

World-Class Instructional Design and Assessment



**Annual Technical Report for
ALTERNATE ACCESS for ELLs[®] English Language Proficiency
Test, Series 502, 2020-2021 Administration**

Annual Technical Report No. 9

Prepared by:

Psychometrics team

WIDA

May 11, 2022

Note: This version of ATR is redacted. Therefore, some of the tables that includes item information are removed.

Table of Contents

1.	<i>Description of Alternate ACCESS for ELLs English Language Proficiency Test</i>	<i>1</i>
1.1.	<i>Purpose of Alternate ACCESS for ELLs.....</i>	<i>1</i>
1.2.	<i>Format of Alternate ACCESS</i>	<i>1</i>
1.2.1	Integration with the Standards	1
1.2.1.	Grade-level Clusters	2
1.2.2.	Language Domains	2
1.2.3.	Language Proficiency Levels	3
1.3.	<i>Test Development.....</i>	<i>4</i>
1.3.1.	Item Development	4
1.3.2.	Field Test	4
1.3.3.	Scaling	4
1.3.4.	Standard Setting.....	6
1.4.	<i>Reporting of Results</i>	<i>8</i>
1.4.1.	Scale Scores	8
1.4.2.	Language Proficiency Level Scores	8
1.5.	<i>Test Administration</i>	<i>10</i>
1.5.1.	Test Administrator Training	10
1.5.2.	Test Security	10
1.5.3.	Test Accommodations	10
1.6.	<i>Scoring.....</i>	<i>10</i>
1.6.1.	Listening and Reading	11
1.6.2.	Writing	11
1.6.3.	Speaking	13
2	<i>An Assessment Use Argument for Alternate ACCESS for ELLs: Focus on Assessment Records</i>	<i>15</i>
2.1	<i>The Generic Validation Framework for Alternate ACCESS</i>	<i>17</i>
2.2	<i>Focus on Assessment Records</i>	<i>18</i>
2.2.1	Breakdown of Claims for the Assessment Records Produced in the Alternate ACCESS for ELLs Assessment Program.....	19
2.3	<i>Evidence for Assessment Records Claims of Alternate ACCESS for ELLs</i>	<i>21</i>
2.4	<i>Summary of Assessment Records Claims, Actions, and Evidence</i>	<i>27</i>
2.5	<i>Visual Guide to Tables and Figures</i>	<i>28</i>
2.5.1	Chapter 4 Visual Guide to Tables and Figures	30
2.5.2	Chapter 6 Visual Guide to Tables and Figures.....	31
3.	<i>Descriptions of Student Results</i>	<i>32</i>

3.1	Participation	32
3.2	Scale Score Results	33
3.2.1	Mean Scale Scores Across Domain and Composite Scores	33
3.2.2	Correlations	33
3.3	Proficiency Level Results	33
3.4	Participation by Disability	34
4.	Student Results	35
4.1	Students excluded from Analysis	35
4.1.1	Out-of-grade-level Test Administration	35
4.2	Participation by Grade-level Cluster	36
4.2.1	Participation by Grade-level Cluster by State	36
4.2.2	Participation by Grade-level Cluster by Gender.....	37
4.2.3	Participation by Grade-level Cluster by Ethnicity.....	37
4.3	Participation by Grade	38
4.3.1	Participation by Grade by State	38
4.3.2	Participation by Grade by Gender	39
4.3.3	Participation by Grade by Ethnicity	39
4.4	Participation by Domain	40
4.4.1	Participation by Grade-level Cluster by Domain.....	40
4.4.2	Participation by Grade by Domain	40
4.5	Scale Scores by Domain and Composite	41
4.5.1	Mean Scale Scores by Domain and Composite	41
4.6	Scale Scores by Grade-level Cluster	43
4.6.1	Mean Scale Scores by Gender	43
4.6.2	Mean Scale Scores by Ethnicity	45
4.7	Scale Scores by Grade	49
4.7.1	Mean Scale Scores by Gender	49
4.7.2	Mean Scale Scores by Ethnicity	55
4.8	Correlations among Scale Scores by Grade-level Cluster	67
4.8.1	Correlations among Scale Scores: Grade-level Cluster 1-2.....	67
4.8.2	Correlations among Scale Scores: Grade-level Cluster 3-5.....	67
4.8.3	Correlations among Scale Scores: Grade-level Cluster 6-8.....	68
4.8.4	Correlations among Scale Scores: Grade-level Cluster 9-12.....	68
4.9	Proficiency Levels	69
4.9.1	Proficiency Level by Grade-level Cluster	69
4.9.2	Proficiency Level by Grade.....	72

