
SPEAKING	1	-0.18	0.15	0.25	0.11
	2	-0.66	0.20	0.22	0.08
	3	-0.29	0.16	0.24	0.10
	4	-0.68	0.20	0.22	0.08
	5	-0.38	0.17	0.24	0.10
	6	0.26	0.11	0.25	0.15
	7	-0.29	0.16	0.24	0.10
	8	-0.51	0.19	0.23	0.09
	9	-0.66	0.20	0.22	0.08
	10	-0.60	0.19	0.22	0.08
	11	-0.30	0.17	0.24	0.10
	12	-0.61	0.19	0.22	0.08
	13	-0.36	0.17	0.24	0.10
	14	-1.07	0.23	0.18	0.05
	15	-0.07	0.14	0.25	0.12
	16	-0.15	0.15	0.25	0.11

Table E.3b: Item Information by Cut Point—Reading and Writing, Grade 2

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	-2.23	0.16	0.05	0.02
	2	-2.72	0.11	0.03	0.01
	3	-1.64	0.21	0.08	0.03
	4	-1.87	0.19	0.07	0.02
	5	-1.77	0.20	0.08	0.03
	6	-1.74	0.20	0.08	0.03
	7	-1.97	0.18	0.06	0.02
	8	-0.82	0.25	0.15	0.06
	9	0.58	0.16	0.25	0.17
	10	-0.09	0.22	0.22	0.11
	11	0.78	0.14	0.25	0.19
	12	1.15	0.11	0.23	0.22
	13	0.61	0.16	0.25	0.18
	14	-0.24	0.23	0.21	0.10
	15	-0.66	0.25	0.17	0.07
	16	0.21	0.19	0.24	0.14
	17	0.75	0.14	0.25	0.19
	18	0.27	0.19	0.24	0.15
	19	0.06	0.21	0.23	0.13
	20	-1.69	0.21	0.08	0.03
	21	1.36	0.09	0.22	0.24
	22	0.38	0.18	0.25	0.16
	23	-0.67	0.25	0.17	0.07
	24	1.40	0.09	0.22	0.24
WRITING	1	-2.44	0.14	0.04	0.01
	2	-2.44	0.14	0.04	0.01
	3	-1.75	0.20	0.08	0.03
	4	-1.25	0.24	0.11	0.04
	5	-2.02	0.18	0.06	0.02
	6	-0.23	0.23	0.21	0.10
	7	0.37	0.18	0.25	0.16
	8	-0.66	0.25	0.17	0.07
	9	-0.76	0.25	0.16	0.07
	10	0.26	0.19	0.24	0.14
	11	0.00	0.21	0.23	0.12
	12	0.24	0.19	0.24	0.14
	13	-1.33	0.23	0.11	0.04
	14	-0.19	0.23	0.21	0.10
	15	-0.11	0.22	0.22	0.11
	16	1.22	0.10	0.23	0.23

Table E.4a: Item Information by Cut Point—Listening and Speaking, Grade 3

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	-2.11	0.21	0.08	0.01
	2	-1.75	0.24	0.10	0.02
	3	-0.77	0.24	0.19	0.05
	4	-1.24	0.25	0.14	0.03
	5	-1.08	0.25	0.16	0.04
	6	-1.99	0.22	0.08	0.02
	7	1.75	0.04	0.15	0.24
	8	-0.74	0.23	0.19	0.05
	9	-0.57	0.22	0.21	0.06
	10	0.06	0.17	0.25	0.10
	11	-0.07	0.18	0.24	0.09
	12	1.04	0.08	0.22	0.19
	13	1.75	0.04	0.15	0.24
	14	0.04	0.17	0.25	0.10
	15	1.91	0.04	0.14	0.25
	16	1.30	0.07	0.20	0.21
	17	0.26	0.15	0.25	0.12
	18	1.15	0.08	0.21	0.20
	19	1.57	0.05	0.17	0.23
	20	-1.32	0.25	0.14	0.03
	21	-1.00	0.25	0.17	0.04
	22	0.97	0.09	0.22	0.18
	23	-0.60	0.22	0.20	0.06
	24	-0.92	0.24	0.18	0.04
SPEAKING	1	-0.18	0.19	0.24	0.08
	2	-0.66	0.23	0.20	0.05
	3	-0.29	0.20	0.23	0.07
	4	-0.68	0.23	0.20	0.05
	5	-0.38	0.21	0.22	0.07
	6	0.26	0.15	0.25	0.11
	7	-0.29	0.20	0.23	0.07
	8	-0.51	0.22	0.21	0.06
	9	-0.66	0.23	0.20	0.05
	10	-0.60	0.22	0.20	0.06
	11	-0.30	0.20	0.23	0.07
	12	-0.61	0.22	0.20	0.06
	13	-0.36	0.20	0.22	0.07
	14	-1.07	0.25	0.16	0.04
	15	-0.07	0.18	0.24	0.09
	16	-0.15	0.19	0.24	0.08

Table E.4b: Item Information by Cut Point—Reading and Writing, Grade 3

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	-2.23	0.13	0.04	0.01
	2	-2.72	0.09	0.03	0.01
	3	-1.64	0.19	0.07	0.02
	4	-1.87	0.17	0.06	0.02
	5	-1.77	0.18	0.06	0.02
	6	-1.74	0.18	0.06	0.02
	7	-1.97	0.16	0.05	0.01
	8	-0.82	0.25	0.13	0.04
	9	0.58	0.18	0.25	0.13
	10	-0.09	0.24	0.20	0.08
	11	0.78	0.16	0.25	0.15
	12	1.15	0.13	0.24	0.19
	13	0.61	0.18	0.25	0.14
	14	-0.24	0.24	0.19	0.07
	15	-0.66	0.25	0.15	0.05
	16	0.21	0.22	0.23	0.10
	17	0.75	0.17	0.25	0.15
	18	0.27	0.21	0.23	0.11
	19	0.06	0.23	0.21	0.09
	20	-1.69	0.19	0.07	0.02
	21	1.36	0.11	0.24	0.21
	22	0.38	0.20	0.24	0.12
	23	-0.67	0.25	0.15	0.05
	24	1.40	0.11	0.23	0.21
WRITING	1	-2.44	0.12	0.03	0.01
	2	-2.44	0.12	0.03	0.01
	3	-1.75	0.18	0.06	0.02
	4	-1.25	0.22	0.10	0.03
	5	-2.02	0.15	0.05	0.01
	6	-0.23	0.24	0.19	0.07
	7	0.37	0.20	0.24	0.12
	8	-0.66	0.25	0.15	0.05
	9	-0.76	0.25	0.14	0.04
	10	0.26	0.21	0.23	0.11
	11	0.00	0.23	0.21	0.09
	12	0.24	0.21	0.23	0.10
	13	-1.33	0.22	0.09	0.03
	14	-0.19	0.24	0.19	0.07
	15	-0.11	0.24	0.20	0.08
	16	1.22	0.12	0.24	0.19

Table E.5a: Item Information by Cut Point—Listening and Speaking, Grade 4

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	-2.11	0.18	0.06	0.01
	2	-1.75	0.21	0.08	0.02
	3	-0.77	0.25	0.16	0.04
	4	-1.24	0.24	0.12	0.03
	5	-1.08	0.25	0.13	0.03
	6	-1.99	0.19	0.07	0.01
	7	1.75	0.06	0.18	0.23
	8	-0.74	0.25	0.17	0.04
	9	-0.57	0.24	0.18	0.05
	10	0.06	0.20	0.23	0.08
	11	-0.07	0.21	0.22	0.07
	12	1.04	0.11	0.24	0.16
	13	1.75	0.06	0.18	0.23
	14	0.04	0.20	0.23	0.08
	15	1.91	0.05	0.17	0.24
	16	1.30	0.09	0.22	0.19
	17	0.26	0.18	0.24	0.10
	18	1.15	0.10	0.23	0.17
	19	1.57	0.07	0.20	0.21
	20	-1.32	0.24	0.11	0.02
	21	-1.00	0.25	0.14	0.03
	22	0.97	0.11	0.24	0.16
	23	-0.60	0.24	0.18	0.05
	24	-0.92	0.25	0.15	0.03
SPEAKING	1	-0.18	0.22	0.22	0.07
	2	-0.66	0.24	0.17	0.04
	3	-0.29	0.22	0.21	0.06
	4	-0.68	0.25	0.17	0.04
	5	-0.38	0.23	0.20	0.06
	6	0.26	0.18	0.24	0.09
	7	-0.29	0.22	0.21	0.06
	8	-0.51	0.24	0.19	0.05
	9	-0.66	0.24	0.17	0.04
	10	-0.60	0.24	0.18	0.05
	11	-0.30	0.23	0.21	0.06
	12	-0.61	0.24	0.18	0.05
	13	-0.36	0.23	0.20	0.06
	14	-1.07	0.25	0.13	0.03
	15	-0.07	0.21	0.22	0.07
	16	-0.15	0.21	0.22	0.07

Table E.5b: Item Information by Cut Point—Reading and Writing, Grade 4

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	-2.23	0.11	0.03	0.01
	2	-2.72	0.08	0.02	0.00
	3	-1.64	0.17	0.05	0.01
	4	-1.87	0.15	0.04	0.01
	5	-1.77	0.15	0.05	0.01
	6	-1.74	0.16	0.05	0.01
	7	-1.97	0.14	0.04	0.01
	8	-0.82	0.24	0.10	0.03
	9	0.58	0.20	0.23	0.10
	10	-0.09	0.25	0.17	0.06
	11	0.78	0.19	0.24	0.12
	12	1.15	0.15	0.25	0.15
	13	0.61	0.20	0.23	0.11
	14	-0.24	0.25	0.16	0.05
	15	-0.66	0.24	0.12	0.04
	16	0.21	0.23	0.20	0.08
	17	0.75	0.19	0.24	0.12
	18	0.27	0.23	0.20	0.08
	19	0.06	0.24	0.18	0.07
	20	-1.69	0.16	0.05	0.01
	21	1.36	0.13	0.25	0.17
	22	0.38	0.22	0.21	0.09
	23	-0.67	0.24	0.12	0.03
	24	1.40	0.13	0.25	0.18
WRITING	1	-2.44	0.10	0.03	0.01
	2	-2.44	0.10	0.03	0.01
	3	-1.75	0.16	0.05	0.01
	4	-1.25	0.20	0.07	0.02
	5	-2.02	0.13	0.04	0.01
	6	-0.23	0.25	0.16	0.05
	7	0.37	0.22	0.21	0.09
	8	-0.66	0.24	0.12	0.04
	9	-0.76	0.24	0.11	0.03
	10	0.26	0.23	0.20	0.08
	11	0.00	0.24	0.18	0.06
	12	0.24	0.23	0.20	0.08
	13	-1.33	0.20	0.07	0.02
	14	-0.19	0.25	0.16	0.05
	15	-0.11	0.25	0.17	0.06
	16	1.22	0.14	0.25	0.16

Table E.6a: Item Information by Cut Point—Listening and Speaking, Grade 5

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	-1.07	0.24	0.11	0.03
	2	-1.31	0.22	0.09	0.02
	3	1.04	0.14	0.25	0.16
	4	-0.01	0.23	0.21	0.07
	5	-0.36	0.25	0.17	0.05
	6	-0.64	0.25	0.15	0.04
	7	-0.68	0.25	0.14	0.04
	8	0.57	0.18	0.24	0.11
	9	0.36	0.20	0.23	0.10
	10	0.26	0.21	0.23	0.09
	11	0.21	0.21	0.22	0.09
	12	0.17	0.22	0.22	0.08
	13	0.73	0.17	0.25	0.13
	14	0.35	0.20	0.23	0.10
	15	0.85	0.15	0.25	0.14
	16	0.35	0.20	0.23	0.10
	17	1.84	0.07	0.20	0.23
	18	-0.55	0.25	0.15	0.04
	19	0.36	0.20	0.23	0.10
	20	-0.10	0.24	0.20	0.07
	21	2.04	0.06	0.18	0.24
	22	0.23	0.21	0.22	0.09
	23	0.44	0.19	0.24	0.10
	24	0.06	0.23	0.21	0.08
	25	0.28	0.21	0.23	0.09
SPEAKING	1	-0.93	0.24	0.12	0.03
	2	-0.67	0.25	0.14	0.04
	3	-0.34	0.25	0.17	0.05
	4	-0.44	0.25	0.17	0.05
	5	0.49	0.19	0.24	0.11
	6	0.08	0.22	0.21	0.08
	7	-0.45	0.25	0.16	0.05
	8	0.20	0.21	0.22	0.09
	9	-0.81	0.25	0.13	0.04
	10	-1.02	0.24	0.11	0.03
	11	-0.59	0.25	0.15	0.04
	12	-1.10	0.23	0.11	0.03
	13	-0.43	0.25	0.17	0.05
	14	-0.42	0.25	0.17	0.05
	15	-0.25	0.24	0.18	0.06
	16	-0.67	0.25	0.14	0.04

Table E.6b: Item Information by Cut Point—Reading and Writing, Grade 5

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	-1.11	0.13	0.05	0.02
	2	-0.99	0.14	0.06	0.02
	3	0.23	0.24	0.15	0.06
	4	0.19	0.24	0.15	0.05
	5	-0.43	0.20	0.09	0.03
	6	-0.05	0.23	0.12	0.04
	7	1.63	0.19	0.25	0.16
	8	1.09	0.23	0.23	0.11
	9	1.83	0.17	0.25	0.18
	10	1.42	0.21	0.24	0.14
	11	1.73	0.18	0.25	0.17
	12	1.59	0.19	0.25	0.16
	13	1.80	0.17	0.25	0.18
	14	1.07	0.23	0.22	0.11
	15	1.47	0.20	0.25	0.15
	16	2.24	0.13	0.23	0.22
	17	0.95	0.24	0.22	0.10
	18	2.29	0.13	0.23	0.22
	19	0.58	0.25	0.18	0.08
	20	1.42	0.21	0.24	0.14
	21	2.12	0.14	0.24	0.21
	22	1.70	0.18	0.25	0.17
	23	1.10	0.23	0.23	0.11
	24	2.52	0.11	0.21	0.24
	25	1.20	0.23	0.23	0.12
	26	1.13	0.23	0.23	0.12
	27	1.14	0.23	0.23	0.12
WRITING	1	-1.13	0.13	0.05	0.02
	2	0.68	0.25	0.19	0.08
	3	-0.47	0.19	0.09	0.03
	4	0.43	0.25	0.17	0.07
	5	-0.02	0.23	0.13	0.05
	6	0.79	0.25	0.20	0.09
	7	0.83	0.25	0.21	0.09
	8	0.45	0.25	0.17	0.07
	9	-0.35	0.20	0.10	0.03
	10	0.61	0.25	0.19	0.08
	11	0.54	0.25	0.18	0.07
	12	0.42	0.25	0.17	0.07
	13	1.51	0.20	0.25	0.15
	14	1.82	0.17	0.25	0.18
	15	1.95	0.16	0.25	0.19
	16	-0.22	0.22	0.11	0.04
	17	1.03	0.24	0.22	0.11
	18	1.56	0.20	0.25	0.16
	19	2.33	0.12	0.23	0.23

Table E.7a: Item Information by Cut Point—Listening and Speaking, Grade 6

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	-1.07	0.22	0.09	0.02
	2	-1.31	0.20	0.07	0.02
	3	1.04	0.16	0.25	0.13
	4	-0.01	0.24	0.19	0.06
	5	-0.36	0.25	0.15	0.04
	6	-0.64	0.25	0.13	0.03
	7	-0.68	0.24	0.12	0.03
	8	0.57	0.20	0.23	0.09
	9	0.36	0.22	0.22	0.08
	10	0.26	0.23	0.21	0.07
	11	0.21	0.23	0.21	0.07
	12	0.17	0.23	0.20	0.07
	13	0.73	0.19	0.24	0.10
	14	0.35	0.22	0.22	0.08
	15	0.85	0.18	0.25	0.11
	16	0.35	0.22	0.22	0.08
	17	1.84	0.09	0.22	0.21
	18	-0.55	0.25	0.13	0.03
	19	0.36	0.22	0.22	0.08
	20	-0.10	0.25	0.18	0.05
	21	2.04	0.08	0.20	0.22
	22	0.23	0.23	0.21	0.07
	23	0.44	0.21	0.22	0.08
	24	0.06	0.24	0.19	0.06
	25	0.28	0.23	0.21	0.07
SPEAKING	1	-0.93	0.23	0.10	0.02
	2	-0.67	0.24	0.12	0.03
	3	-0.34	0.25	0.15	0.04
	4	-0.44	0.25	0.14	0.04
	5	0.49	0.21	0.23	0.09
	6	0.08	0.24	0.19	0.06
	7	-0.45	0.25	0.14	0.04
	8	0.20	0.23	0.20	0.07
	9	-0.81	0.24	0.11	0.03
	10	-1.02	0.23	0.09	0.02
	11	-0.59	0.25	0.13	0.03
	12	-1.10	0.22	0.09	0.02
	13	-0.43	0.25	0.15	0.04
	14	-0.42	0.25	0.15	0.04
	15	-0.25	0.25	0.16	0.05
	16	-0.67	0.24	0.12	0.03

Table E.7b: Item Information by Cut Point—Reading and Writing, Grade 6

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	-1.11	0.12	0.04	0.01
	2	-0.99	0.14	0.05	0.02
	3	0.23	0.24	0.13	0.05
	4	0.19	0.24	0.13	0.05
	5	-0.43	0.19	0.08	0.03
	6	-0.05	0.22	0.11	0.04
	7	1.63	0.20	0.24	0.15
	8	1.09	0.24	0.21	0.10
	9	1.83	0.18	0.25	0.17
	10	1.42	0.22	0.23	0.13
	11	1.73	0.19	0.25	0.16
	12	1.59	0.20	0.24	0.14
	13	1.80	0.18	0.25	0.16
	14	1.07	0.24	0.21	0.10
	15	1.47	0.21	0.24	0.13
	16	2.24	0.14	0.24	0.20
	17	0.95	0.24	0.20	0.09
	18	2.29	0.14	0.24	0.21
	19	0.58	0.25	0.16	0.07
	20	1.42	0.22	0.23	0.13
	21	2.12	0.15	0.25	0.19
	22	1.70	0.19	0.25	0.15
	23	1.10	0.24	0.21	0.10
	24	2.52	0.12	0.23	0.23
	25	1.20	0.23	0.22	0.11
	26	1.13	0.24	0.21	0.10
	27	1.14	0.24	0.21	0.10
WRITING	1	-1.13	0.12	0.04	0.01
	2	0.68	0.25	0.17	0.07
	3	-0.47	0.18	0.08	0.03
	4	0.43	0.25	0.15	0.06
	5	-0.02	0.22	0.11	0.04
	6	0.79	0.25	0.18	0.08
	7	0.83	0.25	0.19	0.08
	8	0.45	0.25	0.15	0.06
	9	-0.35	0.20	0.08	0.03
	10	0.61	0.25	0.17	0.07
	11	0.54	0.25	0.16	0.06
	12	0.42	0.25	0.15	0.06
	13	1.51	0.21	0.24	0.14
	14	1.82	0.18	0.25	0.17
	15	1.95	0.17	0.25	0.18
	16	-0.22	0.21	0.09	0.03
	17	1.03	0.24	0.20	0.10
	18	1.56	0.21	0.24	0.14
	19	2.33	0.13	0.24	0.21

Table E.8a: Item Information by Cut Point—Listening and Speaking, Grade 7

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	-0.66	0.23	0.10	0.03
	2	-0.49	0.24	0.11	0.03
	3	0.42	0.23	0.19	0.07
	4	1.59	0.13	0.25	0.16
	5	2.18	0.09	0.22	0.22
	6	0.92	0.20	0.23	0.10
	7	2.51	0.06	0.19	0.24
	8	1.35	0.16	0.25	0.14
	9	0.45	0.23	0.20	0.07
	10	1.51	0.14	0.25	0.15
	11	2.32	0.08	0.21	0.23
	12	2.14	0.09	0.22	0.21
	13	2.23	0.08	0.22	0.22
	14	2.93	0.04	0.15	0.25
	15	1.01	0.19	0.24	0.11
	16	0.19	0.25	0.17	0.06
	17	0.54	0.23	0.20	0.08
	18	1.07	0.18	0.24	0.11
	19	1.00	0.19	0.24	0.11
	20	0.46	0.23	0.20	0.07
	21	1.77	0.12	0.24	0.18
	22	1.67	0.13	0.25	0.17
	23	0.48	0.23	0.20	0.07
	24	1.81	0.11	0.24	0.18
	25	0.89	0.20	0.23	0.10
SPEAKING	1	0.21	0.24	0.17	0.06
	2	0.57	0.23	0.21	0.08
	3	0.61	0.22	0.21	0.08
	4	0.64	0.22	0.21	0.08
	5	1.15	0.18	0.24	0.12
	6	1.25	0.17	0.25	0.13
	7	0.53	0.23	0.20	0.07
	8	0.28	0.24	0.18	0.06
	9	0.45	0.23	0.20	0.07
	10	0.46	0.23	0.20	0.07
	11	0.88	0.20	0.23	0.10
	12	0.59	0.22	0.21	0.08
	13	0.39	0.24	0.19	0.07
	14	0.89	0.20	0.23	0.10
	15	0.26	0.24	0.18	0.06
	16	0.69	0.22	0.22	0.09

Table E.8b: Item Information by Cut Point—Reading and Writing, Grade 7

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	-1.20	0.10	0.03	0.01
	2	-0.22	0.19	0.08	0.03
	3	-0.97	0.12	0.04	0.02
	4	1.50	0.22	0.23	0.14
	5	0.35	0.24	0.12	0.05
	6	0.31	0.23	0.12	0.05
	7	0.32	0.23	0.12	0.05
	8	1.16	0.24	0.20	0.11
	9	0.73	0.25	0.16	0.07
	10	1.68	0.21	0.24	0.15
	11	1.20	0.24	0.20	0.11
	12	1.24	0.24	0.21	0.11
	13	0.78	0.25	0.16	0.08
	14	1.07	0.25	0.19	0.10
	15	1.65	0.21	0.24	0.15
	16	2.26	0.16	0.25	0.21
	17	1.76	0.20	0.24	0.16
	18	1.30	0.24	0.21	0.12
	19	1.55	0.22	0.23	0.14
	20	1.66	0.21	0.24	0.15
	21	1.86	0.19	0.25	0.17
	22	2.23	0.16	0.25	0.20
	23	1.09	0.25	0.19	0.10
	24	1.48	0.23	0.23	0.13
	25	1.52	0.22	0.23	0.14
	26	1.79	0.20	0.24	0.16
	27	2.16	0.17	0.25	0.20
WRITING	1	-1.02	0.12	0.04	0.01
	2	-0.60	0.16	0.06	0.02
	3	-0.61	0.15	0.06	0.02
	4	-0.77	0.14	0.05	0.02
	5	0.19	0.23	0.11	0.05
	6	0.31	0.23	0.12	0.05
	7	1.69	0.21	0.24	0.15
	8	0.33	0.24	0.12	0.05
	9	0.31	0.23	0.12	0.05
	10	1.08	0.25	0.19	0.10
	11	1.76	0.20	0.24	0.16
	12	0.43	0.24	0.13	0.06
	13	1.13	0.24	0.20	0.10
	14	0.76	0.25	0.16	0.08
	15	1.42	0.23	0.22	0.13
	16	0.43	0.24	0.13	0.06
	17	1.19	0.24	0.20	0.11
	18	0.76	0.25	0.16	0.08
	19	1.92	0.19	0.25	0.18

Table E.9b: Item Information by Cut Point—Listening and Speaking, Grade 8

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	-0.66	0.21	0.08	0.03
	2	-0.49	0.23	0.09	0.03
	3	0.42	0.25	0.17	0.07
	4	1.59	0.16	0.25	0.16
	5	2.18	0.10	0.24	0.22
	6	0.92	0.22	0.21	0.10
	7	2.51	0.08	0.22	0.24
	8	1.35	0.18	0.24	0.14
	9	0.45	0.24	0.17	0.07
	10	1.51	0.16	0.25	0.15
	11	2.32	0.09	0.23	0.23
	12	2.14	0.11	0.24	0.21
	13	2.23	0.10	0.24	0.22
	14	2.93	0.05	0.18	0.25
	15	1.01	0.21	0.22	0.11
	16	0.19	0.25	0.14	0.06
	17	0.54	0.24	0.18	0.08
	18	1.07	0.20	0.22	0.11
	19	1.00	0.21	0.22	0.11
	20	0.46	0.24	0.17	0.07
	21	1.77	0.14	0.25	0.18
	22	1.67	0.15	0.25	0.17
	23	0.48	0.24	0.17	0.07
	24	1.81	0.13	0.25	0.18
	25	0.89	0.22	0.21	0.10
SPEAKING	1	0.21	0.25	0.15	0.06
	2	0.57	0.24	0.18	0.08
	3	0.61	0.24	0.18	0.08
	4	0.64	0.24	0.19	0.08
	5	1.15	0.20	0.23	0.12
	6	1.25	0.19	0.24	0.13
	7	0.53	0.24	0.18	0.07
	8	0.28	0.25	0.15	0.06
	9	0.45	0.24	0.17	0.07
	10	0.46	0.24	0.17	0.07
	11	0.88	0.22	0.21	0.10
	12	0.59	0.24	0.18	0.08
	13	0.39	0.25	0.16	0.07
	14	0.89	0.22	0.21	0.10
	15	0.26	0.25	0.15	0.06
	16	0.69	0.23	0.19	0.09

Table E.9b: Item Information by Cut Point—Reading and Writing, Grade 8

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	-1.20	0.10	0.03	0.01
	2	-0.22	0.18	0.07	0.03
	3	-0.97	0.11	0.03	0.02
	4	1.50	0.23	0.21	0.14
	5	0.35	0.23	0.11	0.05
	6	0.31	0.23	0.10	0.05
	7	0.32	0.23	0.10	0.05
	8	1.16	0.25	0.18	0.11
	9	0.73	0.25	0.14	0.07
	10	1.68	0.22	0.22	0.15
	11	1.20	0.25	0.18	0.11
	12	1.24	0.24	0.19	0.11
	13	0.78	0.25	0.14	0.08
	14	1.07	0.25	0.17	0.10
	15	1.65	0.22	0.22	0.15
	16	2.26	0.16	0.25	0.21
	17	1.76	0.21	0.23	0.16
	18	1.30	0.24	0.19	0.12
	19	1.55	0.23	0.22	0.14
	20	1.66	0.22	0.22	0.15
	21	1.86	0.20	0.24	0.17
	22	2.23	0.17	0.25	0.20
	23	1.09	0.25	0.17	0.10
	24	1.48	0.23	0.21	0.13
	25	1.52	0.23	0.21	0.14
	26	1.79	0.21	0.23	0.16
	27	2.16	0.17	0.25	0.20
WRITING	1	-1.02	0.11	0.03	0.01
	2	-0.60	0.15	0.05	0.02
	3	-0.61	0.15	0.05	0.02
	4	-0.77	0.13	0.04	0.02
	5	0.19	0.22	0.09	0.05
	6	0.31	0.23	0.10	0.05
	7	1.69	0.22	0.23	0.15
	8	0.33	0.23	0.10	0.05
	9	0.31	0.23	0.10	0.05
	10	1.08	0.25	0.17	0.10
	11	1.76	0.21	0.23	0.16
	12	0.43	0.24	0.11	0.06
	13	1.13	0.25	0.18	0.10
	14	0.76	0.25	0.14	0.08
	15	1.42	0.23	0.20	0.13
	16	0.43	0.24	0.11	0.06
	17	1.19	0.25	0.18	0.11
	18	0.76	0.25	0.14	0.08
	19	1.92	0.20	0.24	0.18

Table E.10a: Item Information by Cut Point—Listening and Speaking, Grade 9

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	0.61	0.25	0.14	0.06
	2	0.93	0.25	0.17	0.08
	3	0.12	0.23	0.10	0.04
	4	1.33	0.22	0.21	0.11
	5	2.06	0.16	0.25	0.18
	6	2.23	0.14	0.25	0.19
	7	3.07	0.08	0.21	0.25
	8	1.85	0.18	0.24	0.16
	9	1.46	0.21	0.22	0.12
	10	0.79	0.25	0.16	0.07
	11	0.73	0.25	0.15	0.07
	12	2.99	0.08	0.22	0.24
	13	3.13	0.07	0.21	0.25
	14	1.52	0.21	0.22	0.12
	15	2.55	0.11	0.24	0.22
	16	3.03	0.08	0.21	0.25
	17	2.24	0.14	0.25	0.19
	18	3.04	0.08	0.21	0.25
	19	2.73	0.10	0.23	0.23
	20	2.73	0.10	0.24	0.23
	21	0.93	0.25	0.17	0.08
	22	2.26	0.14	0.25	0.20
	23	2.19	0.15	0.25	0.19
	24	1.63	0.20	0.23	0.14
	25	1.79	0.18	0.24	0.15
SPEAKING	1	1.26	0.23	0.20	0.10
	2	0.89	0.25	0.16	0.08
	3	0.75	0.25	0.15	0.07
	4	1.08	0.24	0.18	0.09
	5	0.94	0.25	0.17	0.08
	6	1.37	0.22	0.21	0.11
	7	1.01	0.24	0.18	0.08
	8	0.69	0.25	0.15	0.06
	9	0.80	0.25	0.16	0.07
	10	0.86	0.25	0.16	0.07
	11	0.55	0.25	0.13	0.06
	12	1.66	0.20	0.23	0.14
	13	0.68	0.25	0.14	0.06
	14	1.19	0.23	0.19	0.10
	15	2.28	0.14	0.25	0.20
	16	0.86	0.25	0.16	0.08

Table E.10b: Item Information by Cut Point—Reading and Writing, Grade 9

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	0.23	0.20	0.07	0.03
	2	0.43	0.22	0.08	0.04
	3	0.89	0.24	0.11	0.06
	4	1.88	0.22	0.20	0.13
	5	1.09	0.25	0.13	0.07
	6	1.05	0.25	0.13	0.07
	7	3.12	0.11	0.24	0.24
	8	2.87	0.13	0.25	0.22
	9	2.58	0.16	0.25	0.20
	10	1.75	0.23	0.19	0.12
	11	2.99	0.12	0.25	0.23
	12	2.92	0.13	0.25	0.22
	13	1.99	0.22	0.21	0.14
	14	1.23	0.25	0.14	0.08
	15	2.16	0.20	0.23	0.15
	16	3.31	0.10	0.23	0.24
	17	2.23	0.19	0.23	0.16
	18	1.13	0.25	0.13	0.07
	19	3.08	0.12	0.25	0.23
	20	2.88	0.13	0.25	0.22
	21	2.88	0.13	0.25	0.22
	22	2.38	0.18	0.24	0.18
	23	2.70	0.15	0.25	0.21
	24	2.32	0.19	0.24	0.17
	25	2.46	0.17	0.24	0.18
	26	1.94	0.22	0.21	0.13
	27	1.66	0.24	0.18	0.11
WRITING	1	0.01	0.18	0.05	0.03
	2	0.34	0.21	0.07	0.04
	3	1.31	0.25	0.15	0.08
	4	2.80	0.14	0.25	0.21
	5	1.99	0.22	0.21	0.14
	6	2.52	0.17	0.25	0.19
	7	1.91	0.22	0.21	0.13
	8	1.56	0.24	0.17	0.10
	9	2.59	0.16	0.25	0.20
	10	1.04	0.25	0.13	0.07
	11	2.22	0.20	0.23	0.16
	12	1.42	0.25	0.16	0.09
	13	3.05	0.12	0.25	0.23
	14	1.83	0.23	0.20	0.12
	15	1.99	0.22	0.21	0.14
	16	1.40	0.25	0.16	0.09
	17	1.43	0.25	0.16	0.09
	18	1.83	0.23	0.20	0.12
	19	2.55	0.16	0.25	0.19

Table E.11a: Item Information by Cut Point—Listening and Speaking, Grade 10

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	0.61	0.25	0.12	0.05
	2	0.93	0.25	0.15	0.07
	3	0.12	0.22	0.08	0.03
	4	1.33	0.23	0.19	0.10
	5	2.06	0.17	0.24	0.16
	6	2.23	0.16	0.25	0.17
	7	3.07	0.08	0.23	0.24
	8	1.85	0.19	0.23	0.14
	9	1.46	0.22	0.20	0.10
	10	0.79	0.25	0.14	0.06
	11	0.73	0.25	0.13	0.06
	12	2.99	0.09	0.23	0.24
	13	3.13	0.08	0.22	0.24
	14	1.52	0.22	0.20	0.11
	15	2.55	0.13	0.25	0.20
	16	3.03	0.09	0.23	0.24
	17	2.24	0.15	0.25	0.18
	18	3.04	0.09	0.23	0.24
	19	2.73	0.11	0.24	0.22
	20	2.73	0.11	0.24	0.22
	21	0.93	0.25	0.15	0.07
	22	2.26	0.15	0.25	0.18
	23	2.19	0.16	0.25	0.17
	24	1.63	0.21	0.21	0.12
	25	1.79	0.20	0.23	0.13
SPEAKING	1	1.26	0.24	0.18	0.09
	2	0.89	0.25	0.14	0.07
	3	0.75	0.25	0.13	0.06
	4	1.08	0.25	0.16	0.08
	5	0.94	0.25	0.15	0.07
	6	1.37	0.23	0.19	0.10
	7	1.01	0.25	0.16	0.07
	8	0.69	0.25	0.13	0.06
	9	0.80	0.25	0.14	0.06
	10	0.86	0.25	0.14	0.06
	11	0.55	0.25	0.11	0.05
	12	1.66	0.21	0.22	0.12
	13	0.68	0.25	0.13	0.05
	14	1.19	0.24	0.17	0.09
	15	2.28	0.15	0.25	0.18
	16	0.86	0.25	0.14	0.06

Table E.11b: Item Information by Cut Point—Reading and Writing, Grade 10

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	0.23	0.19	0.06	0.03
	2	0.43	0.21	0.07	0.03
	3	0.89	0.24	0.10	0.05
	4	1.88	0.23	0.19	0.11
	5	1.09	0.25	0.12	0.06
	6	1.05	0.25	0.12	0.06
	7	3.12	0.12	0.25	0.22
	8	2.87	0.14	0.25	0.20
	9	2.58	0.17	0.24	0.18
	10	1.75	0.24	0.18	0.10
	11	2.99	0.13	0.25	0.21
	12	2.92	0.14	0.25	0.21
	13	1.99	0.22	0.20	0.12
	14	1.23	0.25	0.13	0.07
	15	2.16	0.21	0.22	0.14
	16	3.31	0.10	0.24	0.24
	17	2.23	0.20	0.22	0.14
	18	1.13	0.25	0.12	0.06
	19	3.08	0.12	0.25	0.22
	20	2.88	0.14	0.25	0.21
	21	2.88	0.14	0.25	0.21
	22	2.38	0.19	0.23	0.16
	23	2.70	0.16	0.25	0.19
	24	2.32	0.19	0.23	0.15
	25	2.46	0.18	0.24	0.17
	26	1.94	0.23	0.20	0.12
	27	1.66	0.24	0.17	0.10
WRITING	1	0.01	0.17	0.05	0.02
	2	0.34	0.20	0.07	0.03
	3	1.31	0.25	0.14	0.07
	4	2.80	0.15	0.25	0.20
	5	1.99	0.22	0.20	0.12
	6	2.52	0.18	0.24	0.17
	7	1.91	0.23	0.20	0.12
	8	1.56	0.25	0.16	0.09
	9	2.59	0.17	0.24	0.18
	10	1.04	0.25	0.12	0.06
	11	2.22	0.20	0.22	0.14
	12	1.42	0.25	0.15	0.08
	13	3.05	0.13	0.25	0.22
	14	1.83	0.23	0.19	0.11
	15	1.99	0.22	0.20	0.12
	16	1.40	0.25	0.15	0.08
	17	1.43	0.25	0.15	0.08
	18	1.83	0.23	0.19	0.11
	19	2.55	0.17	0.24	0.18

Table E.12a: Item Information by Cut Point—Listening and Speaking, Grade 11

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	0.61	0.24	0.10	0.04
	2	0.93	0.25	0.13	0.06
	3	0.12	0.21	0.07	0.03
	4	1.33	0.24	0.16	0.08
	5	2.06	0.19	0.23	0.14
	6	2.23	0.18	0.24	0.16
	7	3.07	0.10	0.24	0.23
	8	1.85	0.21	0.21	0.12
	9	1.46	0.24	0.18	0.09
	10	0.79	0.25	0.11	0.05
	11	0.73	0.25	0.11	0.05
	12	2.99	0.11	0.24	0.22
	13	3.13	0.09	0.24	0.23
	14	1.52	0.23	0.18	0.09
	15	2.55	0.14	0.25	0.19
	16	3.03	0.10	0.24	0.23
	17	2.24	0.17	0.24	0.16
	18	3.04	0.10	0.24	0.23
	19	2.73	0.13	0.25	0.20
	20	2.73	0.13	0.25	0.20
	21	0.93	0.25	0.13	0.06
	22	2.26	0.17	0.24	0.16
	23	2.19	0.18	0.24	0.15
	24	1.63	0.23	0.19	0.10
	25	1.79	0.21	0.21	0.12
SPEAKING	1	1.26	0.25	0.16	0.08
	2	0.89	0.25	0.12	0.06
	3	0.75	0.25	0.11	0.05
	4	1.08	0.25	0.14	0.07
	5	0.94	0.25	0.13	0.06
	6	1.37	0.24	0.17	0.08
	7	1.01	0.25	0.13	0.06
	8	0.69	0.24	0.11	0.05
	9	0.80	0.25	0.11	0.05
	10	0.86	0.25	0.12	0.05
	11	0.55	0.24	0.09	0.04
	12	1.66	0.22	0.19	0.10
	13	0.68	0.24	0.10	0.05
	14	1.19	0.25	0.15	0.07
	15	2.28	0.17	0.24	0.16
	16	0.86	0.25	0.12	0.05

Table E.12b: Item Information by Cut Point—Reading and Writing, Grade 11

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	0.23	0.18	0.05	0.02
	2	0.43	0.20	0.06	0.03
	3	0.89	0.23	0.08	0.04
	4	1.88	0.24	0.17	0.10
	5	1.09	0.24	0.10	0.05
	6	1.05	0.24	0.10	0.05
	7	3.12	0.13	0.25	0.21
	8	2.87	0.15	0.24	0.19
	9	2.58	0.18	0.23	0.16
	10	1.75	0.24	0.16	0.09
	11	2.99	0.14	0.25	0.20
	12	2.92	0.15	0.25	0.19
	13	1.99	0.23	0.18	0.11
	14	1.23	0.25	0.11	0.06
	15	2.16	0.22	0.20	0.12
	16	3.31	0.11	0.25	0.22
	17	2.23	0.21	0.20	0.13
	18	1.13	0.25	0.10	0.05
	19	3.08	0.13	0.25	0.21
	20	2.88	0.15	0.25	0.19
	21	2.88	0.15	0.25	0.19
	22	2.38	0.20	0.22	0.14
	23	2.70	0.17	0.24	0.17
	24	2.32	0.20	0.21	0.14
	25	2.46	0.19	0.22	0.15
	26	1.94	0.23	0.18	0.10
	27	1.66	0.25	0.15	0.08
WRITING	1	0.01	0.16	0.04	0.02
	2	0.34	0.19	0.05	0.03
	3	1.31	0.25	0.12	0.06
	4	2.80	0.16	0.24	0.18
	5	1.99	0.23	0.18	0.11
	6	2.52	0.18	0.23	0.15
	7	1.91	0.23	0.17	0.10
	8	1.56	0.25	0.14	0.08
	9	2.59	0.18	0.23	0.16
	10	1.04	0.24	0.10	0.05
	11	2.22	0.21	0.20	0.13
	12	1.42	0.25	0.13	0.07
	13	3.05	0.13	0.25	0.20
	14	1.83	0.24	0.17	0.09
	15	1.99	0.23	0.18	0.11
	16	1.40	0.25	0.13	0.07
	17	1.43	0.25	0.13	0.07
	18	1.83	0.24	0.17	0.09
	19	2.55	0.18	0.23	0.16

Table E.13a: Item Information by Cut Point—Listening and Speaking, Grade 12

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
LISTENING	1	0.61	0.23	0.08	0.04
	2	0.93	0.25	0.10	0.06
	3	0.12	0.19	0.05	0.03
	4	1.33	0.25	0.14	0.08
	5	2.06	0.20	0.21	0.14
	6	2.23	0.19	0.22	0.16
	7	3.07	0.11	0.25	0.23
	8	1.85	0.22	0.19	0.12
	9	1.46	0.24	0.15	0.09
	10	0.79	0.24	0.09	0.05
	11	0.73	0.24	0.09	0.05
	12	2.99	0.12	0.25	0.22
	13	3.13	0.11	0.25	0.23
	14	1.52	0.24	0.15	0.09
	15	2.55	0.16	0.24	0.19
	16	3.03	0.11	0.25	0.23
	17	2.24	0.19	0.22	0.16
	18	3.04	0.11	0.25	0.23
	19	2.73	0.14	0.25	0.20
	20	2.73	0.14	0.25	0.20
	21	0.93	0.25	0.10	0.06
	22	2.26	0.19	0.22	0.16
	23	2.19	0.19	0.22	0.15
	24	1.63	0.24	0.17	0.10
	25	1.79	0.23	0.18	0.12
SPEAKING	1	1.26	0.25	0.13	0.08
	2	0.89	0.25	0.10	0.06
	3	0.75	0.24	0.09	0.05
	4	1.08	0.25	0.11	0.07
	5	0.94	0.25	0.10	0.06
	6	1.37	0.25	0.14	0.08
	7	1.01	0.25	0.11	0.06
	8	0.69	0.24	0.08	0.05
	9	0.80	0.24	0.09	0.05
	10	0.86	0.24	0.10	0.05
	11	0.55	0.23	0.08	0.04
	12	1.66	0.23	0.17	0.10
	13	0.68	0.24	0.08	0.05
	14	1.19	0.25	0.12	0.07
	15	2.28	0.18	0.22	0.16
	16	0.86	0.25	0.10	0.05

Table E.13b: Item Information by Cut Point—Reading and Writing, Grade 12

	Item #	Rasch Difficulty	Item Information at Intermediate Cut	Item Information at Advanced Cut	Item Information at Proficient Cut
READING	1	0.23	0.17	0.04	0.02
	2	0.43	0.19	0.05	0.03
	3	0.89	0.23	0.08	0.04
	4	1.88	0.24	0.16	0.10
	5	1.09	0.24	0.09	0.05
	6	1.05	0.24	0.09	0.05
	7	3.12	0.14	0.25	0.21
	8	2.87	0.16	0.24	0.19
	9	2.58	0.19	0.22	0.16
	10	1.75	0.25	0.14	0.09
	11	2.99	0.15	0.24	0.20
	12	2.92	0.15	0.24	0.19
	13	1.99	0.23	0.17	0.11
	14	1.23	0.25	0.10	0.06
	15	2.16	0.22	0.18	0.12
	16	3.31	0.12	0.25	0.22
	17	2.23	0.22	0.19	0.13
	18	1.13	0.24	0.09	0.05
	19	3.08	0.14	0.25	0.21
	20	2.88	0.16	0.24	0.19
	21	2.88	0.16	0.24	0.19
	22	2.38	0.21	0.20	0.14
	23	2.70	0.18	0.23	0.17
	24	2.32	0.21	0.20	0.14
	25	2.46	0.20	0.21	0.15
	26	1.94	0.24	0.16	0.10
	27	1.66	0.25	0.14	0.08
WRITING	1	0.01	0.15	0.03	0.02
	2	0.34	0.18	0.05	0.03
	3	1.31	0.25	0.11	0.06
	4	2.80	0.17	0.24	0.18
	5	1.99	0.23	0.17	0.11
	6	2.52	0.19	0.22	0.15
	7	1.91	0.24	0.16	0.10
	8	1.56	0.25	0.13	0.08
	9	2.59	0.19	0.22	0.16
	10	1.04	0.24	0.09	0.05
	11	2.22	0.22	0.19	0.13
	12	1.42	0.25	0.11	0.07
	13	3.05	0.14	0.25	0.20
	14	1.83	0.24	0.15	0.09
	15	1.99	0.23	0.17	0.11
	16	1.40	0.25	0.11	0.07
	17	1.43	0.25	0.12	0.07
	18	1.83	0.24	0.15	0.09
	19	2.55	0.19	0.22	0.16

APPENDIX F: CONFIRMATORY FACTOR ANALYSIS (CFA)

F.1: Introduction

Under federal Title III of NCLB, New York state regulations require annual assessment of limited English proficient (LEP) students using 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 principles and consistency of Universal Design and applicable federal and state testing requirements. To meet these requirements, the test was developed for five grade bands (K–1, 2–4, 5–6, 7–8, 9–12) in four modalities (Speaking, Listening, Reading, and Writing) to assess the English language proficiency of ELL students from Grades K–12. The maximum number of points by modality, grade band, and item type—i.e., MC or CR—for the 2012 administration is depicted in the following table.

Table F.1: Maximum Number of Points by Modality and Grade Band for the 2012 NYSESLAT

Maximum Number of Points in NYSESLAT Subtests						
	Speaking	Listening	Reading	Writing		Total Number of Points
				Writing conventions		
Grade Band	CR	MC	MC	MC	CR	
K–1	34	24	15	6	18	97
2–4	34	24	24	12	10	104
5–6	34	25	27	15	10	111
7–8	34	25	27	15	10	111
9–12	34	25	27	15	10	111

F.2: Reporting of the NYSESLAT Scores

The NYSESLAT scores are reported on two different scales: one is a combination of Speaking and Listening modalities, and the other one is a combination of the Reading and Writing modalities. In 2009, NYSED 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. This 2012 report follows the same procedure to keep the consistency of the data construct. In order to test the assumption of unidimensionality and compare it with the currently used two-dimensional model, a CFA was used to examine the factor loadings based on the different combinations of the observed variables.

Because the primary interest of this study was to compare the current reporting latent structure with an alternative preconceived unidimensional structure, a confirmatory factor model was chosen for analysis from the various other factor models available. This study, therefore, tested the following two latent structures:

- A one-factor model with the latent structure being ELPA

- A two-factor model that tested the current latent structure of Speaking and Listening as one factor (SPKLIS), and Reading and Writing (Writing conventions + Writing constructed-response items) as the other factor (WRTRD)

The results of these confirmatory factor analyses are based on the 2012 data set and the full population response. Because 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 four scores for each individual students. Therefore, the factor loadings for the one- and two-factor models were based on the Speaking, Listening, Reading, and Writing modalities. 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: 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 χ^2 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. 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, MacCallum, Browne, and Sugawara (1996) suggest RMSEA ≤ 0.05 as close model fit, with 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. Figures F1 and F2 show the path diagrams for the one- and two-factor models.

Figure F1: Path Diagram for the One-Factor Model

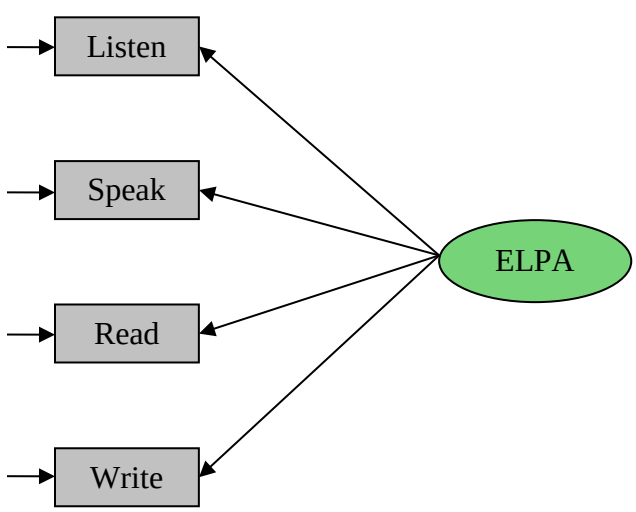
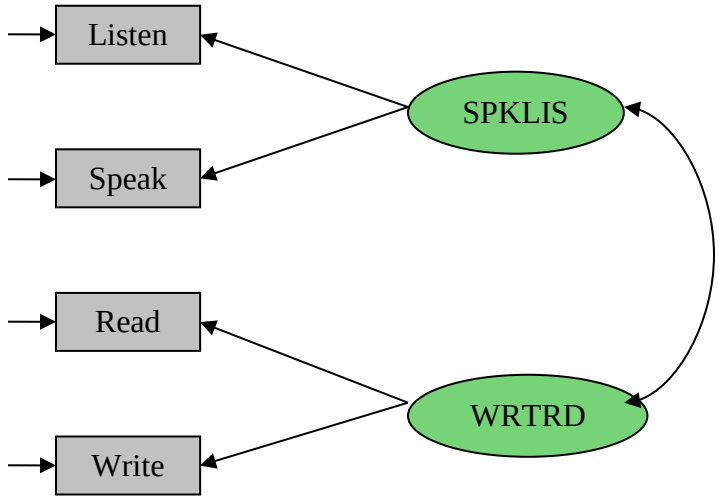


Figure F2—Path Diagram for the Two-Factor Model



In order to estimate factor loadings, the PROC CALIS procedure implemented in SAS was used. All the reported factor loadings are standardized.

F.4: Result and Comments

Grades K–1

Table F.2 shows the summary statistics, Table F.3 shows the standardized factor loadings for the one- and two-factor models, and Table F.4 shows the fit indices for the two models.

Table F.2: Summary Statistics, Grades K–1

Modality	N-Count	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
Listening	51,942	21.102	3.671	-1.933	4.180	0	24
Speaking	51,942	26.796	7.316	-1.776	3.332	0	34
Reading	51,942	11.760	3.483	-0.898	-0.243	0	15
Writing	51,942	18.176	5.348	-0.946	0.187	0	24

Table F.3: Factor Loading Estimates for the One- and Two-Factor Models, Grades K–1

Modality	One-Factor model	Two-Factor Model	
	Λ	Λ_1	Λ_2
Listening	0.716	0.883	--
Speaking	0.658	0.799	--
Reading	0.862	--	0.862
Writing	0.910	--	0.937
		correlation	0.765

Table F.4: Global Fit Indices for the One- and Two-Factor Models, Grades K–1

Fit Indices	One-Factor model	Two-Factor Model
χ^2	13955.411	209.103
GFI	0.884	0.998
AGFI	0.421	0.980
RMSR	0.079	0.005
RMSEA	0.367	0.063

Note: χ^2 with 2 *df*, p <.001 for the one-factor model and χ^2 with 1 *df*, p<.001 for the two-factor model

For the one-factor model, the value of χ^2 is 13955.411(*df*=2, *p*<.001), which is considerably large and significant with 2 *df*. It might be because the data has a large sample size, but it still shows that the one-factor model does not represent the data structure well. Other model-fit indices such as GFI (0.884), AGFI (0.421), RMSR (0.079), and RMSEA (0.367) also show that the one-factor model does not represent data set well and suggest a revision of the model for a better fit.

For the two-factor model keeping with the current dimensionality of test, the value of χ^2 is dramatically decreased from 13955.411(*df*=2, *p*<.001) to 209.103(*df*=1, *p*<.001). Although the p-value of the χ^2 is still in the range of significance at the 0.05 level of rejection, the model-fit significantly improved. All other model-fit indices such as GFI (0.998), AGFI (0.980), RMSR (0.005), and RMSEA (0.063) are in the range of satisfactory model-fit.

Grades 2–4

Table F.5 shows the summary statistics, Table F.6 shows the standardized factor loadings for the one- and two-factor models, and Table F.7 shows the fit indices for the two models.

Table F.5: Summary Statistics, Grades 2–4

Modality	N-Count	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
Listening	58,036	19.476	4.073	-1.493	2.299	0	24
Speaking	58,036	29.199	6.913	-2.499	6.592	0	34
Reading	58,036	17.058	5.052	-0.683	-0.345	0	24
Writing	58,036	15.727	4.641	-1.139	0.783	0	22

Table F.6: Factor Loading Estimates for the One- and Two-Factor Models, Grades 2–4

Modality	One-Factor model	Two-Factor Model	
	Λ	Λ_1	Λ_2
Listening	0.828	0.889	--
Speaking	0.732	0.776	--
Reading	0.844	--	0.850
Writing	0.912	--	0.932
		correlation	0.895

Table F.7: Global Fit Indices for the One- and Two-Factor Models, Grades 2–4

Fit Indices	One-Factor model	Two-Factor Model
χ^2	5907.201	1494.695
GFI	0.953	0.987
AGFI	0.762	0.875
RMSR	0.035	0.014
RMSEA	0.226	0.160

Note: χ^2 with 2 df, p <.001 for the one-factor model and χ^2 with 1 df, p<.001 for the two-factor model

For the one-factor model, the value of χ^2 is 5907.201(*df*=2, *p*<.001), which is extremely large and shows that the model fits well in data structures. The values of GFI and AGFI are 0.953 and 0.762, respectively. While GFI index is close to 1, AGFI is in poor range of model-fit. The value of RMSR is in an acceptable range (0. 035), but the value of RMSEA is in poor range (0.226). All model-fit indices suggest that an alternative model is justifiable.