4.10	<i>Participation by Disability</i>	76
4.10.1	Participation by Disability	76
5.	<i>Analyses of Test Forms: Overview</i>	78
5.1	<i>Background</i>	78
5.1.1	Measurement Models Used	78
5.1.2	Sampling.....	79
5.1.3	Scaling	80
5.1.4	DIF Analyses	80
5.1.5	Reliability of Composites	81
5.1.6	Accuracy and Consistency of Classification	82
5.2	<i>Descriptions</i>	84
5.2.1	Raw Score Information (Figure A and Table A).....	84
5.2.2	Scale Score Information (Figure B and Table B).....	84
5.2.3	Proficiency Level Information (Figure C and Table C).....	85
5.2.4	Equating Summary Table (Table D).....	85
5.2.5	Reliability (Table E).....	85
5.2.6	Test Characteristic Curve (Figure D)	87
5.2.7	Test Information Function (Figure E).....	87
5.2.8	Item Analysis Summary (Table F)	87
5.2.9	Complete Item Analysis Table (Table G).....	88
5.2.10	Complete Raw Score to Scale Score Conversion Chart (Table H).....	88
5.2.11	Raw Score to Proficiency Level Score Conversion Table (Table I).....	89
5.2.12	Accuracy and Consistency of Classification Table (Table J).....	89
5.2.13	Conditional Standard Error of Measurement for Composite Figure (Figure F).....	90
6.	<i>Analyses of Test Forms: Results</i>	91
6.1	<i>Grades: 1-2</i>	91
6.1.1	Listening 1-2.....	91
6.1.2	Reading 1-2.....	97
6.1.3	Speaking 1-2.....	103
6.1.4	Writing 1-2	109
6.1.5	Oral Language Composite 1-2.....	115
6.1.6	Literacy Composite 1-2	118
6.1.7	Comprehension Composite 1-2	122
6.1.8	Overall Composite 1-2.....	126
6.2	<i>Grades: 3-5</i>	130
6.2.1	Listening 3-5.....	130
6.2.2	Reading 3-5.....	136
6.2.3	Speaking 3-5	142
6.2.4	Writing 3-5	147
6.2.5	Oral Language Composite 3-5.....	153
6.2.6	Literacy Composite 3-5	156
6.2.7	Comprehension Composite 3-5	159
6.2.8	Overall Composite 3-5.....	163
6.3	<i>Grades: 6-8</i>	166

6.3.1	Listening 6-8.....	166
6.3.2	Reading 6-8.....	172
6.3.3	Speaking 6-8.....	178
6.3.4	Writing 6-8.....	183
6.3.5	Oral Language Composite 6-8.....	189
6.3.6	Literacy Composite 6-8.....	192
6.3.7	Comprehension Composite 6-8.....	196
6.3.8	Overall Composite 6-8.....	200
6.4	Grades: 9-12.....	203
6.4.1	Listening 9-12.....	203
6.4.2	Reading 9-12.....	209
6.4.3	Speaking 9-12.....	215
6.4.4	Writing 9-12.....	220
6.4.5	Oral Language Composite 9-12.....	226
6.4.6	Literacy Composite 9-12.....	230
6.4.7	Comprehension Composite 9-12.....	234
6.4.8	Overall Composite 9-12.....	237
	References.....	240

1. Description of Alternate ACCESS for ELLs English Language Proficiency Test

1.1. Purpose of Alternate ACCESS for ELLs

The purpose of Alternate ACCESS for ELLs (hereafter, Alternate ACCESS) is to assess the developing English language proficiency (ELP) of English language learners (ELLs) with the most significant cognitive disabilities in Grades 1–12 in the states of the WIDA consortium. The assessment is rooted in the *Alternate English Language Development (ELD) Standards for English Language Learners with Significant Cognitive Disabilities* of the WIDA Consortium. Alternate ACCESS is a first of its kind attempt made by WIDA to assess ELP for ELLs with the most significant cognitive disabilities. As such, the assessment continues to be refined to clarify the construct and to develop a test design that better reflects the diversity of student language use within this population.

The WIDA ELD Standards are corresponded to WIDA Consortium state academic content standards and form the core of the WIDA Consortium’s approach to instructing and testing academic English for ELLs with significant cognitive disabilities. Alternate ACCESS, which was developed based on the WIDA ELD Standards, may thus be described as a standards-based ELP test designed to measure proficiency for ELLs with the most significant cognitive disabilities. It assesses social and instructional English as well as the language associated with Language Arts, Mathematics, and Science within the school context across the four language domains of Listening, Reading, Writing, and Speaking.