With the two-factor model, the value of χ^2 is dramatically decreased from 5907.201 (*df*=2, *p*<.001) to 1494.695(*df*=1, *p*<.001). Although the p-value of the χ^2 is still in the range of significance at the 0.05 level of rejection, the model-fit significantly improved. Both GFI (0.987) and AGFI (0.875) are slightly increased when compared to the one-factor model, whereas RMSR (0.014) and RMSEA (0.160) are slightly decreased relative to one-factor model. Even though the two-factor model does not fit into the data set, it shows the better model fit when compared to one-factor model.

Grades 5–6

Table F.8 shows the summary statistics, Table F.9 shows the standardized factor loadings for the one- and two-factor models, and Table F.10 shows fit indices for the two models.

Table F.8: Summary Statistics, Grades 5–6

Modality	N-Count	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
Listening	28,263	19.641	5.176	-1.411	1.240	0	25
Speaking	28,263	29.060	7.369	-2.270	5.005	0	34
Reading	28,263	18.707	6.252	-0.666	-0.649	0	27
Writing	28,263	18.123	5.282	-1.213	0.820	0	25

Table F.9: Factor Loading Estimates for the One- and Two-Factor Models, Grades 5–6

Modality	One-Factor model	Two-Factor Model	
	Λ	Λ_1	Λ_2
Listening	0.851	0.888	-
Speaking	0.783	0.814	-
Reading	0.826	-	0.830
Writing	0.917	-	0.941
		correlation	0.920

Table F.10: Global Fit Indices for the One- and Two-Factor Models, Grades 5–6

Fit Indices	One-Factor model	Two-Factor Model
χ^2	2796.791	1184.526
GFI	0.956	0.980
AGFI	0.781	0.803
RMSR	0.030	0.018
RMSEA	0.221	0.203

Note: χ^2 with 2 df, p <.001 for the one-factor model and χ^2 with 1 df, p<.001 for the two-factor model

Two-factor model is a slightly better fit to the data when compared to the one-factor model. The values of fit indices for the one- and two-factor models show that there is small improvement in model-fit. Although the p-value of the χ^2 is still in the range of significance at the 0.05 level of rejection, it is decreased from 2796.791(*df*=2, *p*<.001) to 1184.526 (*df*=1, *p*<.001). Both GFI and AGFI also show that the two-factor model is slightly better model-fit to the data. RMSR is in the satisfactory range for both the one- and two-factor model (0.030 and 0.018, respectively). The values of RMSEA for both the one- and two-factor model (0.221 and 0.203, respectively) also show that the two-factor model is slightly better but not superior to the one-factor model.

Grades 7–8

Table F.11 shows the summary statistics, Table F.12 shows the standardized factor loadings for the one- and two-factor models, and Table F.13 shows the fit indices for the two models.

Table F.11: Summary Statistics, Grades 7–8

Modality	N-Count	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
Listening	25,115	18.120	4.997	-0.877	0.053	0	25
Speaking	25,115	27.691	8.483	-1.733	2.317	0	34
Reading	25,115	17.041	5.959	-0.338	-0.918	0	27
Writing	25,115	17.265	5.515	-0.952	0.085	0	25

Table F.12: Factor Loading Estimates for the One- and Two-Factor Models, Grades 7–8

Modality	One-Factor model	Two-Factor Model	
	Λ	Λ_1	Λ_2
Listening	0.833	0.844	--
Speaking	0.781	0.792	--
Reading	0.789	--	0.790
Writing	0.912	--	0.922
		correlation	0.971

Table F.13: Global Fit Indices for the One- and Two-Factor Models, Grades 7–8

Fit Indices	One-Factor model	Two-Factor Model
χ^2	1649.991	1503.533
GFI	0.971	0.972
AGFI	0.856	0.718
RMSR	0.028	0.025
RMSEA	0.181	0.245

Note: χ^2 with 2 df, p <.001 for the one-factor model and χ^2 with 1 df, p<.001 for the two-factor model

Grades 7–8 do not have much improvement using the two-factor model over the one-factor model to fit the data set. The difference of fit indices between the one- and two-factor models is not significantly large. AGFI decreases from the one-factor model to the two-factor model (0.856 and 0.718, respectively). The amount of change in fit indices can be from the estimation error. Overall, there is no significant improvement in model fitting when the one- and two-factor models are used to fit the data set.

Grades 9–12

Table F.14 shows the summary statistics, Table F.15 shows the standardized factor loadings for the one- and two-factor models, and Table F.16 shows the fit indices for the two models.

Table F.14: Summary Statistics, Grades 9–12

Modality	N-Count	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
Listening	44,704	16.686	5.737	-0.541	-0.731	0	25
Speaking	44,704	26.873	8.487	-1.446	1.408	0	34
Reading	44,704	16.335	5.783	-0.310	-0.782	0	27
Writing	44,704	16.422	6.117	-0.695	-0.492	0	25

Table F.15: Factor Loading Estimates for the One- and Two-Factor Models, Grades 9–12

Modality	One-Factor model	Two-Factor Model	
	Λ	Λ_1	Λ_2
Listening	0.851	0.866	-
Speaking	0.775	0.787	-
Reading	0.820	-	0.823
Writing	0.903	-	0.912
		correlation	0.965

Table F.16: Global Fit Indices for the One- and Two-Factor Models, Grades 9–12

Fit Indices	One-Factor model	Two-Factor Model
χ^2	2266.212	1853.231
GFI	0.977	0.980
AGFI	0.887	0.801
RMSR	0.022	0.019
RMSEA	0.159	0.204

Note: χ^2 with 2 df, p <.001 for the one factor model and χ^2 with 1 df, p<.001 for the two-factor model

Grades 9–12 show the same pattern as Grades 7–8. The two-factor model does not improve model fit over the one-factor model. The value of χ^2 decreases from 2266.212 (*df*=2, *p*<.001) to 1853.231(*df*=1, *p*<.001), which could be in the range of estimation error. The amount of change in fit indices can be from the estimation error. Overall, there is no significant improvement in model fitting when the one- and two-factor models are used to fit the data set.

F.3: Conclusion

Overall, the 2012 NYSESLAT dimensionality tests of the latent structure, 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 Reading and Writing (Writing Convention + Writing open-constructed response) as the other—especially with the lower grade bands. With the high grade bands, even though the two-factor model produced a slightly better model fit than the one-factor model, the difference was not as noticeable as the difference observed with the lower grade bands.

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 RMSR index, in conjunction with acceptable values of other fit indices, then the two-factor model for Grade Band K–1 has the most acceptable fit of the model to the data. Besides having acceptable RMSR and RMSEA under the various criteria discussed above, the GFI and the AGFI were close to 1.0.

In Grade Band 2–4, the two-factor model improves the model fit when compared to the one-factor model in conjunction with acceptable values of other fit indices such as GFI, AGFI, and RMSR. However, the improvement is not as significant as the one in Grade Band K–1.

With respect to the analysis of the 5–6, 7–8, and 9–12 grade band 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 significant 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 toward content and substantive issues in selecting items for the future.

APPENDIX G: SCALE SCORE SUMMARY BY SUBGROUP

Table G.1a: Scale Score Summary by Subgroup—Listening and Speaking, Grade K

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
K	LS	Female	12,414	599.46	601	55.39	66
K	LS	Male	13,852	594.33	595	55.19	69
K	LS	American Indian or Alaska Native	75	603.81	601	53.22	45
K	LS	Black or African-American	1,002	597.32	601	58.74	66
K	LS	Asian	5,691	587.85	585	57.89	67
K	LS	Native Hawaiian/Other Pacific Islander	129	590.67	590	59.93	60
K	LS	Hispanic or Latino	16,716	598.37	595	52.69	62
K	LS	White	2,582	605.63	601	61.91	69
K	LS	Multiracial (not Hispanic origin)	71	603.93	601	59.52	58
K	LS	NYC	15,708	594.08	595	57.29	69
K	LS	Big 4 Cities	1,087	581.21	585	55.39	63
K	LS	High Need Urban/Suburban	3,909	598.21	595	52.99	62
K	LS	High Need Rural	204	599.00	595	45.80	55.5
K	LS	Average Need	3,220	607.31	607	52.53	54
K	LS	Low Need	1,425	603.64	601	46.95	54
K	LS	Charter Schools	713	609.36	607	41.37	45
K	LS	0 Years LEP	15,404	593.92	595	56.99	69
K	LS	1 Years LEP	10,240	600.51	601	52.45	58
K	LS	2 Years LEP	489	609.33	607	52.04	65
K	LS	3 Years LEP	28	616.86	621	69.38	52
K	LS	4 Years LEP	3	502.67	471	71.93	133
K	LS	5 Years LEP	5	505.00	475	77.01	15
K	LS	6 Years LEP or More	0	0.00	0	0.00	0
K	LS	Chinese	2,909	581.06	580	57.09	66
K	LS	English	2,614	608.06	607	49.71	61
K	LS	Haitian-Creole	200	590.88	595	64.47	73
K	LS	Korean	167	604.41	607	57.43	61
K	LS	Russian	541	616.32	614	65.81	78
K	LS	Spanish	15,321	597.22	595	52.86	66
K	LS	Other Language	4,514	596.38	595	60.86	69
K	LS	Students without Disabilities	23,235	599.15	601	55.00	62
K	LS	Students with Disabilities	3,031	578.41	576	54.56	59

Table G.1b: Scale Score Summary by Subgroup—Reading and Writing, Grade K

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
K	RW	Female	12,414	555.51	552	54.69	68
K	RW	Male	13,852	545.86	541	54.20	69
K	RW	American Indian or Alaska Native	75	555.77	552	56.15	68
K	RW	Black or African-American	1,002	541.16	536	54.99	67
K	RW	Asian	5,691	571.07	563	62.07	79
K	RW	Native Hawaiian/Other Pacific Islander	129	559.44	552	57.50	73
K	RW	Hispanic or Latino	16,716	542.95	541	49.57	62
K	RW	White	2,582	556.15	552	55.73	68
K	RW	Multiracial (not Hispanic origin)	71	554.06	552	52.18	59
K	RW	NYC	15,708	555.07	552	58.03	73
K	RW	Big 4 Cities	1,087	524.61	521	46.16	57
K	RW	High Need Urban/Suburban	3,909	540.14	536	47.31	62
K	RW	High Need Rural	204	545.82	541	44.26	46
K	RW	Average Need	3,220	545.84	546	46.39	53
K	RW	Low Need	1,425	549.37	546	49.57	60
K	RW	Charter Schools	713	567.62	563	51.50	58
K	RW	0 Years LEP	15,404	554.88	546	57.82	68
K	RW	1 Years LEP	10,240	544.17	541	48.91	58
K	RW	2 Years LEP	489	546.23	541	51.98	53
K	RW	3 Years LEP	28	547.21	557	52.06	81.5
K	RW	4 Years LEP	3	464.67	471	15.50	29
K	RW	5 Years LEP	5	477.00	466	34.61	16
K	RW	6 Years LEP or More	0	0.00	0	0.00	0
K	RW	Chinese	2,909	577.86	570	62.03	82
K	RW	English	2,614	548.03	546	48.31	60
K	RW	Haitian-Creole	200	535.40	531	56.57	71
K	RW	Korean	167	597.41	594	61.49	84
K	RW	Russian	541	579.18	570	63.84	87
K	RW	Spanish	15,321	543.02	541	49.87	62
K	RW	Other Language	4,514	554.70	552	58.78	73
K	RW	Students without Disabilities	23,235	553.12	546	54.28	68
K	RW	Students with Disabilities	3,031	529.68	526	52.94	62

Table G.2a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 1

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
1	LS	Female	11,986	639.45	641	60.85	65
1	LS	Male	13,690	634.17	630	62.45	71
1	LS	American Indian or Alaska Native	61	635.13	630	59.71	71
1	LS	Black or African-American	1,131	636.92	630	61.26	71
1	LS	Asian	5,252	633.69	630	65.09	77
1	LS	Native Hawaiian/Other Pacific Islander	113	633.42	630	62.99	77
1	LS	Hispanic or Latino	16,667	635.61	630	59.18	71
1	LS	White	2,417	650.36	654	69.88	93
1	LS	Multiracial (not Hispanic origin)	35	621.43	621	55.57	51
1	LS	NYC	16,153	634.21	630	63.24	77
1	LS	Big 4 Cities	1,198	618.28	621	61.74	69
1	LS	High Need Urban/Suburban	3,360	635.59	630	59.55	71
1	LS	High Need Rural	193	645.13	641	60.45	65
1	LS	Average Need	2,918	651.67	641	57.56	86
1	LS	Low Need	1,255	647.88	641	54.90	58
1	LS	Charter Schools	599	644.90	641	47.42	58
1	LS	0 Years LEP	14,263	632.12	630	63.49	77
1	LS	1 Years LEP	3,368	627.52	630	68.03	82
1	LS	2 Years LEP	7,012	647.82	641	52.97	58
1	LS	3 Years LEP	948	656.14	654	54.03	86
1	LS	4 Years LEP	27	645.74	641	62.15	86
1	LS	5 Years LEP	6	655.67	752	149.95	266
1	LS	6 Years LEP or More	9	522.11	486	103.10	32
1	LS	Chinese	2,673	629.30	630	62.68	82
1	LS	English	1,983	650.08	641	55.08	58
1	LS	Haitian-Creole	235	636.74	630	67.91	105
1	LS	Korean	140	650.94	641	66.21	82.5
1	LS	Russian	546	656.96	654	75.69	86
1	LS	Spanish	15,651	634.80	630	59.41	71
1	LS	Other Language	4,448	638.54	641	67.98	77
1	LS	Students without Disabilities	22,237	639.50	641	62.01	65
1	LS	Students with Disabilities	3,439	618.05	621	56.74	69

Table G.2b: Scale Score Summary by Subgroup—Reading and Writing, Grade 1

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
1	RW	Female	11,986	634.43	636	66.00	131
1	RW	Male	13,690	620.34	618	68.60	87
1	RW	American Indian or Alaska Native	61	642.93	636	63.05	111
1	RW	Black or African-American	1,131	620.15	618	68.22	87
1	RW	Asian	5,252	650.21	664	65.66	111
1	RW	Native Hawaiian/Other Pacific Islander	113	637.37	636	68.17	131
1	RW	Hispanic or Latino	16,667	618.99	618	66.44	87
1	RW	White	2,417	633.31	636	68.63	122
1	RW	Multiracial (not Hispanic origin)	35	620.94	636	78.43	94
1	RW	NYC	16,153	627.39	636	69.31	87
1	RW	Big 4 Cities	1,198	594.86	585	70.02	95
1	RW	High Need Urban/Suburban	3,360	618.42	618	65.40	87
1	RW	High Need Rural	193	622.15	618	59.92	79
1	RW	Average Need	2,918	635.71	636	61.36	70
1	RW	Low Need	1,255	645.61	636	58.97	111
1	RW	Charter Schools	599	645.45	636	53.26	111
1	RW	0 Years LEP	14,263	625.95	636	69.39	87
1	RW	1 Years LEP	3,368	617.29	618	72.69	101
1	RW	2 Years LEP	7,012	632.59	636	61.59	79
1	RW	3 Years LEP	948	636.32	636	60.42	122
1	RW	4 Years LEP	27	636.07	664	57.08	59
1	RW	5 Years LEP	6	559.83	611.5	169.49	325
1	RW	6 Years LEP or More	9	498.33	481	68.75	39
1	RW	Chinese	2,673	656.35	664	61.56	98
1	RW	English	1,983	633.34	636	61.57	70
1	RW	Haitian-Creole	235	609.29	605	69.99	107
1	RW	Korean	140	670.18	664	56.99	80
1	RW	Russian	546	643.62	664	67.80	111
1	RW	Spanish	15,651	618.37	618	66.81	94
1	RW	Other Language	4,448	633.94	636	70.42	131
1	RW	Students without Disabilities	22,237	631.45	636	66.31	131
1	RW	Students with Disabilities	3,439	597.63	594	69.74	90

Table G.3a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 2

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
2	LS	Female	9,497	665.19	666	48.05	46
2	LS	Male	11,662	660.62	661	48.25	50
2	LS	American Indian or Alaska Native	40	654.38	651	51.54	48
2	LS	Black or African-American	1,066	663.00	666	50.10	50
2	LS	Asian	4,252	658.76	663.5	52.40	54
2	LS	Native Hawaiian/Other Pacific Islander	95	669.22	666	51.59	55
2	LS	Hispanic or Latino	13,943	663.34	666	45.71	46
2	LS	White	1,737	666.50	673	54.62	59
2	LS	Multiracial (not Hispanic origin)	26	663.46	667	53.90	63
2	LS	NYC	13,574	659.92	666	49.39	50
2	LS	Big 4 Cities	1,166	651.25	655	49.59	57
2	LS	High Need Urban/Suburban	2,659	663.06	666	43.37	46
2	LS	High Need Rural	145	673.16	673	49.18	51
2	LS	Average Need	2,228	676.07	680	44.99	46
2	LS	Low Need	993	675.37	673	44.22	46
2	LS	Charter Schools	394	676.92	673	35.99	42
2	LS	0 Years LEP	11,201	658.36	661	50.45	54
2	LS	1 Years LEP	2,226	648.85	655	54.48	60
2	LS	2 Years LEP	2,157	668.41	666	43.01	55
2	LS	3 Years LEP	4,706	675.14	673	38.66	46
2	LS	4 Years LEP	774	673.69	673	38.39	46
2	LS	5 Years LEP	41	674.46	666	35.49	37
2	LS	6 Years LEP or More	16	620.75	648.5	60.75	100.5
2	LS	Chinese	2,026	657.26	661	52.82	54
2	LS	English	1,218	674.99	673	41.65	46
2	LS	Haitian-Creole	248	664.69	663.5	52.27	59
2	LS	Korean	108	668.01	673	61.93	63
2	LS	Russian	374	672.57	673	53.71	51
2	LS	Spanish	13,311	662.61	666	46.04	46
2	LS	Other Language	3,874	660.58	666	52.69	54
2	LS	Students without Disabilities	17,227	665.69	666	48.63	55
2	LS	Students with Disabilities	3,932	649.45	651	43.99	47

Table G.3 b: Scale Score Summary by Subgroup—Reading and Writing, Grade 2

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
2	RW	Female	9,497	617.31	617	44.38	58
2	RW	Male	11,662	611.28	613	45.49	60
2	RW	American Indian or Alaska Native	40	610.58	613	54.23	62
2	RW	Black or African-American	1,066	607.57	608	46.97	58
2	RW	Asian	4,252	623.78	627	48.27	60
2	RW	Native Hawaiian/Other Pacific Islander	95	621.14	627	46.29	68
2	RW	Hispanic or Latino	13,943	610.92	613	42.65	54
2	RW	White	1,737	618.15	622	50.36	62
2	RW	Multiracial (not Hispanic origin)	26	619.12	619.5	51.99	80
2	RW	NYC	13,574	611.76	613	45.23	60
2	RW	Big 4 Cities	1,166	595.36	595	46.59	64
2	RW	High Need Urban/Suburban	2,659	615.66	617	42.05	52
2	RW	High Need Rural	145	619.97	622	44.43	58
2	RW	Average Need	2,228	623.61	622	42.18	49
2	RW	Low Need	993	632.99	632	46.68	58
2	RW	Charter Schools	394	629.89	632	34.68	41
2	RW	0 Years LEP	11,201	610.66	613	45.17	60
2	RW	1 Years LEP	2,226	604.73	604	49.14	66
2	RW	2 Years LEP	2,157	619.53	622	46.04	62
2	RW	3 Years LEP	4,706	623.55	627	40.93	49
2	RW	4 Years LEP	774	616.41	617	40.46	52
2	RW	5 Years LEP	41	610.32	613	42.65	58
2	RW	6 Years LEP or More	16	580.56	577	41.08	58.5
2	RW	Chinese	2,026	627.25	632	47.08	55
2	RW	English	1,218	624.39	627	43.63	55
2	RW	Haitian-Creole	248	606.24	608	46.02	62
2	RW	Korean	108	648.36	649	51.69	53
2	RW	Russian	374	630.48	632	50.87	67
2	RW	Spanish	13,311	610.24	613	42.72	54
2	RW	Other Language	3,874	614.61	617	48.83	66
2	RW	Students without Disabilities	17,227	618.81	622	44.10	58
2	RW	Students with Disabilities	3,932	592.84	595	43.26	55

Table G.4a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 3

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
3	LS	Female	8,643	683.16	688	55.28	54
3	LS	Male	10,381	677.61	680	53.54	58
3	LS	American Indian or Alaska Native	32	663.09	676.5	82.25	93
3	LS	Black or African-American	1,132	679.98	680	54.87	54
3	LS	Asian	3,549	674.51	680	57.88	67
3	LS	Native Hawaiian/Other Pacific Islander	66	671.29	661	60.53	99
3	LS	Hispanic or Latino	12,747	681.40	688	52.47	54
3	LS	White	1,482	683.42	688	59.74	70
3	LS	Multiracial (not Hispanic origin)	16	695.50	697	49.12	68.5
3	LS	NYC	12,386	676.96	680	55.64	58
3	LS	Big 4 Cities	1,104	666.23	673	50.75	59
3	LS	High Need Urban/Suburban	2,416	684.71	688	50.76	48
3	LS	High Need Rural	128	689.46	697	55.90	59
3	LS	Average Need	1,900	695.32	697	50.30	59
3	LS	Low Need	793	691.28	688	53.35	59
3	LS	Charter Schools	297	695.89	688	38.99	52
3	LS	0 Years LEP	10,189	675.83	680	56.77	58
3	LS	1 Years LEP	1,815	653.23	655	62.73	72
3	LS	2 Years LEP	1,516	686.36	688	47.43	54
3	LS	3 Years LEP	1,544	692.96	688	43.90	59
3	LS	4 Years LEP	3,259	697.21	697	41.53	52
3	LS	5 Years LEP	625	694.97	688	40.82	36
3	LS	6 Years LEP or More	57	665.74	680	64.44	63
3	LS	Chinese	1,607	672.18	680	58.86	63
3	LS	English	919	694.77	697	46.76	59
3	LS	Haitian-Creole	283	681.92	688	54.30	54
3	LS	Korean	67	675.91	680	62.86	79
3	LS	Russian	289	686.84	697	63.49	70
3	LS	Spanish	12,297	680.75	688	52.69	54
3	LS	Other Language	3,562	677.21	680	58.07	63
3	LS	Students without Disabilities	15,007	682.91	688	55.75	54
3	LS	Students with Disabilities	4,017	669.75	673	47.69	55

Table G.4b: Scale Score Summary by Subgroup—Reading and Writing, Grade 3

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
3	RW	Female	8,643	643.05	649	49.13	67
3	RW	Male	10,381	636.47	643	49.99	66
3	RW	American Indian or Alaska Native	32	627.09	622.5	65.92	97
3	RW	Black or African-American	1,132	631.52	634.5	50.86	70
3	RW	Asian	3,549	646.11	655	54.45	67
3	RW	Native Hawaiian/Other Pacific Islander	66	637.47	637	60.17	93
3	RW	Hispanic or Latino	12,747	638.24	643	47.34	62
3	RW	White	1,482	640.26	649	54.02	76
3	RW	Multiracial (not Hispanic origin)	16	659.06	655	47.31	54
3	RW	NYC	12,386	637.18	643	50.15	62
3	RW	Big 4 Cities	1,104	622.10	627	50.41	68
3	RW	High Need Urban/Suburban	2,416	642.46	649	46.97	57
3	RW	High Need Rural	128	644.63	649	48.14	65
3	RW	Average Need	1,900	650.01	655	46.21	58
3	RW	Low Need	793	655.86	662	50.50	63
3	RW	Charter Schools	297	661.19	662	35.08	43
3	RW	0 Years LEP	10,189	636.13	643	50.06	66
3	RW	1 Years LEP	1,815	620.27	622	57.22	83
3	RW	2 Years LEP	1,516	645.87	649	49.53	67
3	RW	3 Years LEP	1,544	648.13	649	46.32	58
3	RW	4 Years LEP	3,259	652.98	655	41.34	53
3	RW	5 Years LEP	625	644.39	643	42.01	53
3	RW	6 Years LEP or More	57	618.32	632	55.96	87
3	RW	Chinese	1,607	649.23	655	53.16	58
3	RW	English	919	651.10	655	43.95	53
3	RW	Haitian-Creole	283	628.38	632	52.22	67
3	RW	Korean	67	671.34	670	66.64	72
3	RW	Russian	289	648.76	655	56.05	63
3	RW	Spanish	12,297	637.72	643	47.61	62
3	RW	Other Language	3,562	637.58	643	54.04	66
3	RW	Students without Disabilities	15,007	645.19	649	48.80	63
3	RW	Students with Disabilities	4,017	618.05	622	47.16	62

Table G.5a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 4

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
4	LS	Female	7,981	693.75	697	59.32	59
4	LS	Male	9,872	689.51	697	56.33	64
4	LS	American Indian or Alaska Native	37	679.51	688	71.80	74
4	LS	Black or African-American	1,074	689.67	697	55.62	64
4	LS	Asian	3,156	681.40	688	61.88	79
4	LS	Native Hawaiian/Other Pacific Islander	42	681.62	684	54.26	79
4	LS	Hispanic or Latino	12,180	694.14	697	55.70	59
4	LS	White	1,346	692.12	697	63.85	64
4	LS	Multiracial (not Hispanic origin)	18	695.50	717	68.75	70
4	LS	NYC	11,916	689.33	697	58.57	64
4	LS	Big 4 Cities	1,120	678.80	688	57.56	63
4	LS	High Need Urban/Suburban	2,145	695.03	697	54.88	59
4	LS	High Need Rural	109	685.19	688	57.76	74
4	LS	Average Need	1,674	704.20	709	52.86	52
4	LS	Low Need	682	699.68	697	58.35	52
4	LS	Charter Schools	207	714.04	709	43.97	63
4	LS	0 Years LEP	9,894	690.37	697	59.56	59
4	LS	1 Years LEP	1,839	657.17	661	63.20	81
4	LS	2 Years LEP	1,335	686.27	688	52.88	74
4	LS	3 Years LEP	991	701.10	697	47.94	52
4	LS	4 Years LEP	732	708.11	709	47.95	71
4	LS	5 Years LEP	2,447	711.45	709	42.16	37
4	LS	6 Years LEP or More	594	707.80	709	45.96	45
4	LS	Chinese	1,376	677.97	680	63.06	83
4	LS	English	848	707.46	709	49.67	45
4	LS	Haitian-Creole	296	687.83	697	61.35	74
4	LS	Korean	86	684.53	688	52.21	83
4	LS	Russian	270	695.80	697	63.95	90
4	LS	Spanish	11,689	693.47	697	55.80	59
4	LS	Other Language	3,288	685.70	688	61.55	74
4	LS	Students without Disabilities	13,425	692.33	697	60.51	59
4	LS	Students with Disabilities	4,428	688.59	688	48.20	64

Table G.5b: Scale Score Summary by Subgroup—Reading and Writing, Grade 4

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
4	RW	Female	7,981	658.72	662	52.03	58
4	RW	Male	9,872	653.33	662	52.65	68
4	RW	American Indian or Alaska Native	37	641.46	662	67.28	80
4	RW	Black or African-American	1,074	646.61	655	54.11	63
4	RW	Asian	3,156	658.77	662	57.52	68
4	RW	Native Hawaiian/Other Pacific Islander	42	651.19	658.5	65.22	96
4	RW	Hispanic or Latino	12,180	655.57	662	49.95	63
4	RW	White	1,346	657.74	662	58.37	63
4	RW	Multiracial (not Hispanic origin)	18	671.67	680	66.12	77
4	RW	NYC	11,916	653.88	662	52.66	63
4	RW	Big 4 Cities	1,120	641.74	643	53.92	76
4	RW	High Need Urban/Suburban	2,145	658.47	662	50.69	58
4	RW	High Need Rural	109	653.53	662	51.72	58
4	RW	Average Need	1,674	666.69	670	50.52	53
4	RW	Low Need	682	669.07	670	50.37	61
4	RW	Charter Schools	207	678.82	670	37.82	49
4	RW	0 Years LEP	9,894	654.19	662	51.96	63
4	RW	1 Years LEP	1,839	632.50	632	61.06	83
4	RW	2 Years LEP	1,335	654.68	655	55.34	73
4	RW	3 Years LEP	991	663.16	662	50.80	58
4	RW	4 Years LEP	732	663.61	662	45.32	53
4	RW	5 Years LEP	2,447	672.79	670	42.01	55
4	RW	6 Years LEP or More	594	664.23	670	43.66	53
4	RW	Chinese	1,376	661.13	670	58.28	77
4	RW	English	848	665.87	670	49.05	53
4	RW	Haitian-Creole	296	645.35	649	55.20	82
4	RW	Korean	86	683.67	680	50.10	55
4	RW	Russian	270	666.32	670	59.09	77
4	RW	Spanish	11,689	655.14	662	50.13	63
4	RW	Other Language	3,288	652.31	655	56.96	73
4	RW	Students without Disabilities	13,425	660.86	670	53.03	53
4	RW	Students with Disabilities	4,428	640.22	643	47.36	57

Table G.6a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 5

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
5	LS	Female	6,957	677.90	687	53.78	55
5	LS	Male	8,502	674.32	679	54.22	60
5	LS	American Indian or Alaska Native	28	666.68	679	64.22	82.5
5	LS	Black or African-American	1,033	673.34	679	51.74	64
5	LS	Asian	2,710	664.81	672	55.29	67
5	LS	Native Hawaiian/Other Pacific Islander	53	673.49	679	54.14	67
5	LS	Hispanic or Latino	10,441	679.67	687	53.05	51
5	LS	White	1,178	671.04	679	57.65	71
5	LS	Multiracial (not Hispanic origin)	16	670.81	675.5	51.83	31
5	LS	NYC	10,445	674.39	679	55.07	60
5	LS	Big 4 Cities	1,027	662.75	667	51.40	63
5	LS	High Need Urban/Suburban	1,726	678.88	687	51.45	51
5	LS	High Need Rural	94	683.61	679	55.07	71
5	LS	Average Need	1,311	687.03	696	50.87	56
5	LS	Low Need	565	686.15	687	54.99	62
5	LS	Charter Schools	291	688.11	687	34.30	40
5	LS	0 Years LEP	8,482	676.42	687	56.50	55
5	LS	1 Years LEP	1,879	642.70	647	56.12	73
5	LS	2 Years LEP	1,350	669.61	672	46.38	56
5	LS	3 Years LEP	767	683.54	687	43.42	51
5	LS	4 Years LEP	460	692.11	687	38.63	48
5	LS	5 Years LEP	511	697.46	696	41.32	51
5	LS	6 Years LEP or More	1,982	697.75	696	36.56	51
5	LS	Chinese	1,092	659.56	667	57.03	73
5	LS	English	645	688.25	687	48.78	40
5	LS	Haitian-Creole	280	668.96	672	52.15	67
5	LS	Korean	80	683.96	679	53.63	76
5	LS	Russian	245	674.13	687	58.76	71
5	LS	Spanish	10,129	679.30	687	53.12	51
5	LS	Other Language	2,988	668.42	672	54.83	71
5	LS	Students without Disabilities	11,329	674.24	679	56.70	60
5	LS	Students with Disabilities	4,130	680.57	679	45.69	51

Table G.6b: Scale Score Summary by Subgroup—Reading and Writing, Grade 5

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
5	RW	Female	6,957	686.02	689	48.13	62
5	RW	Male	8,502	677.78	685	48.62	65
5	RW	American Indian or Alaska Native	28	679.96	689	49.13	61.5
5	RW	Black or African-American	1,033	671.04	672	47.01	58
5	RW	Asian	2,710	683.97	689	53.10	72
5	RW	Native Hawaiian/Other Pacific Islander	53	688.94	689	51.05	58
5	RW	Hispanic or Latino	10,441	682.08	689	46.72	58
5	RW	White	1,178	679.41	685	53.40	68
5	RW	Multiracial (not Hispanic origin)	16	685.63	676.5	68.08	110
5	RW	NYC	10,445	680.44	685	48.75	62
5	RW	Big 4 Cities	1,027	663.40	664	45.78	68
5	RW	High Need Urban/Suburban	1,726	685.90	689	47.79	62
5	RW	High Need Rural	94	684.68	687	47.71	55
5	RW	Average Need	1,311	688.06	694	45.59	58
5	RW	Low Need	565	697.82	699	53.14	69
5	RW	Charter Schools	291	694.54	694	37.02	51
5	RW	0 Years LEP	8,482	680.48	685	48.42	62
5	RW	1 Years LEP	1,879	662.24	661	53.89	78
5	RW	2 Years LEP	1,350	682.52	685	48.69	69
5	RW	3 Years LEP	767	690.47	689	47.87	62
5	RW	4 Years LEP	460	692.54	694	43.22	58
5	RW	5 Years LEP	511	693.70	699	44.36	55
5	RW	6 Years LEP or More	1,982	694.36	694	38.99	47
5	RW	Chinese	1,092	686.45	694	54.28	72
5	RW	English	645	688.64	689	46.49	58
5	RW	Haitian-Creole	280	671.40	672	47.78	61
5	RW	Korean	80	711.13	715.5	55.13	63
5	RW	Russian	245	687.13	685	50.80	62
5	RW	Spanish	10,129	681.85	689	46.67	58
5	RW	Other Language	2,988	676.62	680	51.97	71
5	RW	Students without Disabilities	11,329	685.82	694	49.33	62
5	RW	Students with Disabilities	4,130	669.62	672	44.33	58

Table G.7a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 6

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
6	LS	Female	5,643	672.70	679	58.70	67
6	LS	Male	7,521	672.62	679	57.31	67
6	LS	American Indian or Alaska Native	23	654.39	656	43.80	43
6	LS	Black or African-American	992	675.97	679	53.54	64
6	LS	Asian	2,364	660.15	661	56.82	73
6	LS	Native Hawaiian/Other Pacific Islander	57	657.47	661	61.12	93
6	LS	Hispanic or Latino	8,689	676.38	687	57.79	60
6	LS	White	1,030	667.76	672	61.00	74
6	LS	Multiracial (not Hispanic origin)	9	699.22	707	55.29	56
6	LS	NYC	9,048	669.46	679	58.20	71
6	LS	Big 4 Cities	952	665.77	672	55.93	63
6	LS	High Need Urban/Suburban	1,362	677.46	687	57.03	55
6	LS	High Need Rural	96	683.24	687	63.61	71.5
6	LS	Average Need	1,060	689.01	696	57.21	62
6	LS	Low Need	409	686.96	696	56.11	62
6	LS	Charter Schools	237	692.74	687	37.00	35
6	LS	0 Years LEP	6,898	673.39	679	60.65	64
6	LS	1 Years LEP	1,855	639.90	640	56.07	76
6	LS	2 Years LEP	1,359	661.61	661	49.03	54
6	LS	3 Years LEP	751	682.64	679	46.07	51
6	LS	4 Years LEP	331	689.16	687	40.41	40
6	LS	5 Years LEP	276	696.59	696	39.00	51
6	LS	6 Years LEP or More	1,672	703.59	696	41.54	44
6	LS	Chinese	995	653.39	656	59.00	82
6	LS	English	522	687.24	687	55.98	62
6	LS	Haitian-Creole	338	673.48	679	56.69	67
6	LS	Korean	65	669.54	672	51.70	49
6	LS	Russian	235	667.47	672	67.98	90
6	LS	Spanish	8,433	675.95	687	57.78	60
6	LS	Other Language	2,576	666.81	672	55.39	74
6	LS	Students without Disabilities	9,642	668.19	672	60.49	74
6	LS	Students with Disabilities	3,522	684.89	687	48.12	46

Table G.7b: Scale Score Summary by Subgroup—Reading and Writing, Grade 6

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
6	RW	Female	5,643	683.82	689	50.27	69
6	RW	Male	7,521	676.30	680	49.57	68
6	RW	American Indian or Alaska Native	23	660.09	654	42.32	81
6	RW	Black or African-American	992	680.12	685	47.60	62
6	RW	Asian	2,364	681.50	689	54.82	78
6	RW	Native Hawaiian/Other Pacific Islander	57	679.60	676	60.39	78
6	RW	Hispanic or Latino	8,689	679.01	685	47.86	65
6	RW	White	1,030	678.76	680	57.05	71
6	RW	Multiracial (not Hispanic origin)	9	727.78	719	62.66	50
6	RW	NYC	9,048	676.76	680	49.45	68
6	RW	Big 4 Cities	952	670.77	676	49.30	71
6	RW	High Need Urban/Suburban	1,362	683.51	689	49.72	69
6	RW	High Need Rural	96	683.97	689	54.50	73.5
6	RW	Average Need	1,060	693.12	694	51.05	66
6	RW	Low Need	409	700.49	705	53.53	69
6	RW	Charter Schools	237	698.55	699	34.19	39
6	RW	0 Years LEP	6,898	677.12	685	49.02	65
6	RW	1 Years LEP	1,855	663.79	661	55.07	78
6	RW	2 Years LEP	1,359	679.51	680	51.62	68
6	RW	3 Years LEP	751	690.82	694	50.44	58
6	RW	4 Years LEP	331	694.77	694	45.50	51
6	RW	5 Years LEP	276	694.79	694	40.56	51
6	RW	6 Years LEP or More	1,672	696.53	699	40.67	47
6	RW	Chinese	995	679.87	685	54.31	78
6	RW	English	522	692.61	694	49.87	63
6	RW	Haitian-Creole	338	681.53	685	50.97	62
6	RW	Korean	65	706.15	712	59.32	52
6	RW	Russian	235	685.55	685	61.51	78
6	RW	Spanish	8,433	678.57	685	47.84	65
6	RW	Other Language	2,576	678.39	680	53.06	71
6	RW	Students without Disabilities	9,642	682.48	689	51.60	72
6	RW	Students with Disabilities	3,522	671.43	676	44.38	55

Table G.8a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 7

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
7	LS	Female	5,415	691.57	697	51.85	60
7	LS	Male	7,194	692.74	697	51.62	60
7	LS	American Indian or Alaska Native	32	693.63	702	67.84	97.5
7	LS	Black or African-American	1,008	695.24	702	46.70	51.5
7	LS	Asian	2,229	680.74	683	50.11	68
7	LS	Native Hawaiian/Other Pacific Islander	45	684.71	688	51.05	46
7	LS	Hispanic or Latino	8,428	695.10	702	51.94	62
7	LS	White	853	690.90	692	54.47	78
7	LS	Multiracial (not Hispanic origin)	14	685.50	694.5	53.52	63
7	LS	NYC	8,853	690.47	697	52.11	63
7	LS	Big 4 Cities	898	686.97	692	46.37	53
7	LS	High Need Urban/Suburban	1,300	693.40	702	51.92	53
7	LS	High Need Rural	78	697.15	699.5	43.82	50
7	LS	Average Need	905	702.38	708	53.42	55
7	LS	Low Need	399	704.17	708	50.06	70
7	LS	Charter Schools	176	717.88	715	35.47	34
7	LS	0 Years LEP	6,512	693.63	702	53.42	57
7	LS	1 Years LEP	1,972	664.46	662	53.39	73
7	LS	2 Years LEP	1,324	686.73	683	44.94	59
7	LS	3 Years LEP	823	701.39	702	40.98	46
7	LS	4 Years LEP	292	710.09	708	40.70	43
7	LS	5 Years LEP	224	707.09	708	36.08	43
7	LS	6 Years LEP or More	1,443	717.89	715	36.37	45
7	LS	Chinese	979	672.79	676	50.75	69
7	LS	English	432	708.56	715	49.01	59
7	LS	Haitian-Creole	331	696.13	702	51.47	62
7	LS	Korean	60	698.52	699.5	47.39	49.5
7	LS	Russian	174	685.90	697	59.83	77
7	LS	Spanish	8,210	694.62	702	52.01	53
7	LS	Other Language	2,423	688.87	692	48.79	63
7	LS	Students without Disabilities	9,259	688.66	697	54.07	69
7	LS	Students with Disabilities	3,350	702.13	702	43.07	43

Table G.8b: Scale Score Summary by Subgroup—Reading and Writing, Grade 7

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
7	RW	Female	5,415	667.09	670	42.60	57
7	RW	Male	7,194	660.52	662	42.49	61
7	RW	American Indian or Alaska Native	32	658.75	659	45.59	79.5
7	RW	Black or African-American	1,008	663.76	666	41.64	55
7	RW	Asian	2,229	662.68	666	45.26	61
7	RW	Native Hawaiian/Other Pacific Islander	45	662.27	666	43.88	64
7	RW	Hispanic or Latino	8,428	663.56	666	41.76	55
7	RW	White	853	662.70	666	45.60	66
7	RW	Multiracial (not Hispanic origin)	14	660.64	662	37.44	60
7	RW	NYC	8,853	661.84	666	42.41	58
7	RW	Big 4 Cities	898	654.38	656	38.40	52
7	RW	High Need Urban/Suburban	1,300	665.72	670	43.21	60
7	RW	High Need Rural	78	663.54	666	41.08	52
7	RW	Average Need	905	673.82	677	45.01	58
7	RW	Low Need	399	676.28	682	44.56	59
7	RW	Charter Schools	176	683.88	686	32.60	35
7	RW	0 Years LEP	6,512	662.96	666	42.12	58
7	RW	1 Years LEP	1,972	647.79	643	45.69	63
7	RW	2 Years LEP	1,324	662.68	662	42.71	58
7	RW	3 Years LEP	823	670.23	670	39.07	53
7	RW	4 Years LEP	292	678.58	677	41.20	49
7	RW	5 Years LEP	224	676.13	682	36.92	49
7	RW	6 Years LEP or More	1,443	677.97	677	35.79	45
7	RW	Chinese	979	660.13	662	45.47	64
7	RW	English	432	675.34	677	42.66	52
7	RW	Haitian-Creole	331	664.88	670	43.12	58
7	RW	Korean	60	687.52	691	47.07	47.5
7	RW	Russian	174	663.26	666	45.08	64
7	RW	Spanish	8,210	663.13	666	41.65	58
7	RW	Other Language	2,423	662.41	666	44.01	58
7	RW	Students without Disabilities	9,259	665.04	670	43.94	63
7	RW	Students with Disabilities	3,350	658.66	662	38.51	50

Table G.9a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 8

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
8	LS	Female	5,478	694.13	697	54.18	69
8	LS	Male	7,028	695.43	702	54.21	69
8	LS	American Indian or Alaska Native	38	684.00	677.5	65.09	83
8	LS	Black or African-American	954	698.93	702	51.13	59
8	LS	Asian	2,517	684.64	683	52.06	72
8	LS	Native Hawaiian/Other Pacific Islander	66	693.86	690	52.23	56
8	LS	Hispanic or Latino	7,962	697.92	702	54.38	62
8	LS	White	960	693.08	697	57.26	78
8	LS	Multiracial (not Hispanic origin)	9	663.00	659	58.82	63
8	LS	NYC	9,059	694.14	697	54.06	69
8	LS	Big 4 Cities	869	685.36	692	52.09	69
8	LS	High Need Urban/Suburban	1,134	692.85	702	57.30	72
8	LS	High Need Rural	72	693.39	702	54.02	55
8	LS	Average Need	857	704.85	715	53.30	59
8	LS	Low Need	421	708.90	715	50.58	66
8	LS	Charter Schools	94	724.35	722	38.18	45
8	LS	0 Years LEP	6,427	698.62	708	55.06	62
8	LS	1 Years LEP	2,014	664.03	665	52.66	76
8	LS	2 Years LEP	1,572	686.90	683	49.91	59
8	LS	3 Years LEP	864	700.71	697	44.02	54.5
8	LS	4 Years LEP	324	712.53	715	36.16	43
8	LS	5 Years LEP	217	718.63	715	42.03	50
8	LS	6 Years LEP or More	1,073	727.56	722	39.81	40
8	LS	Chinese	1,177	677.12	676	52.20	73
8	LS	English	389	712.02	715	48.09	63
8	LS	Haitian-Creole	304	701.91	702	50.79	53.5
8	LS	Korean	80	702.39	702	52.25	70
8	LS	Russian	236	691.54	692	59.03	89
8	LS	Spanish	7,797	697.51	702	54.59	66
8	LS	Other Language	2,523	691.54	692	52.65	63
8	LS	Students without Disabilities	9,355	690.36	697	55.70	78
8	LS	Students with Disabilities	3,151	708.22	708	46.99	43

Table G.9b: Scale Score Summary by Subgroup—Reading and Writing, Grade 8

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
8	RW	Female	5,478	671.30	673	43.99	58
8	RW	Male	7,028	665.48	670	43.43	60
8	RW	American Indian or Alaska Native	38	650.97	646	49.49	79
8	RW	Black or African-American	954	669.05	670	43.13	57
8	RW	Asian	2,517	668.68	670	45.37	65
8	RW	Native Hawaiian/Other Pacific Islander	66	667.38	670	41.34	57
8	RW	Hispanic or Latino	7,962	667.71	670	42.90	57
8	RW	White	960	668.70	670	46.97	65
8	RW	Multiracial (not Hispanic origin)	9	660.78	662	54.79	78
8	RW	NYC	9,059	667.41	670	43.36	57
8	RW	Big 4 Cities	869	655.05	656	43.57	59
8	RW	High Need Urban/Suburban	1,134	666.92	670	43.36	65
8	RW	High Need Rural	72	670.19	671.5	45.37	63.5
8	RW	Average Need	857	677.87	682	44.71	56
8	RW	Low Need	421	685.05	686	43.37	55
8	RW	Charter Schools	94	693.18	691	31.02	31
8	RW	0 Years LEP	6,427	668.61	673	43.10	53
8	RW	1 Years LEP	2,014	651.43	649	46.48	68
8	RW	2 Years LEP	1,572	666.00	666	43.90	60
8	RW	3 Years LEP	864	675.82	677	40.92	62
8	RW	4 Years LEP	324	681.97	682	37.36	49
8	RW	5 Years LEP	217	684.49	686	37.33	49
8	RW	6 Years LEP or More	1,073	685.05	686	35.16	42
8	RW	Chinese	1,177	668.24	670	46.18	65
8	RW	English	389	681.95	686	41.84	52
8	RW	Haitian-Creole	304	672.11	671.5	40.85	53
8	RW	Korean	80	692.19	696	50.21	42.5
8	RW	Russian	236	675.87	677	45.60	63.5
8	RW	Spanish	7,797	667.27	670	42.89	57
8	RW	Other Language	2,523	666.15	666	44.89	60
8	RW	Students without Disabilities	9,355	669.06	673	44.57	65
8	RW	Students with Disabilities	3,151	664.95	670	41.16	48

Table G.10a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 9

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
9	LS	Female	6,066	697.16	701	60.68	82
9	LS	Male	8,222	699.03	705	61.74	90
9	LS	American Indian or Alaska Native	43	671.98	660	60.07	82
9	LS	Black or African-American	1,354	700.58	705	52.50	68
9	LS	Asian	3,201	684.84	683	53.97	76
9	LS	Native Hawaiian/Other Pacific Islander	55	685.64	683	57.04	79
9	LS	Hispanic or Latino	8,647	702.42	714	64.49	90
9	LS	White	974	703.24	705	60.00	82
9	LS	Multiracial (not Hispanic origin)	14	731.93	741.5	44.92	67
9	LS	NYC	10,055	698.53	705	61.50	90
9	LS	Big 4 Cities	1,038	700.76	701	61.12	76
9	LS	High Need Urban/Suburban	1,514	688.48	690	62.28	96
9	LS	High Need Rural	75	706.15	714	54.03	76
9	LS	Average Need	994	696.84	701	58.78	85
9	LS	Low Need	462	711.05	716	60.00	82
9	LS	Charter Schools	150	725.29	730	41.88	56
9	LS	0 Years LEP	7,162	703.07	714	64.75	99
9	LS	1 Years LEP	2,806	667.26	665	54.45	76
9	LS	2 Years LEP	1,795	689.15	690	49.88	70
9	LS	3 Years LEP	1,042	710.03	709	48.46	56
9	LS	4 Years LEP	337	722.30	724	47.03	60
9	LS	5 Years LEP	200	728.72	730	44.20	54
9	LS	6 Years LEP or More	909	745.66	744	41.93	46
9	LS	Chinese	1,478	676.00	677	53.57	74
9	LS	English	461	713.05	714	57.04	79
9	LS	Haitian-Creole	450	706.55	709	50.41	62
9	LS	Korean	76	709.64	705	43.50	47
9	LS	Russian	204	710.54	714	61.87	77.5
9	LS	Spanish	8,450	702.00	714	64.59	90
9	LS	Other Language	3,169	694.19	693	54.93	76
9	LS	Students without Disabilities	11,607	691.03	693	61.89	90
9	LS	Students with Disabilities	2,681	729.46	730	47.36	48

Table G.10b: Scale Score Summary by Subgroup—Reading and Writing, Grade 9

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
9	RW	Female	6,066	687.35	690	43.52	67
9	RW	Male	8,222	685.89	687	44.25	70
9	RW	American Indian or Alaska Native	43	667.74	659	38.17	50
9	RW	Black or African-American	1,354	688.60	690	41.87	56
9	RW	Asian	3,201	683.08	680	43.85	65
9	RW	Native Hawaiian/Other Pacific Islander	55	684.53	690	43.09	71
9	RW	Hispanic or Latino	8,647	687.05	690	43.90	67
9	RW	White	974	690.39	690	46.66	64
9	RW	Multiracial (not Hispanic origin)	14	725.36	732	51.60	67
9	RW	NYC	10,055	687.73	690	43.86	67
9	RW	Big 4 Cities	1,038	678.26	677	44.20	60
9	RW	High Need Urban/Suburban	1,514	678.49	677	42.54	67
9	RW	High Need Rural	75	686.03	687	38.70	59
9	RW	Average Need	994	684.91	687	44.29	73
9	RW	Low Need	462	701.42	707	44.80	67
9	RW	Charter Schools	150	707.00	707	31.82	37
9	RW	0 Years LEP	7,162	689.62	693	45.05	64
9	RW	1 Years LEP	2,806	670.05	665	41.78	60
9	RW	2 Years LEP	1,795	682.17	680	41.28	61
9	RW	3 Years LEP	1,042	690.52	693	40.08	53
9	RW	4 Years LEP	337	700.04	704	39.32	50
9	RW	5 Years LEP	200	696.86	697	36.52	48
9	RW	6 Years LEP or More	909	710.09	711	34.73	39
9	RW	Chinese	1,478	682.08	680	44.45	64
9	RW	English	461	696.56	700	41.74	56
9	RW	Haitian-Creole	450	693.40	693	40.10	52
9	RW	Korean	76	707.30	700	40.21	56.5
9	RW	Russian	204	700.65	704	50.17	65.5
9	RW	Spanish	8,450	686.75	690	43.95	70
9	RW	Other Language	3,169	684.08	684	43.56	65
9	RW	Students without Disabilities	11,607	684.70	684	44.47	65
9	RW	Students with Disabilities	2,681	694.32	697	40.69	49

Table G.11a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 10

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
10	LS	Female	5,956	701.06	701	55.34	74
10	LS	Male	7,333	702.89	705	56.24	79
10	LS	American Indian or Alaska Native	38	693.97	699	57.38	76
10	LS	Black or African-American	1,303	708.51	709	45.75	59
10	LS	Asian	3,478	687.51	687	49.12	61
10	LS	Native Hawaiian/Other Pacific Islander	46	692.30	697	40.64	59
10	LS	Hispanic or Latino	7,435	706.42	709	59.10	79
10	LS	White	986	712.93	714	56.13	76
10	LS	Multiracial (not Hispanic origin)	3	695.67	701	16.65	32
10	LS	NYC	9,906	700.26	701	56.73	74
10	LS	Big 4 Cities	745	702.28	697	51.84	71
10	LS	High Need Urban/Suburban	1,156	702.99	705	55.45	79
10	LS	High Need Rural	53	724.45	724	47.44	63
10	LS	Average Need	875	708.29	709	51.65	73
10	LS	Low Need	433	715.76	718	48.15	57
10	LS	Charter Schools	121	736.87	736	45.19	50
10	LS	0 Years LEP	5,830	710.49	714	60.38	82
10	LS	1 Years LEP	2,972	677.72	677	47.77	59
10	LS	2 Years LEP	2,062	691.80	690	47.16	58
10	LS	3 Years LEP	1,179	705.88	705	47.11	65
10	LS	4 Years LEP	361	724.23	724	47.39	60
10	LS	5 Years LEP	227	726.82	730	45.68	52
10	LS	6 Years LEP or More	635	742.81	744	41.65	40
10	LS	Chinese	1,975	676.09	674	47.09	57
10	LS	English	408	709.70	718	50.25	67
10	LS	Haitian-Creole	454	711.89	709	47.55	53
10	LS	Korean	58	708.74	716	57.23	53
10	LS	Russian	238	712.66	711.5	61.64	82
10	LS	Spanish	7,250	706.51	709	59.21	79
10	LS	Other Language	2,906	705.05	705	48.76	62
10	LS	Students without Disabilities	11,363	697.82	697	55.23	76
10	LS	Students with Disabilities	1,926	727.17	730	52.79	63