Major purposes of Alternate ACCESS include¹:

- To meet federal accountability requirements for assessment practice for ELLs and students with the most significant cognitive disabilities as specified in The Every Student Succeeds Act (ESSA; 2015) and the Individuals with Disabilities Education Act (IDEA; 2004)
- To provide educators with a measure sensitive to ELP growth of ELLs with significant cognitive disabilities

1.2. Format of Alternate ACCESS

1.2.1 Integration with the Standards

The design of Alternate ACCESS is built upon the foundational WIDA ELD Standards. The four WIDA ELD Standards represented are:

Standard 1—Social and Instructional Language

ELLs communicate in English for **social and instructional** purposes in the school setting.

Standard 2— Language of Language Arts

¹ From the WIDA Alternate ACCESS website, <https://wida.wisc.edu/assess/alt-access>

ELLs communicate information, ideas, and concepts necessary for academic success in the content area of **Language Arts**.

Standard 3—Language of Mathematics

ELLs communicate information, ideas, and concepts necessary for academic success in the content area of **Mathematics**.

Standard 4—Language of Science

ELLs communicate information, ideas, and concepts necessary for academic success in the content area of **Science**.

For practical purposes, the four Standards are abbreviated as follows in this report:

Social and Instructional language: SI

Language of English Language Arts: LA

Language of Mathematics: MA Language

of Science: SC

The selected response items and performance-based tasks on Alternate ACCESS target these four Standards.

1.2.1. Grade-level Clusters

The WIDA ELD Standards describe developing ELP for five grade-level clusters. These are PreK- K, 1-2, 3-5, 6-8, and 9-12. A Kindergarten version of Alternate ACCESS, however, is not currently available. Thus, Alternate ACCESS is organized into the following grade-level clusters: 1-2, 3-5, 6-8, and 9-12.²

1.2.2. Language Domains

The Alternate ACCESS test includes individual sections to assess each of four language domains: Listening, Reading, Speaking, and Writing.

² The organization of grade-level clusters is based on the 2007 WIDA ELP Standards (WIDA, 2007).

1.2.3. Language Proficiency Levels

Alternate ACCESS assesses growth in ELP over six levels. These six levels include three newly developed language proficiency levels and three levels derived from the WIDA ELD Standards for the general population. The most basic proficiency level is A1: ‘Initiating,’ and the most advanced stage of language proficiency described is P3: ‘Developing’. The first three levels of the Alternate ELD proficiency levels, A1 – A3, are language proficiency antecedents to the existing WIDA ELD P1 that applies to the general student population. An important aspect of the Alternate ELD levels (A1 – A3) is that they represent small chunks of language growth within P1. A highlight of this structure is that progress in language acquisition for students with significant cognitive disabilities can be identified in smaller and narrower gradations. Figure 1.2.4A below presents a conceptualization of the proficiency levels assessed in Alternate ACCESS. In this figure, P1 has been stretched for illustrative purposes to display levels A1 – A3.

ACCESS. In this figure, PL1 has been stretched for illustrative purposes to display levels A1 – A3.

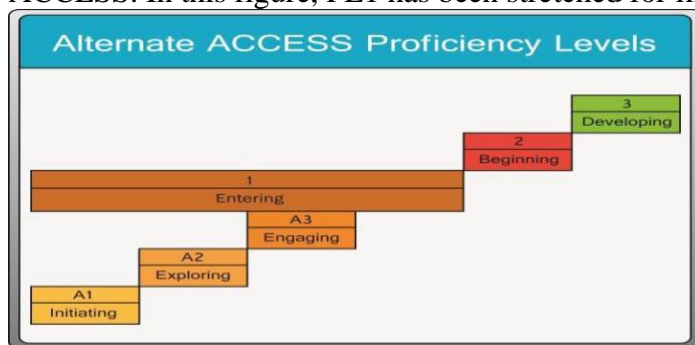


Figure 1.2.4A. Alternate ACCESS Proficiency Levels

These language proficiency levels are thoroughly embedded in the WIDA ELD Standards in a two-pronged fashion.