Table G.11b: Scale Score Summary by Subgroup—Reading and Writing, Grade 10

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
10	RW	Female	5,956	693.06	693	41.77	58
10	RW	Male	7,333	690.13	690	41.40	58
10	RW	American Indian or Alaska Native	38	684.61	688.5	39.35	59
10	RW	Black or African-American	1,303	693.18	693	38.24	52
10	RW	Asian	3,478	688.81	690	41.23	56
10	RW	Native Hawaiian/Other Pacific Islander	46	686.11	684	35.66	36
10	RW	Hispanic or Latino	7,435	691.24	693	41.54	58
10	RW	White	986	700.54	700	46.30	61
10	RW	Multiracial (not Hispanic origin)	3	680.33	677	29.14	58
10	RW	NYC	9,906	690.84	690	41.69	58
10	RW	Big 4 Cities	745	681.15	680	38.09	54
10	RW	High Need Urban/Suburban	1,156	690.22	693	42.39	65
10	RW	High Need Rural	53	703.17	704	45.87	47
10	RW	Average Need	875	696.18	697	39.87	56
10	RW	Low Need	433	709.49	711	40.25	51
10	RW	Charter Schools	121	712.45	711	29.60	32
10	RW	0 Years LEP	5,830	695.09	697	42.59	59
10	RW	1 Years LEP	2,972	680.50	677	40.22	57
10	RW	2 Years LEP	2,062	687.49	687	40.44	56
10	RW	3 Years LEP	1,179	693.18	693	39.08	55
10	RW	4 Years LEP	361	702.00	704	37.39	52
10	RW	5 Years LEP	227	702.20	704	37.07	52
10	RW	6 Years LEP or More	635	708.42	711	36.87	45
10	RW	Chinese	1,975	686.33	687	41.59	59
10	RW	English	408	696.73	700	42.34	56
10	RW	Haitian-Creole	454	695.67	697	38.21	49
10	RW	Korean	58	700.41	711	46.96	49
10	RW	Russian	238	708.16	707	50.58	67
10	RW	Spanish	7,250	691.21	693	41.41	58
10	RW	Other Language	2,906	692.57	693	40.94	55
10	RW	Students without Disabilities	11,363	691.45	690	41.15	58
10	RW	Students with Disabilities	1,926	691.40	697	44.09	59

Table G.12a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 11

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
11	LS	Female	4,730	709.35	709	50.36	67
11	LS	Male	5,045	712.34	714	50.87	64
11	LS	American Indian or Alaska Native	25	702.80	718	43.56	27
11	LS	Black or African-American	936	719.83	718	43.27	51
11	LS	Asian	2,764	703.28	701	43.32	56
11	LS	Native Hawaiian/Other Pacific Islander	31	714.55	697	59.14	88
11	LS	Hispanic or Latino	5,410	711.87	714	54.35	67
11	LS	White	604	723.13	724	52.60	60
11	LS	Multiracial (not Hispanic origin)	5	720.80	730	19.10	27
11	LS	NYC	7,054	709.15	709	50.67	67
11	LS	Big 4 Cities	510	714.10	714	49.57	61
11	LS	High Need Urban/Suburban	935	712.00	714	52.65	67
11	LS	High Need Rural	72	717.03	730	52.32	68
11	LS	Average Need	745	719.69	724	48.04	63
11	LS	Low Need	412	715.06	714	49.74	57
11	LS	Charter Schools	47	730.85	730	40.38	59
11	LS	0 Years LEP	3,608	719.80	724	55.30	66
11	LS	1 Years LEP	1,898	692.51	693	47.44	62
11	LS	2 Years LEP	2,012	700.18	697	43.18	53
11	LS	3 Years LEP	1,157	710.84	709	42.31	53
11	LS	4 Years LEP	329	720.38	718	41.55	51
11	LS	5 Years LEP	221	726.56	730	44.75	52
11	LS	6 Years LEP or More	539	744.86	744	43.56	40
11	LS	Chinese	1,494	693.87	693	42.30	50
11	LS	English	345	720.62	724	50.03	63
11	LS	Haitian-Creole	329	720.79	718	40.49	47
11	LS	Korean	44	720.34	724	42.14	59.5
11	LS	Russian	117	727.97	724	51.50	56
11	LS	Spanish	5,263	711.62	714	54.27	67
11	LS	Other Language	2,183	716.65	714	45.44	57
11	LS	Students without Disabilities	8,506	709.08	709	49.30	67
11	LS	Students with Disabilities	1,269	723.01	730	57.43	67

Table G.12b: Scale Score Summary by Subgroup—Reading and Writing, Grade 11

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
11	RW	Female	4,730	701.84	704	39.82	52
11	RW	Male	5,045	701.45	704	41.39	55
11	RW	American Indian or Alaska Native	25	701.68	707	38.14	37
11	RW	Black or African-American	936	705.14	704	35.89	45
11	RW	Asian	2,764	704.63	707	39.06	49
11	RW	Native Hawaiian/Other Pacific Islander	31	702.29	704	49.22	67
11	RW	Hispanic or Latino	5,410	698.51	700	41.49	53
11	RW	White	604	710.42	711	43.86	54
11	RW	Multiracial (not Hispanic origin)	5	715.80	729	29.85	14
11	RW	NYC	7,054	701.57	704	40.72	52
11	RW	Big 4 Cities	510	696.05	697	36.69	46
11	RW	High Need Urban/Suburban	935	696.25	697	39.50	56
11	RW	High Need Rural	72	704.88	704	43.28	58
11	RW	Average Need	745	707.80	711	40.55	51
11	RW	Low Need	412	708.71	711	43.65	57
11	RW	Charter Schools	47	715.70	715	36.92	54
11	RW	0 Years LEP	3,608	703.40	707	43.28	49
11	RW	1 Years LEP	1,898	696.53	697	39.76	56
11	RW	2 Years LEP	2,012	699.21	697	40.07	53
11	RW	3 Years LEP	1,157	703.82	704	37.45	52
11	RW	4 Years LEP	329	703.25	704	34.79	40
11	RW	5 Years LEP	221	703.69	707	39.49	45
11	RW	6 Years LEP or More	539	711.02	711	34.53	45
11	RW	Chinese	1,494	704.75	704	40.44	49
11	RW	English	345	708.52	711	39.82	48
11	RW	Haitian-Creole	329	706.99	707	33.70	45
11	RW	Korean	44	725.41	726.5	53.67	56.5
11	RW	Russian	117	720.12	720	46.15	54
11	RW	Spanish	5,263	698.17	700	41.44	53
11	RW	Other Language	2,183	704.50	707	38.23	49
11	RW	Students without Disabilities	8,506	703.45	704	38.68	52
11	RW	Students with Disabilities	1,269	689.48	697	50.22	62

Table G.13a: Scale Score Summary by Subgroup—Listening and Speaking, Grade 12

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
12	LS	Female	3,559	717.28	718	49.29	57
12	LS	Male	3,793	718.74	724	51.82	63
12	LS	American Indian or Alaska Native	9	721.00	736	52.88	56
12	LS	Black or African-American	655	722.98	724	42.29	47
12	LS	Asian	2,035	710.80	709	46.06	53
12	LS	Native Hawaiian/Other Pacific Islander	15	727.00	724	36.76	60
12	LS	Hispanic or Latino	4,112	720.08	724	52.76	63
12	LS	White	523	723.60	730	56.91	60
12	LS	Multiracial (not Hispanic origin)	3	721.33	705	76.81	151
12	LS	NYC	5,465	715.86	718	51.80	61
12	LS	Big 4 Cities	382	713.91	714	46.13	57
12	LS	High Need Urban/Suburban	620	723.27	724	48.96	56
12	LS	High Need Rural	25	743.52	744	43.97	70
12	LS	Average Need	522	729.26	730	44.45	48
12	LS	Low Need	288	729.02	724	44.04	48
12	LS	Charter Schools	50	729.08	730	38.50	71
12	LS	0 Years LEP	3,135	721.31	724	56.04	63
12	LS	1 Years LEP	871	706.37	709	44.24	50
12	LS	2 Years LEP	1,240	708.35	709	45.46	56
12	LS	3 Years LEP	1,023	715.27	714	45.76	57
12	LS	4 Years LEP	332	727.53	730	40.52	48
12	LS	5 Years LEP	232	724.53	724	44.97	43
12	LS	6 Years LEP or More	510	738.06	744	44.28	50
12	LS	Chinese	1,102	704.07	701	45.41	56
12	LS	English	224	723.76	724	56.44	67
12	LS	Haitian-Creole	243	722.90	724	44.94	56
12	LS	Korean	67	713.67	718	48.18	54
12	LS	Russian	105	728.81	724	56.18	59
12	LS	Spanish	4,020	720.05	724	52.71	66
12	LS	Other Language	1,591	720.55	724	46.54	47
12	LS	Students without Disabilities	5,891	723.23	724	45.14	60
12	LS	Students with Disabilities	1,461	697.09	705	64.27	87

Table G.13b: Scale Score Summary by Subgroup—Reading and Writing, Grade 12

Grade	Test	Subgroup	N-Count	Mean	Median	SD	IQR
12	RW	Female	3,559	704.31	707	42.42	49
12	RW	Male	3,793	700.76	707	45.83	52
12	RW	American Indian or Alaska Native	9	715.11	720	47.09	47
12	RW	Black or African-American	655	706.02	707	40.33	51
12	RW	Asian	2,035	707.49	707	42.56	51
12	RW	Native Hawaiian/Other Pacific Islander	15	720.27	711	28.19	29
12	RW	Hispanic or Latino	4,112	698.78	704	44.46	55
12	RW	White	523	706.77	715	50.92	57
12	RW	Multiracial (not Hispanic origin)	3	720.33	729	20.43	38
12	RW	NYC	5,465	700.99	707	45.77	52
12	RW	Big 4 Cities	382	696.75	700	38.21	49
12	RW	High Need Urban/Suburban	620	701.22	705.5	38.98	49
12	RW	High Need Rural	25	720.92	711	34.84	44
12	RW	Average Need	522	712.54	715	36.92	42
12	RW	Low Need	288	720.46	720	40.72	50
12	RW	Charter Schools	50	707.04	707	31.83	51
12	RW	0 Years LEP	3,135	696.52	704	50.25	58
12	RW	1 Years LEP	871	707.77	707	38.67	51
12	RW	2 Years LEP	1,240	707.00	707	40.15	45
12	RW	3 Years LEP	1,023	704.60	704	37.57	49
12	RW	4 Years LEP	332	710.55	711	35.31	45
12	RW	5 Years LEP	232	705.31	707	39.03	45
12	RW	6 Years LEP or More	510	708.34	711	37.95	48
12	RW	Chinese	1,102	709.79	711	41.84	48
12	RW	English	224	708.03	715	47.73	51
12	RW	Haitian-Creole	243	708.08	711	41.75	57
12	RW	Korean	67	709.61	720	58.22	60
12	RW	Russian	105	716.74	724	45.91	44
12	RW	Spanish	4,020	698.57	704	44.17	55
12	RW	Other Language	1,591	704.41	707	44.05	45
12	RW	Students without Disabilities	5,891	711.70	711	34.96	45
12	RW	Students with Disabilities	1,461	665.31	668	56.57	80

APPENDIX H: PROFICIENCY PERCENTAGES BY SUBGROUP

Table H.1a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade K

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
K	LS	Female	12,414	5	25	35	35
K	LS	Male	13,852	6	28	36	31
K	LS	American Indian or Alaska Native	75	4	20	45	31
K	LS	Black or African-American	1,002	7	24	35	34
K	LS	Asian	5,691	8	31	33	28
K	LS	Native Hawaiian/Other Pacific Islander	75	4	20	45	31
K	LS	Hispanic or Latino	16,716	5	26	37	33
K	LS	White	2,582	6	22	32	40
K	LS	Multiracial (not Hispanic origin)	71	4	21	42	32
K	LS	NYC	15,708	7	29	34	31
K	LS	Big 4 Cities	1,087	11	31	35	23
K	LS	High Need Urban/Suburban	3,909	5	26	36	33
K	LS	High Need Rural	204	1	25	42	31
K	LS	Average Need	3,220	3	20	40	38
K	LS	Low Need	1,425	3	20	41	36
K	LS	Charter Schools	713	1	19	41	40
K	LS	0 Years LEP	15,404	7	28	34	31
K	LS	1 Years LEP	10,240	4	24	38	34
K	LS	2 Years LEP	489	3	19	38	40
K	LS	3 Years LEP	28	7	11	29	54
K	LS	4 Years LEP	3	67	0	33	0
K	LS	5 Years LEP	5	80	0	0	20
K	LS	6 Years LEP or More	0	--	--	--	--
K	LS	Chinese	2,909	9	36	31	24
K	LS	English	2,614	2	19	40	39
K	LS	Haitian-Creole	200	12	25	30	34
K	LS	Korean	167	8	13	38	41
K	LS	Russian	541	6	18	28	48
K	LS	Spanish	15,321	5	27	37	32
K	LS	Other Language	4,514	8	25	33	34
K	LS	Students without Disabilities	23,235	5	25	36	34
K	LS	Students with Disabilities	3,031	9	37	34	20

Table H.1b: Proficiency Percentages by Subgroup—Reading and Writing, Grade K

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
K	RW	Female	12,414	40	32	13	14
K	RW	Male	13,852	48	29	11	11
K	RW	American Indian or Alaska Native	75	41	28	17	13
K	RW	Black or African-American	1,002	52	28	10	10
K	RW	Asian	5,691	31	30	16	23
K	RW	Native Hawaiian/Other Pacific Islander	75	41	28	17	13
K	RW	Hispanic or Latino	16,716	50	31	11	9
K	RW	White	2,582	39	33	13	14
K	RW	Multiracial (not Hispanic origin)	71	41	32	18	8
K	RW	NYC	15,708	42	30	13	15
K	RW	Big 4 Cities	1,087	66	23	7	4
K	RW	High Need Urban/Suburban	3,909	52	31	10	7
K	RW	High Need Rural	204	46	37	10	6
K	RW	Average Need	3,220	45	36	12	8
K	RW	Low Need	1,425	45	33	12	11
K	RW	Charter Schools	713	29	37	17	17
K	RW	0 Years LEP	15,404	42	30	13	15
K	RW	1 Years LEP	10,240	48	32	11	9
K	RW	2 Years LEP	489	46	36	9	9
K	RW	3 Years LEP	28	43	25	21	11
K	RW	4 Years LEP	3	100	0	0	0
K	RW	5 Years LEP	5	100	0	0	0
K	RW	6 Years LEP or More	0	--	--	--	--
K	RW	Chinese	2,909	27	30	17	26
K	RW	English	2,614	44	36	11	9
K	RW	Haitian-Creole	200	59	23	11	8
K	RW	Korean	167	19	18	20	43
K	RW	Russian	541	28	27	18	27
K	RW	Spanish	15,321	50	31	11	9
K	RW	Other Language	4,514	41	30	13	16
K	RW	Students without Disabilities	23,235	42	32	13	13
K	RW	Students with Disabilities	3,031	61	25	8	7

Table H.1c: Proficiency Percentages by Subgroup—Grade K Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
K	Overall	Female	12,414	41	35	15	9
K	Overall	Male	13,852	49	31	13	7
K	Overall	American Indian or Alaska Native	75	43	28	21	8
K	Overall	Black or African-American	1,002	52	29	12	7
K	Overall	Asian	5,691	32	35	20	13
K	Overall	Native Hawaiian/Other Pacific Islander	129	42	26	19	12
K	Overall	Hispanic or Latino	16,716	50	32	12	6
K	Overall	White	2,582	40	34	15	11
K	Overall	Multiracial (not Hispanic origin)	71	41	34	17	8
K	Overall	NYC	15,708	42	33	15	10
K	Overall	Big 4 Cities	1,087	66	24	8	2
K	Overall	High Need Urban/Suburban	3,909	52	32	11	5
K	Overall	High Need Rural	204	46	38	12	4
K	Overall	Average Need	3,220	45	37	12	6
K	Overall	Low Need	1,425	45	34	14	6
K	Overall	Charter Schools	713	29	39	21	12
K	Overall	0 Years LEP	15,404	42	33	15	10
K	Overall	1 Years LEP	10,240	48	33	12	6
K	Overall	2 Years LEP	489	46	37	10	8
K	Overall	3 Years LEP	28	43	25	21	11
K	Overall	4 Years LEP	3	100	0	0	0
K	Overall	5 Years LEP	5	100	0	0	0
K	Overall	6 Years LEP or More	0	NA	NA	NA	NA
K	Overall	Chinese	2,909	28	37	22	13
K	Overall	English	2,614	44	37	13	6
K	Overall	Haitian-Creole	200	60	22	12	6
K	Overall	Korean	167	20	22	34	24
K	Overall	Russian	541	29	28	21	21
K	Overall	Spanish	15,321	50	32	12	6
K	Overall	Other Language	4,514	41	33	16	10
K	Overall	Students without Disabilities	23,235	43	34	15	9
K	Overall	Students with Disabilities	3,031	61	27	8	4

Table H.2a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 1

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
1	LS	Female	11,986	3	11	43	43
1	LS	Male	13,690	4	13	43	40
1	LS	American Indian or Alaska Native	61	3	11	48	38
1	LS	Black or African-American	1,131	3	12	44	42
1	LS	Asian	5,252	5	15	40	41
1	LS	Native Hawaiian/Other Pacific Islander	113	4	17	38	41
1	LS	Hispanic or Latino	16,667	3	12	46	40
1	LS	White	2,417	4	11	33	52
1	LS	Multiracial (not Hispanic origin)	35	3	17	57	23
1	LS	NYC	16,153	4	13	42	40
1	LS	Big 4 Cities	1,198	6	18	46	30
1	LS	High Need Urban/Suburban	3,360	3	12	47	38
1	LS	High Need Rural	193	2	10	39	49
1	LS	Average Need	2,918	2	7	42	49
1	LS	Low Need	1,255	1	8	45	46
1	LS	Charter Schools	599	0	5	53	42
1	LS	0 Years LEP	14,263	5	14	42	39
1	LS	1 Years LEP	3,368	6	16	41	37
1	LS	2 Years LEP	7,012	0	8	46	46
1	LS	3 Years LEP	948	0	6	42	52
1	LS	4 Years LEP	27	7	7	44	41
1	LS	5 Years LEP	6	33	0	0	67
1	LS	6 Years LEP or More	9	78	0	11	11
1	LS	Chinese	2,673	5	15	43	37
1	LS	English	1,983	1	7	44	48
1	LS	Haitian-Creole	235	5	12	41	41
1	LS	Korean	140	4	8	41	48
1	LS	Russian	546	5	10	27	58
1	LS	Spanish	15,651	4	12	45	39
1	LS	Other Language	4,448	4	14	37	45
1	LS	Students without Disabilities	22,237	4	11	42	43
1	LS	Students with Disabilities	3,439	4	19	50	27

Table H.2b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 1

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
1	RW	Female	11,986	14	21	20	44
1	RW	Male	13,690	20	25	20	36
1	RW	American Indian or Alaska Native	61	10	21	23	46
1	RW	Black or African-American	1,131	19	25	21	36
1	RW	Asian	5,252	11	15	17	57
1	RW	Native Hawaiian/Other Pacific Islander	113	15	22	14	49
1	RW	Hispanic or Latino	16,667	19	26	21	34
1	RW	White	2,417	15	21	20	44
1	RW	Multiracial (not Hispanic origin)	35	23	14	23	40
1	RW	NYC	16,153	18	22	19	41
1	RW	Big 4 Cities	1,198	34	27	17	23
1	RW	High Need Urban/Suburban	3,360	18	28	21	33
1	RW	High Need Rural	193	12	35	22	31
1	RW	Average Need	2,918	11	23	23	43
1	RW	Low Need	1,255	9	20	21	50
1	RW	Charter Schools	599	5	22	26	47
1	RW	0 Years LEP	14,263	18	22	19	40
1	RW	1 Years LEP	3,368	23	23	18	36
1	RW	2 Years LEP	7,012	13	25	22	41
1	RW	3 Years LEP	948	11	24	23	43
1	RW	4 Years LEP	27	11	15	22	52
1	RW	5 Years LEP	6	33	17	17	33
1	RW	6 Years LEP or More	9	78	11	11	0
1	RW	Chinese	2,673	8	14	17	61
1	RW	English	1,983	12	25	22	41
1	RW	Haitian-Creole	235	26	27	17	30
1	RW	Korean	140	4	10	16	69
1	RW	Russian	546	12	17	19	51
1	RW	Spanish	15,651	20	26	20	34
1	RW	Other Language	4,448	16	19	19	46
1	RW	Students without Disabilities	22,237	15	22	20	42
1	RW	Students with Disabilities	3,439	30	28	17	24

Table H.2c: Proficiency Percentages by Subgroup—Grade 1 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
1	Overall	Female	11,986	15	24	35	27
1	Overall	Male	13,690	20	27	31	21
1	Overall	American Indian or Alaska Native	61	10	23	44	23
1	Overall	Black or African-American	1,131	19	27	31	23
1	Overall	Asian	5,252	11	20	37	31
1	Overall	Native Hawaiian/Other Pacific Islander	113	16	27	30	27
1	Overall	Hispanic or Latino	16,667	20	28	32	20
1	Overall	White	2,417	15	23	29	33
1	Overall	Multiracial (not Hispanic origin)	35	23	20	43	14
1	Overall	NYC	16,153	18	25	32	24
1	Overall	Big 4 Cities	1,198	34	29	24	13
1	Overall	High Need Urban/Suburban	3,360	19	30	32	20
1	Overall	High Need Rural	193	12	38	31	19
1	Overall	Average Need	2,918	12	25	35	29
1	Overall	Low Need	1,255	9	22	38	30
1	Overall	Charter Schools	599	5	24	46	25
1	Overall	0 Years LEP	14,263	19	25	33	23
1	Overall	1 Years LEP	3,368	23	26	28	22
1	Overall	2 Years LEP	7,012	13	26	35	26
1	Overall	3 Years LEP	948	11	25	35	29
1	Overall	4 Years LEP	27	15	19	41	26
1	Overall	5 Years LEP	6	33	17	17	33
1	Overall	6 Years LEP or More	9	78	11	0	11
1	Overall	Chinese	2,673	9	20	40	30
1	Overall	English	1,983	12	27	34	27
1	Overall	Haitian-Creole	235	27	28	24	21
1	Overall	Korean	140	4	15	41	39
1	Overall	Russian	546	13	19	27	42
1	Overall	Spanish	15,651	20	28	32	20
1	Overall	Other Language	4,448	16	22	32	30
1	Overall	Students without Disabilities	22,237	16	25	34	26
1	Overall	Students with Disabilities	3,439	31	31	27	11

Table H.3a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 2

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
2	LS	Female	9,497	2	6	31	61
2	LS	Male	11,662	2	7	35	56
2	LS	American Indian or Alaska Native	40	3	13	40	45
2	LS	Black or African-American	1,066	2	7	35	57
2	LS	Asian	4,252	3	9	32	55
2	LS	Native Hawaiian/Other Pacific Islander	95	2	6	33	59
2	LS	Hispanic or Latino	13,943	2	6	34	58
2	LS	White	1,737	2	9	26	62
2	LS	Multiracial (not Hispanic origin)	26	4	12	27	58
2	LS	NYC	13,574	3	7	34	56
2	LS	Big 4 Cities	1,166	3	12	38	47
2	LS	High Need Urban/Suburban	2,659	2	4	36	58
2	LS	High Need Rural	145	1	6	28	66
2	LS	Average Need	2,228	1	4	26	69
2	LS	Low Need	993	1	3	28	67
2	LS	Charter Schools	394	0	1	30	70
2	LS	0 Years LEP	11,201	3	8	33	55
2	LS	1 Years LEP	2,226	4	13	37	46
2	LS	2 Years LEP	2,157	0	5	36	59
2	LS	3 Years LEP	4,706	0	2	29	68
2	LS	4 Years LEP	774	0	2	30	67
2	LS	5 Years LEP	41	0	0	41	59
2	LS	6 Years LEP or More	16	13	25	25	38
2	LS	Chinese	2,026	4	8	34	53
2	LS	English	1,218	1	3	27	70
2	LS	Haitian-Creole	248	2	7	35	56
2	LS	Korean	108	5	5	27	64
2	LS	Russian	374	2	8	23	67
2	LS	Spanish	13,311	2	6	34	58
2	LS	Other Language	3,874	3	10	31	57
2	LS	Students without Disabilities	17,227	3	6	30	61
2	LS	Students with Disabilities	3,932	2	9	46	43

Table H.3b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 2

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
2	RW	Female	9,497	13	42	33	12
2	RW	Male	11,662	17	43	29	10
2	RW	American Indian or Alaska Native	40	13	48	35	5
2	RW	Black or African-American	1,066	21	43	28	9
2	RW	Asian	4,252	13	34	35	18
2	RW	Native Hawaiian/Other Pacific Islander	95	14	36	35	16
2	RW	Hispanic or Latino	13,943	16	46	30	8
2	RW	White	1,737	17	37	30	16
2	RW	Multiracial (not Hispanic origin)	26	19	35	23	23
2	RW	NYC	13,574	17	44	30	10
2	RW	Big 4 Cities	1,166	28	45	21	6
2	RW	High Need Urban/Suburban	2,659	13	42	35	10
2	RW	High Need Rural	145	12	39	37	13
2	RW	Average Need	2,228	10	40	35	15
2	RW	Low Need	993	8	34	36	22
2	RW	Charter Schools	394	4	41	43	12
2	RW	0 Years LEP	11,201	17	44	30	9
2	RW	1 Years LEP	2,226	24	42	24	10
2	RW	2 Years LEP	2,157	14	39	32	15
2	RW	3 Years LEP	4,706	9	40	37	13
2	RW	4 Years LEP	774	11	49	30	9
2	RW	5 Years LEP	41	17	49	29	5
2	RW	6 Years LEP or More	16	38	44	19	0
2	RW	Chinese	2,026	11	34	37	19
2	RW	English	1,218	10	39	35	16
2	RW	Haitian-Creole	248	21	43	28	8
2	RW	Korean	108	8	15	44	32
2	RW	Russian	374	11	35	31	23
2	RW	Spanish	13,311	16	46	30	8
2	RW	Other Language	3,874	18	38	31	13
2	RW	Students without Disabilities	17,227	13	41	34	13
2	RW	Students with Disabilities	3,932	27	50	19	4

Table H.3c: Proficiency Percentages by Subgroup—Grade 2 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
2	Overall	Female	9,497	13	42	34	11
2	Overall	Male	11,662	17	43	30	10
2	Overall	American Indian or Alaska Native	40	13	48	38	3
2	Overall	Black or African-American	1,066	21	43	28	9
2	Overall	Asian	4,252	13	34	37	16
2	Overall	Native Hawaiian/Other Pacific Islander	95	14	36	36	15
2	Overall	Hispanic or Latino	13,943	16	46	30	8
2	Overall	White	1,737	17	37	31	15
2	Overall	Multiracial (not Hispanic origin)	26	19	35	23	23
2	Overall	NYC	13,574	17	43	30	10
2	Overall	Big 4 Cities	1,166	28	45	22	5
2	Overall	High Need Urban/Suburban	2,659	13	42	35	9
2	Overall	High Need Rural	145	12	39	37	13
2	Overall	Average Need	2,228	10	40	36	14
2	Overall	Low Need	993	8	34	37	21
2	Overall	Charter Schools	394	4	41	44	12
2	Overall	0 Years LEP	11,201	17	44	30	9
2	Overall	1 Years LEP	2,226	24	42	25	10
2	Overall	2 Years LEP	2,157	14	39	33	14
2	Overall	3 Years LEP	4,706	9	40	38	13
2	Overall	4 Years LEP	774	11	49	31	9
2	Overall	5 Years LEP	41	17	49	29	5
2	Overall	6 Years LEP or More	16	38	44	0	19
2	Overall	Chinese	2,026	11	34	38	17
2	Overall	English	1,218	10	39	36	15
2	Overall	Haitian-Creole	248	21	43	28	8
2	Overall	Korean	108	8	16	45	31
2	Overall	Russian	374	11	35	31	23
2	Overall	Spanish	13,311	16	46	30	8
2	Overall	Other Language	3,874	18	38	32	13
2	Overall	Students without Disabilities	17,227	13	41	35	12
2	Overall	Students with Disabilities	3,932	27	50	19	3

Table H.4a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 3

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
3	LS	Female	8,643	3	5	31	60
3	LS	Male	10,381	3	7	35	55
3	LS	American Indian or Alaska Native	32	13	9	28	50
3	LS	Black or African-American	1,132	3	6	34	57
3	LS	Asian	3,549	4	9	34	53
3	LS	Native Hawaiian/Other Pacific Islander	66	2	15	41	42
3	LS	Hispanic or Latino	12,747	3	5	34	58
3	LS	White	1,482	4	7	29	60
3	LS	Multiracial (not Hispanic origin)	16	0	6	25	69
3	LS	NYC	12,386	4	7	34	55
3	LS	Big 4 Cities	1,104	3	10	42	45
3	LS	High Need Urban/Suburban	2,416	2	4	34	59
3	LS	High Need Rural	128	5	3	27	65
3	LS	Average Need	1,900	2	3	27	68
3	LS	Low Need	793	2	5	27	66
3	LS	Charter Schools	297	0	1	30	69
3	LS	0 Years LEP	10,189	5	7	33	55
3	LS	1 Years LEP	1,815	7	16	42	35
3	LS	2 Years LEP	1,516	0	5	36	59
3	LS	3 Years LEP	1,544	0	2	33	64
3	LS	4 Years LEP	3,259	0	1	30	69
3	LS	5 Years LEP	625	0	1	31	68
3	LS	6 Years LEP or More	57	12	5	28	54
3	LS	Chinese	1,607	5	9	35	51
3	LS	English	919	1	3	28	68
3	LS	Haitian-Creole	283	3	5	33	59
3	LS	Korean	67	3	7	39	51
3	LS	Russian	289	3	10	23	64
3	LS	Spanish	12,297	3	5	34	58
3	LS	Other Language	3,562	4	9	32	55
3	LS	Students without Disabilities	15,007	4	6	30	61
3	LS	Students with Disabilities	4,017	2	7	47	45

Table H.4b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 3

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
3	RW	Female	8,643	11	23	41	25
3	RW	Male	10,381	14	26	40	21
3	RW	American Indian or Alaska Native	32	25	25	25	25
3	RW	Black or African-American	1,132	16	29	35	20
3	RW	Asian	3,549	13	20	37	31
3	RW	Native Hawaiian/Other Pacific Islander	66	21	21	30	27
3	RW	Hispanic or Latino	12,747	12	26	42	21
3	RW	White	1,482	15	22	37	25
3	RW	Multiracial (not Hispanic origin)	16	6	13	50	31
3	RW	NYC	12,386	14	25	40	21
3	RW	Big 4 Cities	1,104	21	33	31	15
3	RW	High Need Urban/Suburban	2,416	10	24	42	24
3	RW	High Need Rural	128	9	23	41	27
3	RW	Average Need	1,900	7	22	41	29
3	RW	Low Need	793	8	18	36	38
3	RW	Charter Schools	297	1	17	49	33
3	RW	0 Years LEP	10,189	14	25	41	21
3	RW	1 Years LEP	1,815	26	30	27	17
3	RW	2 Years LEP	1,516	10	24	38	28
3	RW	3 Years LEP	1,544	8	23	42	27
3	RW	4 Years LEP	3,259	5	21	45	29
3	RW	5 Years LEP	625	8	27	42	24
3	RW	6 Years LEP or More	57	28	21	33	18
3	RW	Chinese	1,607	12	17	38	33
3	RW	English	919	6	22	42	30
3	RW	Haitian-Creole	283	20	27	35	19
3	RW	Korean	67	9	15	30	46
3	RW	Russian	289	12	21	37	30
3	RW	Spanish	12,297	12	26	41	20
3	RW	Other Language	3,562	16	24	36	25
3	RW	Students without Disabilities	15,007	11	21	42	26
3	RW	Students with Disabilities	4,017	20	37	32	10

Table H.4c: Proficiency Percentages by Subgroup—Grade 3 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
3	Overall	Female	8,643	11	23	43	22
3	Overall	Male	10,381	14	26	42	19
3	Overall	American Indian or Alaska Native	32	25	25	25	25
3	Overall	Black or African-American	1,132	16	29	38	17
3	Overall	Asian	3,549	13	20	40	27
3	Overall	Native Hawaiian/Other Pacific Islander	66	21	21	32	26
3	Overall	Hispanic or Latino	12,747	12	26	44	18
3	Overall	White	1,482	15	22	38	24
3	Overall	Multiracial (not Hispanic origin)	16	6	13	56	25
3	Overall	NYC	12,386	14	25	42	19
3	Overall	Big 4 Cities	1,104	21	33	34	13
3	Overall	High Need Urban/Suburban	2,416	10	24	45	20
3	Overall	High Need Rural	128	9	23	41	26
3	Overall	Average Need	1,900	8	22	43	28
3	Overall	Low Need	793	8	18	40	34
3	Overall	Charter Schools	297	1	17	53	29
3	Overall	0 Years LEP	10,189	14	25	43	18
3	Overall	1 Years LEP	1,815	26	30	29	15
3	Overall	2 Years LEP	1,516	10	24	41	24
3	Overall	3 Years LEP	1,544	8	23	45	24
3	Overall	4 Years LEP	3,259	5	21	48	27
3	Overall	5 Years LEP	625	8	27	44	22
3	Overall	6 Years LEP or More	57	28	21	35	16
3	Overall	Chinese	1,607	12	17	43	27
3	Overall	English	919	6	22	45	27
3	Overall	Haitian-Creole	283	20	27	37	17
3	Overall	Korean	67	9	16	37	37
3	Overall	Russian	289	12	21	39	27
3	Overall	Spanish	12,297	12	26	44	18
3	Overall	Other Language	3,562	16	24	38	23
3	Overall	Students without Disabilities	15,007	11	21	45	23
3	Overall	Students with Disabilities	4,017	20	37	33	9

Table H.5a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 4

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
4	LS	Female	7,981	4	7	27	62
4	LS	Male	9,872	3	7	31	58
4	LS	American Indian or Alaska Native	37	3	8	35	54
4	LS	Black or African-American	1,074	3	8	30	59
4	LS	Asian	3,156	5	10	33	51
4	LS	Native Hawaiian/Other Pacific Islander	42	5	7	38	50
4	LS	Hispanic or Latino	12,180	3	6	29	62
4	LS	White	1,346	5	8	25	62
4	LS	Multiracial (not Hispanic origin)	18	6	11	22	61
4	LS	NYC	11,916	4	7	30	59
4	LS	Big 4 Cities	1,120	5	9	35	51
4	LS	High Need Urban/Suburban	2,145	3	5	29	63
4	LS	High Need Rural	109	4	10	35	51
4	LS	Average Need	1,674	1	5	25	69
4	LS	Low Need	682	3	5	27	65
4	LS	Charter Schools	207	0	0	24	76
4	LS	0 Years LEP	9,894	5	7	28	60
4	LS	1 Years LEP	1,839	9	19	38	35
4	LS	2 Years LEP	1,335	1	9	40	50
4	LS	3 Years LEP	991	1	4	30	65
4	LS	4 Years LEP	732	1	2	29	69
4	LS	5 Years LEP	2,447	0	1	24	75
4	LS	6 Years LEP or More	594	1	1	25	73
4	LS	Chinese	1,376	7	10	35	48
4	LS	English	848	1	3	26	70
4	LS	Haitian-Creole	296	5	8	29	58
4	LS	Korean	86	2	8	38	51
4	LS	Russian	270	4	9	21	65
4	LS	Spanish	11,689	3	6	29	62
4	LS	Other Language	3,288	4	10	30	56
4	LS	Students without Disabilities	13,425	4	7	26	62
4	LS	Students with Disabilities	4,428	2	5	39	55

Table H.5b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 4

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
4	RW	Female	7,981	10	20	50	20
4	RW	Male	9,872	12	22	48	18
4	RW	American Indian or Alaska Native	37	24	16	43	16
4	RW	Black or African-American	1,074	14	27	44	15
4	RW	Asian	3,156	13	20	43	24
4	RW	Native Hawaiian/Other Pacific Islander	42	17	24	31	29
4	RW	Hispanic or Latino	12,180	10	21	51	17
4	RW	White	1,346	13	19	46	23
4	RW	Multiracial (not Hispanic origin)	18	17	11	39	33
4	RW	NYC	11,916	12	21	49	18
4	RW	Big 4 Cities	1,120	18	27	41	14
4	RW	High Need Urban/Suburban	2,145	10	21	50	20
4	RW	High Need Rural	109	14	26	41	19
4	RW	Average Need	1,674	7	18	50	25
4	RW	Low Need	682	6	19	49	26
4	RW	Charter Schools	207	1	10	61	28
4	RW	0 Years LEP	9,894	12	20	51	17
4	RW	1 Years LEP	1,839	26	29	32	13
4	RW	2 Years LEP	1,335	13	25	42	20
4	RW	3 Years LEP	991	7	21	49	23
4	RW	4 Years LEP	732	6	21	51	22
4	RW	5 Years LEP	2,447	3	16	56	25
4	RW	6 Years LEP or More	594	5	21	55	20
4	RW	Chinese	1,376	13	17	43	26
4	RW	English	848	6	19	54	21
4	RW	Haitian-Creole	296	16	30	38	17
4	RW	Korean	86	3	14	48	35
4	RW	Russian	270	10	20	42	28
4	RW	Spanish	11,689	11	21	51	17
4	RW	Other Language	3,288	15	22	44	20
4	RW	Students without Disabilities	13,425	11	17	50	22
4	RW	Students with Disabilities	4,428	13	33	45	9

Table H.5c: Proficiency Percentages by Subgroup—Grade 4 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
4	Overall	Female	7,981	11	20	52	18
4	Overall	Male	9,872	12	22	49	16
4	Overall	American Indian or Alaska Native	37	24	16	43	16
4	Overall	Black or African-American	1,074	14	27	44	15
4	Overall	Asian	3,156	13	20	46	21
4	Overall	Native Hawaiian/Other Pacific Islander	42	19	21	31	29
4	Overall	Hispanic or Latino	12,180	11	21	52	16
4	Overall	White	1,346	14	19	46	22
4	Overall	Multiracial (not Hispanic origin)	18	17	11	39	33
4	Overall	NYC	11,916	12	21	50	16
4	Overall	Big 4 Cities	1,120	18	27	42	13
4	Overall	High Need Urban/Suburban	2,145	10	21	52	18
4	Overall	High Need Rural	109	15	27	44	15
4	Overall	Average Need	1,674	7	18	51	23
4	Overall	Low Need	682	6	19	51	24
4	Overall	Charter Schools	207	1	10	64	25
4	Overall	0 Years LEP	9,894	12	20	52	16
4	Overall	1 Years LEP	1,839	26	29	33	12
4	Overall	2 Years LEP	1,335	13	25	44	18
4	Overall	3 Years LEP	991	7	21	50	22
4	Overall	4 Years LEP	732	6	21	53	20
4	Overall	5 Years LEP	2,447	3	17	57	23
4	Overall	6 Years LEP or More	594	5	21	56	19
4	Overall	Chinese	1,376	13	18	47	22
4	Overall	English	848	6	19	55	19
4	Overall	Haitian-Creole	296	16	30	39	16
4	Overall	Korean	86	5	14	52	29
4	Overall	Russian	270	10	20	42	27
4	Overall	Spanish	11,689	11	22	52	16
4	Overall	Other Language	3,288	15	22	45	18
4	Overall	Students without Disabilities	13,425	11	17	52	20
4	Overall	Students with Disabilities	4,428	13	33	46	8

Table H.6a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 5

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
5	LS	Female	6,957	5	11	33	51
5	LS	Male	8,502	6	12	35	47
5	LS	American Indian or Alaska Native	28	14	14	25	46
5	LS	Black or African-American	1,033	4	15	36	45
5	LS	Asian	2,710	7	19	34	40
5	LS	Native Hawaiian/Other Pacific Islander	53	4	19	32	45
5	LS	Hispanic or Latino	10,441	5	9	34	52
5	LS	White	1,178	7	16	30	47
5	LS	Multiracial (not Hispanic origin)	16	6	13	44	38
5	LS	NYC	10,445	6	13	34	48
5	LS	Big 4 Cities	1,027	6	17	39	37
5	LS	High Need Urban/Suburban	1,726	5	9	35	52
5	LS	High Need Rural	94	2	13	36	49
5	LS	Average Need	1,311	4	9	29	58
5	LS	Low Need	565	4	10	28	58
5	LS	Charter Schools	291	0	2	45	52
5	LS	0 Years LEP	8,482	6	11	31	51
5	LS	1 Years LEP	1,879	12	28	37	23
5	LS	2 Years LEP	1,350	2	16	43	38
5	LS	3 Years LEP	767	1	9	40	50
5	LS	4 Years LEP	460	0	3	38	58
5	LS	5 Years LEP	511	0	4	35	62
5	LS	6 Years LEP or More	1,982	0	2	31	67
5	LS	Chinese	1,092	8	20	36	35
5	LS	English	645	3	7	32	58
5	LS	Haitian-Creole	280	4	15	38	43
5	LS	Korean	80	4	8	40	49
5	LS	Russian	245	7	15	27	51
5	LS	Spanish	10,129	5	9	34	52
5	LS	Other Language	2,988	6	18	33	44
5	LS	Students without Disabilities	11,329	6	13	31	49
5	LS	Students with Disabilities	4,130	2	8	41	48

Table H.6b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 5

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
5	RW	Female	6,957	10	16	35	39
5	RW	Male	8,502	14	19	35	32
5	RW	American Indian or Alaska Native	28	14	14	29	43
5	RW	Black or African-American	1,033	14	23	38	25
5	RW	Asian	2,710	14	17	29	40
5	RW	Native Hawaiian/Other Pacific Islander	53	9	13	30	47
5	RW	Hispanic or Latino	10,441	11	17	37	35
5	RW	White	1,178	16	16	33	35
5	RW	Multiracial (not Hispanic origin)	16	19	19	19	44
5	RW	NYC	10,445	13	17	35	35
5	RW	Big 4 Cities	1,027	19	27	33	21
5	RW	High Need Urban/Suburban	1,726	10	17	33	40
5	RW	High Need Rural	94	7	18	36	38
5	RW	Average Need	1,311	9	14	38	39
5	RW	Low Need	565	8	12	32	48
5	RW	Charter Schools	291	1	15	42	42
5	RW	0 Years LEP	8,482	13	16	36	35
5	RW	1 Years LEP	1,879	24	25	28	23
5	RW	2 Years LEP	1,350	11	19	33	36
5	RW	3 Years LEP	767	7	19	36	39
5	RW	4 Years LEP	460	5	18	36	41
5	RW	5 Years LEP	511	6	15	34	45
5	RW	6 Years LEP or More	1,982	3	14	39	44
5	RW	Chinese	1,092	13	17	27	42
5	RW	English	645	8	15	40	38
5	RW	Haitian-Creole	280	13	24	38	26
5	RW	Korean	80	5	10	24	61
5	RW	Russian	245	11	15	36	38
5	RW	Spanish	10,129	11	17	37	35
5	RW	Other Language	2,988	16	19	32	34
5	RW	Students without Disabilities	11,329	12	15	34	40
5	RW	Students with Disabilities	4,130	13	25	39	23

Table H.6c: Proficiency Percentages by Subgroup—Grade 5 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
5	Overall	Female	6,957	11	17	42	30
5	Overall	Male	8,502	14	20	41	25
5	Overall	American Indian or Alaska Native	28	18	14	36	32
5	Overall	Black or African-American	1,033	14	25	41	20
5	Overall	Asian	2,710	15	20	37	28
5	Overall	Native Hawaiian/Other Pacific Islander	53	11	19	32	38
5	Overall	Hispanic or Latino	10,441	12	18	42	28
5	Overall	White	1,178	16	17	39	28
5	Overall	Multiracial (not Hispanic origin)	16	19	19	44	19
5	Overall	NYC	10,445	13	19	41	27
5	Overall	Big 4 Cities	1,027	19	27	38	16
5	Overall	High Need Urban/Suburban	1,726	10	17	42	30
5	Overall	High Need Rural	94	7	21	40	31
5	Overall	Average Need	1,311	9	15	43	32
5	Overall	Low Need	565	9	15	39	38
5	Overall	Charter Schools	291	1	15	55	28
5	Overall	0 Years LEP	8,482	13	17	41	28
5	Overall	1 Years LEP	1,879	26	27	32	15
5	Overall	2 Years LEP	1,350	12	22	41	25
5	Overall	3 Years LEP	767	7	20	44	29
5	Overall	4 Years LEP	460	5	18	45	32
5	Overall	5 Years LEP	511	6	16	43	35
5	Overall	6 Years LEP or More	1,982	3	14	47	36
5	Overall	Chinese	1,092	15	20	38	26
5	Overall	English	645	8	16	45	31
5	Overall	Haitian-Creole	280	14	26	41	19
5	Overall	Korean	80	6	10	41	43
5	Overall	Russian	245	11	18	40	31
5	Overall	Spanish	10,129	12	18	43	28
5	Overall	Other Language	2,988	16	21	37	26
5	Overall	Students without Disabilities	11,329	12	16	40	31
5	Overall	Students with Disabilities	4,130	13	25	43	18

Table H.7a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 6

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
6	LS	Female	5,643	9	15	37	39
6	LS	Male	7,521	8	15	38	39
6	LS	American Indian or Alaska Native	23	4	30	48	17
6	LS	Black or African-American	992	6	15	38	41
6	LS	Asian	2,364	10	23	38	29
6	LS	Native Hawaiian/Other Pacific Islander	57	16	23	30	32
6	LS	Hispanic or Latino	8,689	8	12	37	42
6	LS	White	1,030	11	18	35	36
6	LS	Multiracial (not Hispanic origin)	9	0	11	33	56
6	LS	NYC	9,048	9	16	38	37
6	LS	Big 4 Cities	952	9	17	40	34
6	LS	High Need Urban/Suburban	1,362	8	12	38	43
6	LS	High Need Rural	96	9	13	32	46
6	LS	Average Need	1,060	5	11	32	52
6	LS	Low Need	409	5	12	32	52
6	LS	Charter Schools	237	0	5	46	49
6	LS	0 Years LEP	6,898	10	12	36	42
6	LS	1 Years LEP	1,855	17	31	35	17
6	LS	2 Years LEP	1,359	5	25	46	24
6	LS	3 Years LEP	751	2	11	45	42
6	LS	4 Years LEP	331	1	9	43	47
6	LS	5 Years LEP	276	0	5	40	54
6	LS	6 Years LEP or More	1,672	1	3	35	61
6	LS	Chinese	995	13	27	34	26
6	LS	English	522	6	10	37	47
6	LS	Haitian-Creole	338	7	17	38	38
6	LS	Korean	65	5	18	40	37
6	LS	Russian	235	12	22	29	37
6	LS	Spanish	8,433	8	12	38	42
6	LS	Other Language	2,576	8	19	39	34
6	LS	Students without Disabilities	9,642	10	17	36	37
6	LS	Students with Disabilities	3,522	3	9	42	45

Table H.7b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 6

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
6	RW	Female	5,643	13	19	36	32
6	RW	Male	7,521	15	24	35	26
6	RW	American Indian or Alaska Native	23	22	35	26	17
6	RW	Black or African-American	992	13	22	39	26
6	RW	Asian	2,364	17	20	30	32
6	RW	Native Hawaiian/Other Pacific Islander	57	19	19	31	30
6	RW	Hispanic or Latino	8,689	14	22	37	27
6	RW	White	1,030	18	22	31	30
6	RW	Multiracial (not Hispanic origin)	9	0	11	22	67
6	RW	NYC	9,048	16	22	35	26
6	RW	Big 4 Cities	952	18	24	35	22
6	RW	High Need Urban/Suburban	1,362	12	22	33	33
6	RW	High Need Rural	96	15	21	31	33
6	RW	Average Need	1,060	8	20	35	37
6	RW	Low Need	409	8	17	27	48
6	RW	Charter Schools	237	2	14	51	34
6	RW	0 Years LEP	6,898	16	21	37	27
6	RW	1 Years LEP	1,855	26	27	26	21
6	RW	2 Years LEP	1,359	14	26	32	28
6	RW	3 Years LEP	751	8	22	34	35
6	RW	4 Years LEP	331	6	18	38	38
6	RW	5 Years LEP	276	3	20	41	36
6	RW	6 Years LEP or More	1,672	3	18	41	38
6	RW	Chinese	995	18	20	31	31
6	RW	English	522	9	20	33	39
6	RW	Haitian-Creole	338	13	21	38	28
6	RW	Korean	65	8	9	31	52
6	RW	Russian	235	17	18	30	34
6	RW	Spanish	8,433	14	22	37	27
6	RW	Other Language	2,576	17	22	32	29
6	RW	Students without Disabilities	9,642	15	19	34	32
6	RW	Students with Disabilities	3,522	13	29	39	19

Table H.7c: Proficiency Percentages by Subgroup—Grade 6 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
6	Overall	Female	5,643	15	21	43	21
6	Overall	Male	7,521	16	26	40	18
6	Overall	American Indian or Alaska Native	23	22	39	35	4
6	Overall	Black or African-American	992	13	24	43	20
6	Overall	Asian	2,364	18	25	38	19
6	Overall	Native Hawaiian/Other Pacific Islander	57	23	23	32	23
6	Overall	Hispanic or Latino	8,689	15	23	43	19
6	Overall	White	1,030	19	24	37	20
6	Overall	Multiracial (not Hispanic origin)	9	0	11	44	44
6	Overall	NYC	9,048	17	24	41	18
6	Overall	Big 4 Cities	952	19	24	41	15
6	Overall	High Need Urban/Suburban	1,362	13	23	42	22
6	Overall	High Need Rural	96	15	26	34	25
6	Overall	Average Need	1,060	9	21	42	28
6	Overall	Low Need	409	8	18	40	33
6	Overall	Charter Schools	237	3	14	62	22
6	Overall	0 Years LEP	6,898	17	22	42	19
6	Overall	1 Years LEP	1,855	28	31	30	11
6	Overall	2 Years LEP	1,359	15	29	40	16
6	Overall	3 Years LEP	751	9	23	44	24
6	Overall	4 Years LEP	331	6	19	50	25
6	Overall	5 Years LEP	276	3	21	51	25
6	Overall	6 Years LEP or More	1,672	3	18	49	29
6	Overall	Chinese	995	19	26	37	18
6	Overall	English	522	10	20	43	26
6	Overall	Haitian-Creole	338	13	25	42	20
6	Overall	Korean	65	8	17	46	29
6	Overall	Russian	235	19	23	31	27
6	Overall	Spanish	8,433	15	23	43	19
6	Overall	Other Language	2,576	17	24	39	19
6	Overall	Students without Disabilities	9,642	16	21	41	21
6	Overall	Students with Disabilities	3,522	13	29	43	14

Table H.8a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 7

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
7	LS	Female	5,415	5	15	37	43
7	LS	Male	7,194	5	14	36	44
7	LS	American Indian or Alaska Native	32	6	25	25	44
7	LS	Black or African-American	1,008	3	12	40	45
7	LS	Asian	2,229	5	22	40	32
7	LS	Native Hawaiian/Other Pacific Islander	45	4	16	44	36
7	LS	Hispanic or Latino	8,428	6	12	35	47
7	LS	White	853	6	17	34	43
7	LS	Multiracial (not Hispanic origin)	14	7	14	43	36
7	LS	NYC	8,853	6	15	37	42
7	LS	Big 4 Cities	898	4	16	42	38
7	LS	High Need Urban/Suburban	1,300	6	13	36	45
7	LS	High Need Rural	78	3	6	45	46
7	LS	Average Need	905	5	12	30	53
7	LS	Low Need	399	3	11	34	52
7	LS	Charter Schools	176	1	2	32	65
7	LS	0 Years LEP	6,512	6	13	34	47
7	LS	1 Years LEP	1,972	10	31	36	22
7	LS	2 Years LEP	1,324	2	19	45	34
7	LS	3 Years LEP	823	0	11	45	44
7	LS	4 Years LEP	292	1	5	39	55
7	LS	5 Years LEP	224	1	4	41	54
7	LS	6 Years LEP or More	1,443	1	2	33	65
7	LS	Chinese	979	7	27	39	27
7	LS	English	432	2	8	32	57
7	LS	Haitian-Creole	331	4	12	37	47
7	LS	Korean	60	3	12	43	42
7	LS	Russian	174	9	22	27	43
7	LS	Spanish	8,210	6	13	35	47
7	LS	Other Language	2,423	4	17	40	39
7	LS	Students without Disabilities	9,259	6	17	34	42
7	LS	Students with Disabilities	3,350	2	7	42	49

Table H.8b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 7

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
7	RW	Female	5,415	19	36	31	13
7	RW	Male	7,194	24	38	29	10
7	RW	American Indian or Alaska Native	32	34	22	31	13
7	RW	Black or African-American	1,008	19	40	29	12
7	RW	Asian	2,229	25	36	27	13
7	RW	Native Hawaiian/Other Pacific Islander	45	27	31	33	9
7	RW	Hispanic or Latino	8,428	21	38	31	11
7	RW	White	853	24	35	27	14
7	RW	Multiracial (not Hispanic origin)	14	14	50	29	7
7	RW	NYC	8,853	23	37	29	11
7	RW	Big 4 Cities	898	24	46	23	6
7	RW	High Need Urban/Suburban	1,300	20	36	30	14
7	RW	High Need Rural	78	21	42	26	12
7	RW	Average Need	905	17	32	33	18
7	RW	Low Need	399	15	30	37	19
7	RW	Charter Schools	176	6	29	46	19
7	RW	0 Years LEP	6,512	22	36	31	10
7	RW	1 Years LEP	1,972	37	36	18	9
7	RW	2 Years LEP	1,324	21	41	26	12
7	RW	3 Years LEP	823	15	40	31	14
7	RW	4 Years LEP	292	11	39	33	17
7	RW	5 Years LEP	224	10	35	42	13
7	RW	6 Years LEP or More	1,443	8	39	38	16
7	RW	Chinese	979	28	35	24	13
7	RW	English	432	14	34	34	18
7	RW	Haitian-Creole	331	21	36	28	15
7	RW	Korean	60	10	27	40	23
7	RW	Russian	174	26	33	26	14
7	RW	Spanish	8,210	21	38	30	11
7	RW	Other Language	2,423	23	37	29	11
7	RW	Students without Disabilities	9,259	22	34	30	13
7	RW	Students with Disabilities	3,350	20	46	28	6

Table H.8c: Proficiency Percentages by Subgroup—Grade 7 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
7	Overall	Female	5,415	19	36	33	12
7	Overall	Male	7,194	24	38	30	9
7	Overall	American Indian or Alaska Native	32	34	25	28	13
7	Overall	Black or African-American	1,008	20	40	30	11
7	Overall	Asian	2,229	25	36	29	10
7	Overall	Native Hawaiian/Other Pacific Islander	45	27	31	33	9
7	Overall	Hispanic or Latino	8,428	21	38	32	10
7	Overall	White	853	25	35	28	12
7	Overall	Multiracial (not Hispanic origin)	14	14	50	29	7
7	Overall	NYC	8,853	23	38	30	9
7	Overall	Big 4 Cities	898	25	46	25	5
7	Overall	High Need Urban/Suburban	1,300	20	36	31	12
7	Overall	High Need Rural	78	21	42	26	12
7	Overall	Average Need	905	17	32	34	16
7	Overall	Low Need	399	15	30	39	16
7	Overall	Charter Schools	176	6	29	49	16
7	Overall	0 Years LEP	6,512	22	36	33	9
7	Overall	1 Years LEP	1,972	38	36	19	8
7	Overall	2 Years LEP	1,324	22	41	27	10
7	Overall	3 Years LEP	823	15	40	33	12
7	Overall	4 Years LEP	292	11	39	34	16
7	Overall	5 Years LEP	224	10	35	42	13
7	Overall	6 Years LEP or More	1,443	8	39	39	15
7	Overall	Chinese	979	28	36	26	10
7	Overall	English	432	15	34	35	17
7	Overall	Haitian-Creole	331	22	36	28	14
7	Overall	Korean	60	10	28	42	20
7	Overall	Russian	174	27	32	29	11
7	Overall	Spanish	8,210	21	38	31	9
7	Overall	Other Language	2,423	23	37	30	10
7	Overall	Students without Disabilities	8,210	21	38	31	9
7	Overall	Students with Disabilities	3,350	20	46	29	6

Table H.9a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 8

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
8	LS	Female	5,478	6	19	30	45
8	LS	Male	7,028	6	18	30	46
8	LS	American Indian or Alaska Native	38	8	32	29	32
8	LS	Black or African-American	954	4	15	35	46
8	LS	Asian	2,517	6	26	33	35
8	LS	Native Hawaiian/Other Pacific Islander	66	2	24	35	39
8	LS	Hispanic or Latino	7,962	6	16	29	49
8	LS	White	960	6	23	26	45
8	LS	Multiracial (not Hispanic origin)	9	22	33	11	33
8	LS	NYC	9,059	6	19	31	45
8	LS	Big 4 Cities	869	7	23	33	38
8	LS	High Need Urban/Suburban	1,134	7	19	28	46
8	LS	High Need Rural	72	4	18	33	44
8	LS	Average Need	857	5	14	25	56
8	LS	Low Need	421	4	11	27	57
8	LS	Charter Schools	94	1	2	30	67
8	LS	0 Years LEP	6,427	6	15	28	50
8	LS	1 Years LEP	2,014	12	35	31	22
8	LS	2 Years LEP	1,572	4	25	38	34
8	LS	3 Years LEP	864	1	16	39	44
8	LS	4 Years LEP	324	0	7	36	57
8	LS	5 Years LEP	217	0	4	32	63
8	LS	6 Years LEP or More	1,073	0	4	21	75
8	LS	Chinese	1,177	7	31	33	29
8	LS	English	389	1	13	29	58
8	LS	Haitian-Creole	304	4	13	37	45
8	LS	Korean	80	3	15	35	48
8	LS	Russian	236	5	30	22	44
8	LS	Spanish	7,797	6	16	29	49
8	LS	Other Language	2,523	5	21	33	41
8	LS	Students without Disabilities	9,355	7	22	29	42
8	LS	Students with Disabilities	3,151	3	9	32	56

Table H.9b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 8

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
8	RW	Female	5,478	19	36	27	17
8	RW	Male	7,028	22	38	27	13
8	RW	American Indian or Alaska Native	38	29	39	21	11
8	RW	Black or African-American	954	20	39	26	16
8	RW	Asian	2,517	21	37	25	17
8	RW	Native Hawaiian/Other Pacific Islander	66	17	48	20	15
8	RW	Hispanic or Latino	7,962	21	37	28	14
8	RW	White	960	21	36	25	18
8	RW	Multiracial (not Hispanic origin)	9	33	22	22	22
8	RW	NYC	9,059	21	38	27	14
8	RW	Big 4 Cities	869	29	41	20	9
8	RW	High Need Urban/Suburban	1,134	23	36	25	16
8	RW	High Need Rural	72	18	36	26	19
8	RW	Average Need	857	16	31	33	20
8	RW	Low Need	421	10	32	32	26
8	RW	Charter Schools	94	2	26	49	23
8	RW	0 Years LEP	6,427	20	37	29	14
8	RW	1 Years LEP	2,014	36	37	16	11
8	RW	2 Years LEP	1,572	22	41	23	14
8	RW	3 Years LEP	864	15	37	28	19
8	RW	4 Years LEP	324	8	40	29	23
8	RW	5 Years LEP	217	10	35	31	24
8	RW	6 Years LEP or More	1,073	6	35	38	21
8	RW	Chinese	1,177	23	35	25	18
8	RW	English	389	12	34	30	24
8	RW	Haitian-Creole	304	16	42	24	17
8	RW	Korean	80	10	23	40	28
8	RW	Russian	236	16	38	26	20
8	RW	Spanish	7,797	21	37	28	14
8	RW	Other Language	2,523	22	39	24	15
8	RW	Students without Disabilities	9,355	21	35	27	16
8	RW	Students with Disabilities	3,151	19	43	28	10

Table H.9c: Proficiency Percentages by Subgroup—Grade 8 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
8	Overall	Female	5,478	19	36	29	15
8	Overall	Male	7,028	22	38	28	12
8	Overall	American Indian or Alaska Native	38	32	37	24	8
8	Overall	Black or African-American	954	20	39	28	14
8	Overall	Asian	2,517	22	38	26	15
8	Overall	Native Hawaiian/Other Pacific Islander	66	17	48	20	15
8	Overall	Hispanic or Latino	7,962	21	37	30	12
8	Overall	White	960	22	36	26	16
8	Overall	Multiracial (not Hispanic origin)	9	33	22	22	22
8	Overall	NYC	9,059	21	38	28	12
8	Overall	Big 4 Cities	869	30	41	21	8
8	Overall	High Need Urban/Suburban	1,134	23	36	26	15
8	Overall	High Need Rural	72	18	36	28	18
8	Overall	Average Need	857	16	32	34	18
8	Overall	Low Need	421	11	32	34	23
8	Overall	Charter Schools	94	2	26	51	21
8	Overall	0 Years LEP	6,427	20	37	30	13
8	Overall	1 Years LEP	2,014	37	37	17	9
8	Overall	2 Years LEP	1,572	23	42	24	12
8	Overall	3 Years LEP	864	15	38	30	17
8	Overall	4 Years LEP	324	8	40	31	20
8	Overall	5 Years LEP	217	10	35	33	22
8	Overall	6 Years LEP or More	1,073	6	35	40	19
8	Overall	Chinese	1,177	24	37	25	14
8	Overall	English	389	12	33	33	22
8	Overall	Haitian-Creole	304	16	42	28	13
8	Overall	Korean	80	10	23	43	25
8	Overall	Russian	236	16	39	25	19
8	Overall	Spanish	7,797	21	37	29	12
8	Overall	Other Language	2,523	22	39	25	14
8	Overall	Students without Disabilities	9,355	22	36	28	14
8	Overall	Students with Disabilities	3,151	19	43	29	9

Table H.10a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 9

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
9	LS	Female	6,066	12	25	22	41
9	LS	Male	8,222	12	24	20	44
9	LS	American Indian or Alaska Native	43	16	49	12	23
9	LS	Black or African-American	1,354	7	26	28	39
9	LS	Asian	3,201	12	34	25	29
9	LS	Native Hawaiian/Other Pacific Islander	55	16	31	25	27
9	LS	Hispanic or Latino	8,647	13	20	18	48
9	LS	White	974	8	26	24	42
9	LS	Multiracial (not Hispanic origin)	14	0	14	21	64
9	LS	NYC	10,055	12	24	21	43
9	LS	Big 4 Cities	1,038	8	28	23	41
9	LS	High Need Urban/Suburban	1,514	17	25	20	37
9	LS	High Need Rural	75	7	24	21	48
9	LS	Average Need	994	11	27	21	41
9	LS	Low Need	462	8	21	21	50
9	LS	Charter Schools	150	2	10	26	62
9	LS	0 Years LEP	7,162	13	20	18	49
9	LS	1 Years LEP	2,806	21	39	20	20
9	LS	2 Years LEP	1,795	7	34	30	29
9	LS	3 Years LEP	1,042	3	21	32	44
9	LS	4 Years LEP	337	2	13	26	58
9	LS	5 Years LEP	200	1	10	28	62
9	LS	6 Years LEP or More	909	0	3	16	80
9	LS	Chinese	1,478	15	37	24	24
9	LS	English	461	6	21	25	48
9	LS	Haitian-Creole	450	5	24	28	44
9	LS	Korean	76	1	14	42	42
9	LS	Russian	204	9	19	23	50
9	LS	Spanish	8,450	14	20	18	48
9	LS	Other Language	3,169	9	31	25	35
9	LS	Students without Disabilities	11,607	14	28	21	37
9	LS	Students with Disabilities	2,681	3	8	22	67

Table H.10b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 9

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
9	RW	Female	6,066	17	41	24	18
9	RW	Male	8,222	19	41	22	18
9	RW	American Indian or Alaska Native	43	37	42	14	7
9	RW	Black or African-American	1,354	14	44	24	17
9	RW	Asian	3,201	20	44	21	16
9	RW	Native Hawaiian/Other Pacific Islander	55	22	44	15	20
9	RW	Hispanic or Latino	8,647	19	39	24	19
9	RW	White	974	16	41	22	21
9	RW	Multiracial (not Hispanic origin)	14	0	29	14	57
9	RW	NYC	10,055	17	41	23	19
9	RW	Big 4 Cities	1,038	23	45	19	13
9	RW	High Need Urban/Suburban	1,514	24	41	22	13
9	RW	High Need Rural	75	13	44	27	16
9	RW	Average Need	994	20	38	23	19
9	RW	Low Need	462	14	30	25	31
9	RW	Charter Schools	150	2	36	39	23
9	RW	0 Years LEP	7,162	17	38	24	21
9	RW	1 Years LEP	2,806	30	46	14	11
9	RW	2 Years LEP	1,795	19	46	21	14
9	RW	3 Years LEP	1,042	12	45	26	17
9	RW	4 Years LEP	337	8	35	32	24
9	RW	5 Years LEP	200	8	43	28	21
9	RW	6 Years LEP or More	909	4	29	38	29
9	RW	Chinese	1,478	21	42	21	15
9	RW	English	461	10	39	27	23
9	RW	Haitian-Creole	450	10	46	25	19
9	RW	Korean	76	1	45	22	32
9	RW	Russian	204	12	35	25	28
9	RW	Spanish	8,450	19	39	24	19
9	RW	Other Language	3,169	19	44	21	16
9	RW	Students without Disabilities	11,607	20	41	21	18
9	RW	Students with Disabilities	2,681	10	40	30	20

Table H.10c: Proficiency Percentages by Subgroup—Grade 9 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
9	Overall	Female	6,066	20	39	24	16
9	Overall	Male	8,222	22	39	23	17
9	Overall	American Indian or Alaska Native	43	37	42	14	7
9	Overall	Black or African-American	1,354	16	44	25	16
9	Overall	Asian	3,201	22	43	21	13
9	Overall	Native Hawaiian/Other Pacific Islander	55	29	38	15	18
9	Overall	Hispanic or Latino	8,647	22	36	24	18
9	Overall	White	974	18	40	22	20
9	Overall	Multiracial (not Hispanic origin)	14	0	29	14	57
9	Overall	NYC	10,055	20	39	23	17
9	Overall	Big 4 Cities	1,038	24	45	19	13
9	Overall	High Need Urban/Suburban	1,514	28	39	22	12
9	Overall	High Need Rural	75	15	43	28	15
9	Overall	Average Need	994	21	38	24	17
9	Overall	Low Need	462	15	30	28	27
9	Overall	Charter Schools	150	4	36	39	21
9	Overall	0 Years LEP	7,162	21	36	24	19
9	Overall	1 Years LEP	2,806	34	43	14	9
9	Overall	2 Years LEP	1,795	21	46	22	11
9	Overall	3 Years LEP	1,042	13	45	27	16
9	Overall	4 Years LEP	337	8	36	33	23
9	Overall	5 Years LEP	200	8	43	29	20
9	Overall	6 Years LEP or More	909	4	29	39	29
9	Overall	Chinese	1,478	24	43	22	12
9	Overall	English	461	11	40	28	21
9	Overall	Haitian-Creole	450	12	46	26	17
9	Overall	Korean	76	3	45	30	22
9	Overall	Russian	204	15	34	25	26
9	Overall	Spanish	8,450	22	36	24	18
9	Overall	Other Language	3,169	21	43	21	15
9	Overall	Students without Disabilities	11,607	24	39	22	16
9	Overall	Students with Disabilities	2,681	11	40	31	19

Table H.11a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 10

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
10	LS	Female	5,956	8	31	24	36
10	LS	Male	7,333	9	29	24	38
10	LS	American Indian or Alaska Native	38	8	26	32	34
10	LS	Black or African-American	1,303	3	28	32	38
10	LS	Asian	3,478	9	40	28	23
10	LS	Native Hawaiian/Other Pacific Islander	46	4	37	28	30
10	LS	Hispanic or Latino	7,435	10	26	21	43
10	LS	White	986	5	25	27	43
10	LS	Multiracial (not Hispanic origin)	3	0	33	67	0
10	LS	NYC	9,906	9	31	24	36
10	LS	Big 4 Cities	745	5	35	26	34
10	LS	High Need Urban/Suburban	1,156	9	28	24	39
10	LS	High Need Rural	53	2	17	25	57
10	LS	Average Need	875	6	26	26	41
10	LS	Low Need	433	3	22	29	46
10	LS	Charter Schools	121	3	6	27	64
10	LS	0 Years LEP	5,830	9	23	21	47
10	LS	1 Years LEP	2,972	14	44	25	17
10	LS	2 Years LEP	2,062	7	38	30	25
10	LS	3 Years LEP	1,179	3	32	30	36
10	LS	4 Years LEP	361	1	17	32	51
10	LS	5 Years LEP	227	2	15	25	58
10	LS	6 Years LEP or More	635	1	6	17	76
10	LS	Chinese	1,975	13	46	25	16
10	LS	English	408	5	26	24	45
10	LS	Haitian-Creole	454	3	24	34	39
10	LS	Korean	58	7	19	31	43
10	LS	Russian	238	8	26	23	44
10	LS	Spanish	7,250	10	26	21	43
10	LS	Other Language	2,906	4	30	30	35
10	LS	Students without Disabilities	11,363	9	33	25	33
10	LS	Students with Disabilities	1,926	5	13	22	61

Table H.11b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 10

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
10	RW	Female	5,956	13	44	26	17
10	RW	Male	7,333	15	44	24	16
10	RW	American Indian or Alaska Native	38	16	47	26	11
10	RW	Black or African-American	1,303	11	48	26	15
10	RW	Asian	3,478	15	47	23	15
10	RW	Native Hawaiian/Other Pacific Islander	46	9	61	26	4
10	RW	Hispanic or Latino	7,435	15	42	26	17
10	RW	White	986	11	42	24	23
10	RW	Multiracial (not Hispanic origin)	3	0	67	33	0
10	RW	NYC	9,906	15	45	24	16
10	RW	Big 4 Cities	745	18	52	19	11
10	RW	High Need Urban/Suburban	1,156	17	40	26	17
10	RW	High Need Rural	53	6	43	30	21
10	RW	Average Need	875	11	42	28	19
10	RW	Low Need	433	6	35	32	27
10	RW	Charter Schools	121	1	36	42	21
10	RW	0 Years LEP	5,830	14	40	28	19
10	RW	1 Years LEP	2,972	21	51	18	10
10	RW	2 Years LEP	2,062	14	50	22	14
10	RW	3 Years LEP	1,179	12	46	26	16
10	RW	4 Years LEP	361	9	39	30	22
10	RW	5 Years LEP	227	8	38	30	23
10	RW	6 Years LEP or More	635	6	32	36	26
10	RW	Chinese	1,975	17	48	21	14
10	RW	English	408	12	40	30	18
10	RW	Haitian-Creole	454	10	46	28	16
10	RW	Korean	58	14	31	38	17
10	RW	Russian	238	10	37	21	33
10	RW	Spanish	7,250	15	43	26	17
10	RW	Other Language	2,906	12	46	25	16
10	RW	Students without Disabilities	11,363	14	45	24	17
10	RW	Students with Disabilities	1,926	15	40	29	16

Table H.11c: Proficiency Percentages by Subgroup—Grade 10 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
10	Overall	Female	5,956	16	44	26	14
10	Overall	Male	7,333	17	45	24	14
10	Overall	American Indian or Alaska Native	38	16	50	26	8
10	Overall	Black or African-American	1,303	12	49	27	13
10	Overall	Asian	3,478	17	49	23	10
10	Overall	Native Hawaiian/Other Pacific Islander	46	13	61	22	4
10	Overall	Hispanic or Latino	7,435	18	42	26	15
10	Overall	White	986	12	42	25	21
10	Overall	Multiracial (not Hispanic origin)	3	0	67	33	0
10	Overall	NYC	9,906	17	45	24	14
10	Overall	Big 4 Cities	745	18	52	19	10
10	Overall	High Need Urban/Suburban	1,156	19	41	27	14
10	Overall	High Need Rural	53	6	43	30	21
10	Overall	Average Need	875	13	42	29	16
10	Overall	Low Need	433	7	37	34	22
10	Overall	Charter Schools	121	4	34	42	20
10	Overall	0 Years LEP	5,830	16	40	28	17
10	Overall	1 Years LEP	2,972	24	51	17	7
10	Overall	2 Years LEP	2,062	16	52	22	10
10	Overall	3 Years LEP	1,179	13	46	27	14
10	Overall	4 Years LEP	361	9	40	32	19
10	Overall	5 Years LEP	227	9	37	33	21
10	Overall	6 Years LEP or More	635	6	32	36	26
10	Overall	Chinese	1,975	20	51	20	8
10	Overall	English	408	12	43	29	16
10	Overall	Haitian-Creole	454	12	45	29	14
10	Overall	Korean	58	16	29	40	16
10	Overall	Russian	238	12	37	22	29
10	Overall	Spanish	7,250	18	42	26	15
10	Overall	Other Language	2,906	13	48	25	14
10	Overall	Students without Disabilities	11,363	17	45	24	14
10	Overall	Students with Disabilities	1,926	15	40	30	15

Table H.12a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 11

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
11	LS	Female	4,730	6	32	26	35
11	LS	Male	5,045	6	29	26	39
11	LS	American Indian or Alaska Native	25	8	16	52	24
11	LS	Black or African-American	936	2	25	33	40
11	LS	Asian	2,764	4	40	29	27
11	LS	Native Hawaiian/Other Pacific Islander	31	3	45	13	39
11	LS	Hispanic or Latino	5,410	9	28	23	40
11	LS	White	604	4	22	30	44
11	LS	Multiracial (not Hispanic origin)	5	0	20	20	60
11	LS	NYC	7,054	7	32	26	35
11	LS	Big 4 Cities	510	4	28	31	37
11	LS	High Need Urban/Suburban	935	7	30	24	38
11	LS	High Need Rural	72	10	21	15	54
11	LS	Average Need	745	5	24	23	47
11	LS	Low Need	412	5	28	25	42
11	LS	Charter Schools	47	0	17	32	51
11	LS	0 Years LEP	3,608	7	22	24	47
11	LS	1 Years LEP	1,898	11	42	25	22
11	LS	2 Years LEP	2,012	5	42	29	24
11	LS	3 Years LEP	1,157	3	33	32	32
11	LS	4 Years LEP	329	1	25	32	42
11	LS	5 Years LEP	221	3	19	25	53
11	LS	6 Years LEP or More	539	1	10	18	71
11	LS	Chinese	1,494	5	48	28	18
11	LS	English	345	6	22	25	47
11	LS	Haitian-Creole	329	1	23	36	40
11	LS	Korean	44	5	20	30	45
11	LS	Russian	117	3	20	30	48
11	LS	Spanish	5,263	8	28	23	40
11	LS	Other Language	2,183	3	28	30	39
11	LS	Students without Disabilities	8,506	6	33	26	35
11	LS	Students with Disabilities	1,269	7	17	24	52

Table H.12b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 11

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
11	RW	Female	4,730	9	46	28	17
11	RW	Male	5,045	10	45	26	18
11	RW	American Indian or Alaska Native	25	4	56	24	16
11	RW	Black or African-American	936	6	49	28	17
11	RW	Asian	2,764	7	47	27	19
11	RW	Native Hawaiian/Other Pacific Islander	31	6	45	26	23
11	RW	Hispanic or Latino	5,410	12	45	27	16
11	RW	White	604	8	38	27	27
11	RW	Multiracial (not Hispanic origin)	5	0	20	60	20
11	RW	NYC	7,054	10	46	27	17
11	RW	Big 4 Cities	510	8	57	22	13
11	RW	High Need Urban/Suburban	935	14	48	24	15
11	RW	High Need Rural	72	10	46	22	22
11	RW	Average Need	745	8	41	29	22
11	RW	Low Need	412	9	37	27	27
11	RW	Charter Schools	47	2	36	34	28
11	RW	0 Years LEP	3,608	11	40	30	19
11	RW	1 Years LEP	1,898	13	49	23	15
11	RW	2 Years LEP	2,012	9	52	23	16
11	RW	3 Years LEP	1,157	7	48	27	18
11	RW	4 Years LEP	329	6	49	32	14
11	RW	5 Years LEP	221	9	44	28	20
11	RW	6 Years LEP or More	539	5	42	32	21
11	RW	Chinese	1,494	8	47	26	19
11	RW	English	345	10	38	28	24
11	RW	Haitian-Creole	329	6	45	32	17
11	RW	Korean	44	9	18	41	32
11	RW	Russian	117	6	29	29	36
11	RW	Spanish	5,263	12	46	27	15
11	RW	Other Language	2,183	7	47	26	19
11	RW	Students without Disabilities	8,506	8	46	27	18
11	RW	Students with Disabilities	1,269	20	43	25	13

Table H.12c: Proficiency Percentages by Subgroup—Grade 11 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
11	Overall	Female	4,730	11	49	27	13
11	Overall	Male	5,045	12	47	26	15
11	Overall	American Indian or Alaska Native	25	8	52	28	12
11	Overall	Black or African-American	936	6	51	29	14
11	Overall	Asian	2,764	9	53	26	12
11	Overall	Native Hawaiian/Other Pacific Islander	31	6	45	29	19
11	Overall	Hispanic or Latino	5,410	15	45	26	14
11	Overall	White	604	8	41	28	23
11	Overall	Multiracial (not Hispanic origin)	5	0	40	40	20
11	Overall	NYC	7,054	12	48	26	13
11	Overall	Big 4 Cities	510	10	56	23	11
11	Overall	High Need Urban/Suburban	935	15	49	23	13
11	Overall	High Need Rural	72	11	46	22	21
11	Overall	Average Need	745	10	41	30	19
11	Overall	Low Need	412	10	40	29	21
11	Overall	Charter Schools	47	2	40	34	23
11	Overall	0 Years LEP	3,608	13	41	30	16
11	Overall	1 Years LEP	1,898	17	52	21	10
11	Overall	2 Years LEP	2,012	11	55	23	11
11	Overall	3 Years LEP	1,157	8	52	26	14
11	Overall	4 Years LEP	329	6	50	33	11
11	Overall	5 Years LEP	221	10	43	28	19
11	Overall	6 Years LEP or More	539	5	43	32	21
11	Overall	Chinese	1,494	10	57	24	9
11	Overall	English	345	11	39	29	21
11	Overall	Haitian-Creole	329	6	46	34	14
11	Overall	Korean	44	9	27	39	25
11	Overall	Russian	117	6	32	29	32
11	Overall	Spanish	5,263	15	46	26	13
11	Overall	Other Language	2,183	7	50	27	16
11	Overall	Students without Disabilities	8,506	11	49	26	14
11	Overall	Students with Disabilities	1,269	20	43	26	12

Table H.13a: Proficiency Percentages by Subgroup—Listening and Speaking, Grade 12

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
12	LS	Female	3,559	5	32	22	41
12	LS	Male	3,793	6	29	20	45
12	LS	American Indian or Alaska Native	9	11	33	0	56
12	LS	Black or African-American	655	2	27	24	46
12	LS	Asian	2,035	4	38	24	34
12	LS	Native Hawaiian/Other Pacific Islander	15	0	27	27	47
12	LS	Hispanic or Latino	4,112	6	29	19	46
12	LS	White	523	8	21	20	51
12	LS	Multiracial (not Hispanic origin)	3	0	33	33	33
12	LS	NYC	5,465	6	33	20	41
12	LS	Big 4 Cities	382	6	32	24	37
12	LS	High Need Urban/Suburban	620	5	24	22	49
12	LS	High Need Rural	25	0	24	8	68
12	LS	Average Need	522	3	21	22	54
12	LS	Low Need	288	2	20	29	48
12	LS	Charter Schools	50	0	32	16	52
12	LS	0 Years LEP	3,135	7	27	17	49
12	LS	1 Years LEP	871	6	40	23	31
12	LS	2 Years LEP	1,240	5	40	24	31
12	LS	3 Years LEP	1,023	4	36	25	35
12	LS	4 Years LEP	332	2	23	24	52
12	LS	5 Years LEP	232	4	24	25	47
12	LS	6 Years LEP or More	510	3	15	18	64
12	LS	Chinese	1,102	4	46	23	27
12	LS	English	224	6	24	22	47
12	LS	Haitian-Creole	243	2	27	21	49
12	LS	Korean	67	6	27	27	40
12	LS	Russian	105	5	20	26	50
12	LS	Spanish	4,020	6	29	19	46
12	LS	Other Language	1,591	5	26	24	45
12	LS	Students without Disabilities	5,891	2	31	22	45
12	LS	Students with Disabilities	1,461	18	32	15	35

Table H.13b: Proficiency Percentages by Subgroup—Reading and Writing, Grade 12

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
12	RW	Female	3,559	11	46	24	19
12	RW	Male	3,793	12	46	23	19
12	RW	American Indian or Alaska Native	9	11	33	22	33
12	RW	Black or African-American	655	9	48	23	21
12	RW	Asian	2,035	8	47	23	22
12	RW	Native Hawaiian/Other Pacific Islander	15	0	53	27	20
12	RW	Hispanic or Latino	4,112	14	46	24	16
12	RW	White	523	14	35	25	26
12	RW	Multiracial (not Hispanic origin)	3	0	33	67	0
12	RW	NYC	5,465	13	45	23	19
12	RW	Big 4 Cities	382	13	53	21	13
12	RW	High Need Urban/Suburban	620	10	49	27	15
12	RW	High Need Rural	25	0	52	20	28
12	RW	Average Need	522	5	43	29	23
12	RW	Low Need	288	4	39	26	32
12	RW	Charter Schools	50	4	58	22	16
12	RW	0 Years LEP	3,135	17	42	22	18
12	RW	1 Years LEP	871	7	49	22	21
12	RW	2 Years LEP	1,240	8	48	25	20
12	RW	3 Years LEP	1,023	7	53	22	18
12	RW	4 Years LEP	332	6	45	30	19
12	RW	5 Years LEP	232	8	47	26	19
12	RW	6 Years LEP or More	510	7	44	27	22
12	RW	Chinese	1,102	7	46	23	24
12	RW	English	224	11	38	25	26
12	RW	Haitian-Creole	243	9	46	20	25
12	RW	Korean	67	10	31	24	34
12	RW	Russian	105	8	29	29	35
12	RW	Spanish	4,020	14	46	24	16
12	RW	Other Language	1,591	10	47	23	20
12	RW	Students without Disabilities	5,891	4	48	26	22
12	RW	Students with Disabilities	1,461	42	36	14	8

Table H.13c: Proficiency Percentages by Subgroup—Grade 12 Overall

				Proficiency Levels			
Grade	**Test**	**Subgroup**	**N-Count**	**1**	**2**	**3**	**4**
12	Overall	Female	3,559	12	48	25	15
12	Overall	Male	3,793	14	49	22	15
12	Overall	American Indian or Alaska Native	9	11	44	11	33
12	Overall	Black or African-American	655	9	49	23	18
12	Overall	Asian	2,035	9	53	23	15
12	Overall	Native Hawaiian/Other Pacific Islander	15	0	68	20	13
12	Overall	Hispanic or Latino	4,112	15	47	23	14
12	Overall	White	523	15	38	25	23
12	Overall	Multiracial (not Hispanic origin)	3	0	33	67	0
12	Overall	NYC	5,465	14	48	22	15
12	Overall	Big 4 Cities	382	13	55	21	10
12	Overall	High Need Urban/Suburban	620	10	50	27	13
12	Overall	High Need Rural	25	0	52	24	24
12	Overall	Average Need	522	5	46	30	19
12	Overall	Low Need	288	4	42	30	25
12	Overall	Charter Schools	50	4	66	18	12
12	Overall	0 Years LEP	3,135	18	43	23	16
12	Overall	1 Years LEP	871	9	54	21	15
12	Overall	2 Years LEP	1,240	9	53	24	13
12	Overall	3 Years LEP	1,023	8	56	22	14
12	Overall	4 Years LEP	332	6	47	31	16
12	Overall	5 Years LEP	232	9	50	27	15
12	Overall	6 Years LEP or More	510	8	44	28	19
12	Overall	Chinese	1,102	9	55	22	14
12	Overall	English	224	12	40	28	21
12	Overall	Haitian-Creole	243	9	47	22	22
12	Overall	Korean	67	10	34	34	21
12	Overall	Russian	105	10	34	30	27
12	Overall	Spanish	4,020	15	48	23	14
12	Overall	Other Language	1,591	10	49	24	17
12	Overall	Students without Disabilities	5,891	5	52	26	17
12	Overall	Students with Disabilities	1,461	42	36	14	7

APPENDIX I: EXIT RATE BY SUBGROUP

Table I.1: Exit Rate by Subgroup—Grade K

Grade	Subgroup	N-Count	Percent
K	Female	12,414	47
K	Male	13,852	53
K	American Indian or Alaska Native	75	0
K	Black or African-American	1,002	4
K	Asian	5,691	22
K	Native Hawaiian/Other Pacific Islander	129	0
K	Hispanic or Latino	16,716	64
K	White	2,582	10
K	Multiracial (not Hispanic origin)	71	0
K	NYC	15,708	60
K	Big 4 Cities	1,087	4
K	High Need Urban/Suburban	3,909	15
K	High Need Rural	204	1
K	Average Need	3,220	12
K	Low Need	1,425	5
K	Charter Schools	713	3
K	0 Years LEP	15,404	59
K	1 Years LEP	10,240	39
K	2 Years LEP	489	2
K	3 Years LEP	28	0
K	4 Years LEP	3	0
K	5 Years LEP	5	0
K	6 Years LEP or More	0	0
K	Chinese	2,909	11
K	English	2,614	10
K	Haitian-Creole	200	1
K	Korean	167	1
K	Russian	541	2
K	Spanish	15,321	58
K	Other Language	4,514	17
K	Students without Disabilities	23,235	88
K	Students with Disabilities	3,031	12

Table I.2: Exit Rate by Subgroup—Grade 1

Grade	Subgroup	N-Count	Percent
1	Female	11,986	47
1	Male	13,690	53
1	American Indian or Alaska Native	61	0
1	Black or African-American	1,131	4
1	Asian	5,252	20
1	Native Hawaiian/Other Pacific Islander	113	0
1	Hispanic or Latino	16,667	65
1	White	2,417	9
1	Multiracial (not Hispanic origin)	35	0
1	NYC	16,153	63
1	Big 4 Cities	1,198	5
1	High Need Urban/Suburban	3,360	13
1	High Need Rural	193	1
1	Average Need	2,918	11
1	Low Need	1,255	5
1	Charter Schools	599	2
1	0 Years LEP	14,263	56
1	1 Years LEP	3,368	13
1	2 Years LEP	7,012	27
1	3 Years LEP	948	4
1	4 Years LEP	27	0
1	5 Years LEP	6	0
1	6 Years LEP or More	9	0
1	Chinese	2,673	10
1	English	1,983	8
1	Haitian-Creole	235	1
1	Korean	140	1
1	Russian	546	2
1	Spanish	15,651	61
1	Other Language	4,448	17
1	Students without Disabilities	22,237	87
1	Students with Disabilities	3,439	13

Table I.3: Exit Rate by Subgroup—Grade 2

Grade	Subgroup	N-Count	Percent
2	Female	9,497	45
2	Male	11,662	55
2	American Indian or Alaska Native	40	0
2	Black or African-American	1,066	5
2	Asian	4,252	20
2	Native Hawaiian/Other Pacific Islander	95	0
2	Hispanic or Latino	13,943	66
2	White	1,737	8
2	Multiracial (not Hispanic origin)	26	0
2	NYC	13,574	64
2	Big 4 Cities	1,166	6
2	High Need Urban/Suburban	2,659	13
2	High Need Rural	145	1
2	Average Need	2,228	11
2	Low Need	993	5
2	Charter Schools	394	2
2	0 Years LEP	11,201	53
2	1 Years LEP	2,226	11
2	2 Years LEP	2,157	10
2	3 Years LEP	4,706	22
2	4 Years LEP	774	4
2	5 Years LEP	41	0
2	6 Years LEP or More	16	0
2	Chinese	2,026	10
2	English	1,218	6
2	Haitian-Creole	248	1
2	Korean	108	1
2	Russian	374	2
2	Spanish	13,311	63
2	Other Language	3,874	18
2	Students without Disabilities	17,227	81
2	Students with Disabilities	3,932	19

Table I.4: Exit Rate by Subgroup—Grade 3

Grade	Subgroup	N-Count	Percent
3	Female	8,643	45
3	Male	10,381	55
3	American Indian or Alaska Native	32	0
3	Black or African-American	1,132	6
3	Asian	3,549	19
3	Native Hawaiian/Other Pacific Islander	66	0
3	Hispanic or Latino	12,747	67
3	White	1,482	8
3	Multiracial (not Hispanic origin)	16	0
3	NYC	12,386	65
3	Big 4 Cities	1,104	6
3	High Need Urban/Suburban	2,416	13
3	High Need Rural	128	1
3	Average Need	1,900	10
3	Low Need	793	4
3	Charter Schools	297	2
3	0 Years LEP	10,189	54
3	1 Years LEP	1,815	10
3	2 Years LEP	1,516	8
3	3 Years LEP	1,544	8
3	4 Years LEP	3,259	17
3	5 Years LEP	625	3
3	6 Years LEP or More	56	0
3	Chinese	1,607	8
3	English	919	5
3	Haitian-Creole	283	1
3	Korean	67	0
3	Russian	289	2
3	Spanish	12,297	65
3	Other Language	3,562	19
3	Students without Disabilities	15,007	79
3	Students with Disabilities	4,017	21

Table I.5: Exit Rate by Subgroup—Grade 4

Grade	Subgroup	N-Count	Percent
4	Female	7,981	45
4	Male	9,872	55
4	American Indian or Alaska Native	37	0
4	Black or African-American	1,074	6
4	Asian	3,156	18
4	Native Hawaiian/Other Pacific Islander	42	0
4	Hispanic or Latino	12,180	68
4	White	1,346	8
4	Multiracial (not Hispanic origin)	18	0
4	NYC	11,916	67
4	Big 4 Cities	1,120	6
4	High Need Urban/Suburban	2,145	12
4	High Need Rural	109	1
4	Average Need	1,674	9
4	Low Need	682	4
4	Charter Schools	207	1
4	0 Years LEP	9,894	55
4	1 Years LEP	1,839	10
4	2 Years LEP	1,335	7
4	3 Years LEP	991	6
4	4 Years LEP	732	4
4	5 Years LEP	2,447	14
4	6 Years LEP or More	594	3
4	Chinese	1,376	8
4	English	848	5
4	Haitian-Creole	296	2
4	Korean	86	0
4	Russian	270	2
4	Spanish	11,689	65
4	Other Language	3,288	18
4	Students without Disabilities	13,425	75
4	Students with Disabilities	4,428	25

Table I.6: Exit Rate by Subgroup—Grade 5

Grade	Subgroup	N-Count	Percent
5	Female	6,957	45
5	Male	8,502	55
5	American Indian or Alaska Native	28	0
5	Black or African-American	1,033	7
5	Asian	2,710	18
5	Native Hawaiian/Other Pacific Islander	53	0
5	Hispanic or Latino	10,441	68
5	White	1,178	8
5	Multiracial (not Hispanic origin)	16	0
5	NYC	10,445	68
5	Big 4 Cities	1,027	7
5	High Need Urban/Suburban	1,726	11
5	High Need Rural	94	1
5	Average Need	1,311	8
5	Low Need	565	4
5	Charter Schools	291	2
5	0 Years LEP	8,482	55
5	1 Years LEP	1,879	12
5	2 Years LEP	1,350	9
5	3 Years LEP	767	5
5	4 Years LEP	460	3
5	5 Years LEP	511	3
5	6 Years LEP or More	1,982	13
5	Chinese	1,092	7
5	English	645	4
5	Haitian-Creole	280	2
5	Korean	80	1
5	Russian	245	2
5	Spanish	10,129	66
5	Other Language	2,988	19
5	Students without Disabilities	11,329	73
5	Students with Disabilities	4,130	27

Table I.7: Exit Rate by Subgroup—Grade 6

Grade	Subgroup	N-Count	Percent
6	Female	5,643	43
6	Male	7,521	57
6	American Indian or Alaska Native	23	0
6	Black or African-American	992	8
6	Asian	2,364	18
6	Native Hawaiian/Other Pacific Islander	57	0
6	Hispanic or Latino	8,689	66
6	White	1,030	8
6	Multiracial (not Hispanic origin)	9	0
6	NYC	9,048	69
6	Big 4 Cities	952	7
6	High Need Urban/Suburban	1,362	10
6	High Need Rural	96	1
6	Average Need	1,060	8
6	Low Need	409	3
6	Charter Schools	237	2
6	0 Years LEP	6,898	52
6	1 Years LEP	1,855	14
6	2 Years LEP	1,359	10
6	3 Years LEP	751	6
6	4 Years LEP	331	3
6	5 Years LEP	276	2
6	6 Years LEP or More	1,672	13
6	Chinese	995	8
6	English	522	4
6	Haitian-Creole	338	3
6	Korean	65	0
6	Russian	235	2
6	Spanish	8,433	64
6	Other Language	2,576	20
6	Students without Disabilities	9,642	73
6	Students with Disabilities	3,522	27

Table I.8: Exit Rate by Subgroup—Grade 7

Grade	Subgroup	N-Count	Percent
7	Female	5,415	43
7	Male	7,194	57
7	American Indian or Alaska Native	32	0
7	Black or African-American	1,008	8
7	Asian	2,229	18
7	Native Hawaiian/Other Pacific Islander	45	0
7	Hispanic or Latino	8,428	67
7	White	853	7
7	Multiracial (not Hispanic origin)	14	0
7	NYC	8,853	70
7	Big 4 Cities	898	7
7	High Need Urban/Suburban	1,300	10
7	High Need Rural	78	1
7	Average Need	905	7
7	Low Need	399	3
7	Charter Schools	176	1
7	0 Years LEP	6,512	52
7	1 Years LEP	1,972	16
7	2 Years LEP	1,324	11
7	3 Years LEP	823	7
7	4 Years LEP	292	2
7	5 Years LEP	224	2
7	6 Years LEP or More	1,443	11
7	Chinese	979	8
7	English	432	3
7	Haitian-Creole	331	3
7	Korean	60	0
7	Russian	174	1
7	Spanish	8,210	65
7	Other Language	2,423	19
7	Students without Disabilities	9,259	73
7	Students with Disabilities	3,350	27

Table I.9: Exit Rate by Subgroup—Grade 8

Grade	Subgroup	N-Count	Percent
8	Female	5,478	44
8	Male	7,028	56
8	American Indian or Alaska Native	38	0
8	Black or African-American	954	8
8	Asian	2,517	20
8	Native Hawaiian/Other Pacific Islander	66	1
8	Hispanic or Latino	7,962	64
8	White	960	8
8	Multiracial (not Hispanic origin)	9	0
8	NYC	9,059	72
8	Big 4 Cities	869	7
8	High Need Urban/Suburban	1,134	9
8	High Need Rural	72	1
8	Average Need	857	7
8	Low Need	421	3
8	Charter Schools	94	1
8	0 Years LEP	6,427	51
8	1 Years LEP	2,014	16
8	2 Years LEP	1,572	13
8	3 Years LEP	864	7
8	4 Years LEP	324	3
8	5 Years LEP	217	2
8	6 Years LEP or More	1,073	9
8	Chinese	1,177	9
8	English	389	3
8	Haitian-Creole	304	2
8	Korean	80	1
8	Russian	236	2
8	Spanish	7,797	62
8	Other Language	2,523	20
8	Students without Disabilities	9,355	75
8	Students with Disabilities	3,151	25

Table I.10: Exit Rate by Subgroup—Grade 9

Grade	Subgroup	N-Count	Percent
9	Female	6,066	42
9	Male	8,222	58
9	American Indian or Alaska Native	43	0
9	Black or African-American	1,354	9
9	Asian	3,201	22
9	Native Hawaiian/Other Pacific Islander	55	0
9	Hispanic or Latino	8,647	61
9	White	974	7
9	Multiracial (not Hispanic origin)	14	0
9	NYC	10,055	70
9	Big 4 Cities	1,038	7
9	High Need Urban/Suburban	1,514	11
9	High Need Rural	75	1
9	Average Need	994	7
9	Low Need	462	3
9	Charter Schools	150	1
9	0 Years LEP	7,162	50
9	1 Years LEP	2,806	20
9	2 Years LEP	1,795	13
9	3 Years LEP	1,042	7
9	4 Years LEP	337	2
9	5 Years LEP	200	1
9	6 Years LEP or More	908	6
9	Chinese	1,478	10
9	English	461	3
9	Haitian-Creole	450	3
9	Korean	76	1
9	Russian	204	1
9	Spanish	8,450	59
9	Other Language	3,169	22
9	Students without Disabilities	11,607	81
9	Students with Disabilities	2,681	19

Table I.11: Exit Rate by Subgroup—Grade 10

Grade	Subgroup	N-Count	Percent
10	Female	5,956	45
10	Male	7,333	55
10	American Indian or Alaska Native	38	0
10	Black or African-American	1,303	10
10	Asian	3,478	26
10	Native Hawaiian/Other Pacific Islander	46	0
10	Hispanic or Latino	7,435	56
10	White	986	7
10	Multiracial (not Hispanic origin)	3	0
10	NYC	9,906	75
10	Big 4 Cities	745	6
10	High Need Urban/Suburban	1,156	9
10	High Need Rural	53	0
10	Average Need	875	7
10	Low Need	433	3
10	Charter Schools	121	1
10	0 Years LEP	5,830	44
10	1 Years LEP	2,972	22
10	2 Years LEP	2,062	16
10	3 Years LEP	1,179	9
10	4 Years LEP	361	3
10	5 Years LEP	227	2
10	6 Years LEP or More	635	5
10	Chinese	1,975	15
10	English	408	3
10	Haitian-Creole	454	3
10	Korean	58	0
10	Russian	238	2
10	Spanish	7,250	55
10	Other Language	2,906	22
10	Students without Disabilities	11,363	86
10	Students with Disabilities	1,926	14

Table I.12: Exit Rate by Subgroup—Grade 11

Grade	Subgroup	N-Count	Percent
11	Female	4,730	48
11	Male	5,045	52
11	American Indian or Alaska Native	25	0
11	Black or African-American	936	10
11	Asian	2,764	28
11	Native Hawaiian/Other Pacific Islander	31	0
11	Hispanic or Latino	5,410	55
11	White	604	6
11	Multiracial (not Hispanic origin)	5	0
11	NYC	7,054	72
11	Big 4 Cities	510	5
11	High Need Urban/Suburban	935	10
11	High Need Rural	72	1
11	Average Need	745	8
11	Low Need	412	4
11	Charter Schools	47	0
11	0 Years LEP	3,608	37
11	1 Years LEP	1,898	19
11	2 Years LEP	2,012	21
11	3 Years LEP	1,157	12
11	4 Years LEP	329	3
11	5 Years LEP	221	2
11	6 Years LEP or More	539	6
11	Chinese	1,494	15
11	English	345	4
11	Haitian-Creole	329	3
11	Korean	44	0
11	Russian	117	1
11	Spanish	5,263	54
11	Other Language	2,183	22
11	Students without Disabilities	8,506	87
11	Students with Disabilities	1,269	13

Table I.13: Exit Rate by Subgroup—Grade 12

Grade	Subgroup	N-Count	Percent
12	Female	3,559	48
12	Male	3,793	52
12	American Indian or Alaska Native	9	0
12	Black or African-American	655	9
12	Asian	2,035	28
12	Native Hawaiian/Other Pacific Islander	15	0
12	Hispanic or Latino	4,112	56
12	White	523	7
12	Multiracial (not Hispanic origin)	3	0
12	NYC	5,465	74
12	Big 4 Cities	382	5
12	High Need Urban/Suburban	620	8
12	High Need Rural	25	0
12	Average Need	522	7
12	Low Need	288	4
12	Charter Schools	50	1
12	0 Years LEP	3,135	43
12	1 Years LEP	871	12
12	2 Years LEP	1,240	17
12	3 Years LEP	1,023	14
12	4 Years LEP	332	5
12	5 Years LEP	232	3
12	6 Years LEP or More	509	7
12	Chinese	1,102	15
12	English	224	3
12	Haitian-Creole	243	3
12	Korean	67	1
12	Russian	105	1
12	Spanish	4,020	55
12	Other Language	1,591	22
12	Students without Disabilities	5,891	80
12	Students with Disabilities	1,461	20

APPENDIX J: DIF STATISTICS

Table J.1: Results of DIF Analyses

Grade Band	Modality	Item Seq. Number	DIF			
			Male/ Female	Asian/ Other	Hispanic/ Other	White/ Other
K–1	Listening	1	A	B+	A	A
		2	A	B+	A	A
		9	A	B+	B-	A
		27	A	B-	B+	A
		28	A	A	B+	A
	Speaking	1	A	B-	A	A
		2	A	B-	B+	A
		3	A	B-	A	A
		4	A	B-	A	A
		8	A	B-	A	A
		9	A	B-	A	A
		10	A	B-	A	A
		11	A	B-	B+	A
		12	A	B-	A	A
		14	A	B-	B+	A
		17	A	B-	A	A
		18	A	B-	A	A
	Writing	2	A	B-	A	A
		9	A	B+	A	A
		10	A	B+	A	A
		12	A	B+	B-	A
		13	A	B+	A	A
		14	A	C+	B-	A
		15	A	B+	B-	A
2–4	Listening	6	A	A	B+	A
		10	A	A	A	B
		12	A	B+	A	A
		23	A	B+	A	A
		24	A	A	A	B
	Speaking	1	A	B-	A	A
		2	A	B-	B+	A
		3	A	B-	A	A
		4	A	C-	B+	A
		6	A	B-	A	A
		18	A	B-	A	A
	Writing	3	A	B-	A	A
		5	A	B-	B+	A
		11	A	C-	B+	A

Table J.1 (continued): Results of DIF Analyses

Grade Band	Modality	Item Seq. Number	DIF			
			Male/ Female	Asian/ Other	Hispanic/ Other	White/ Other
5–6	Listening	15	A	B+	A	A
		16	A	B-	B+	A
	Speaking	1	A	B-	A	A
		2	A	B-	A	A
		5	A	B-	A	A
	Reading	4	A	B-	B+	B
		10	A	B+	B-	A
		26	A	B-	A	A
	Writing	2	A	A	B+	B
		3	A	B-	B+	A
		6	A	C-	B+	A
		8	A	B-	B+	A
		11	A	B-	A	A
7–8	Listening	1	A	A	B+	A
		2	A	C-	C+	A
		9	A	B+	A	A
		23	A	B+	A	A
	Speaking	2	A	B-	A	A
		3	A	B-	A	A
		14	A	B-	A	A
	Reading	3	A	C-	C+	C
		13	A	B+	A	A
		15	A	B+	A	B
		16	A	B+	A	A
		21	A	A	B-	A
	Writing	1	A	B+	A	A
		7	A	B-	A	A
		10	A	B-	A	A
		19	B	A	A	A

Table J.1 (continued): Results of DIF Analyses

Grade Band	Modality	Item Seq. Number	DIF			
			Male/ Female	Asian/ Other	Hispanic/ Other	White/ Other
9–12	Listening	1	A	C-	C+	A
		7	A	B-	C+	A
		14	A	B+	A	A
		17	A	B+	A	A
		22	A	C+	B-	A
	Reading	2	A	C+	A	A
		3	A	C-	A	A
		4	A	B-	B+	B
		6	A	B+	B-	A
		16	B	A	A	A
		21	A	B+	A	A
		23	A	B-	C+	A
	Writing	1	A	A	A	C
		2	A	C-	C+	A
		3	A	B+	B-	A
		9	A	A	A	B
		13	A	C+	C-	A
		19	B	A	A	A

Note: "+" indicates in favor of Males or Other groups, and "-" 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.

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

Andrich, A. & Luo, G. (2004). *Modern measurement and analysis in social science*. Perth, Western Australia: Murdoch University.

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.

Dorans, N. J., & Schmitt, A. P. (1991). *Constructed response and differential item functioning: A pragmatic approach*. ETS Research Report No. RR 91-47. Princeton, NJ: Educational Testing Service.

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. & Béguin, 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.

Linacre, J. M. (2009). *A user's guide to WINSTEPS: 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.

No Child Left Behind Act of 2001, Pub. L. No. 107–110, 115 Stat.1425 (2002).

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.

Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement, 30*, 233–251.