First, they appear in the **performance definitions**. According to the WIDA ELD Standards, the performance definitions provide a global overview of the stages of the language acquisition process. As such, they complement the **Alternate Model Performance Indicators (AMPIs)** for each language proficiency level (see the next paragraph for further description of the AMPIs). The performance definitions are based on three criteria. The first is students’ increasing comprehension and production of the technical language required for success in the academic content areas. The second criterion is students’ demonstration of oral interaction or writing of increasing linguistic complexity. The final criterion is the increasing development of phonological, syntactic, and semantic understanding in receptive skills or control in usage in productive language skills.

Second, the language proficiency levels of the WIDA ELD Standards are fully embedded in the accompanying AMPIs, which exemplify the Standards. The AMPIs describe the expectations for ELLs with significant cognitive disabilities for each of the four **Standards**, at the four different **grade-level clusters**, across four **language domains**, and at each of the **language proficiency levels**. The sequence of these five AMPIs together describes a logical progression and accumulation of skills on the path from the lowest level of ELP to full proficiency for academic success. This progression is called a “strand.”

Each selected-response item or performance-based task on Alternate ACCESS is carefully developed, reviewed, piloted, and field tested.

1.3. Test Development

1.3.1. Item Development

Items developed for Alternate ACCESS were field tested on Form 100 and included on Form 101. The initial item writing for Alternate ACCESS was done as part of a U.S. Department of Education, Enhanced Assessment Grant at the University of Wisconsin. The subsequent pool of items was then refined by the CAL test development team. An internal review of the items was conducted, and items were chosen for further development based on how well they fit the Standards and AMPs. The chosen items were refined by CAL staff before proceeding through further test development activities.

Upon internal revision and development of test forms, CAL conducted the following test development activities, each followed by further internal review and revisions: Bias and Content Reviews, Pilot Testing, and WIDA/SEA's Forms Review. Details regarding this portion of the test development cycle can be found in the *Alternate ACCESS for ELLs Technical Report for Form 100*.

1.3.2. Field Test

Field testing of Alternate ACCESS Form 100 was conducted from March 12 to June 1, 2012. The purpose of the field test was to collect data on items and tasks, to judge the strength of individual items and tasks, to develop the Alternate ACCESS reporting scale, and to conduct the Standard Setting Study.

In total, 1,912 students in Grades 1-12 in 15 WIDA states participated in the field test. Participating SEAs encouraged educators in their states to sign up for the field test through the regular ACCESS for ELLs test ordering site provided by MetriTech, Inc. The administrations were labeled as an operational field test, meaning states had the option of designating participation in the testing as a field test activity or as the first operational testing opportunity of the Alternate ACCESS program. For more details about the field test please refer to the *Alternate ACCESS for ELLs Technical Report for Form 100*.

1.3.3. Scaling

Scaling is the process of developing a standard scale that maintains a consistent meaning across test administrations. Reporting scores on such a scale allows users to interpret test scores.

For Alternate ACCESS, a three-digit scale score (910 to 960) was selected to aid in score interpretation. The scale needed an interpretive center point across domains and composites, so the centering value of 935 was chosen to represent the midpoint of the cut score between proficiency levels A3 and P1 for the 3-5 grade-level cluster (see "Creating the Composite Scores" on the next page for more information about the composites). This is analogous to the ACCESS for ELLs scale, where the score of 350 is set as the center value and represents the cut score between proficiency levels P3 and P4 for Grade 5 (for more information see Kenyon, 2006).

Because the test blueprints across grade-level clusters by domain are the same and the Alternate PLs and AMPIs for the test tasks across grade-level clusters pose nearly identical linguistic challenges and differ only in the topics presented, it is desirable to have common cut scores across grade-level clusters by domain. In order to derive these common cut scores, however, test scores from all grade-level clusters need to be placed on a common scale. A common Rasch logit scale was developed to put the task parameters across grade-level clusters on the same scale, allowing test scores from all grade-level clusters to be placed on a common scale. Because the same scoring rules are used to convert students' original responses to raw scores by domain, a single rating scale was modeled across all grade-level clusters by domain. This was achieved by imposing the same threshold parameters across the four grade-level clusters by domain. Through this scaling process, task parameters as well as test scores across grade-level clusters are put on the same scale. The procedure for developing the reporting scale for Alternate ACCESS was complex, but involved a number of basic steps. These were carried out separately for each domain until the last stage, when the separate domain scales were combined to form the composite scores. These steps, as conducted following the field test administration, are briefly summarized here. For more details about the field test please refer to the *Alternate ACCESS for ELLs Technical Report for Form 100*.

Scaling Design: The measurement model that formed the basis of the Alternate ACCESS scaling analyses was the Rasch Rating Scale Model (Andrich, 1978), as this model is appropriate for polytomously scored test tasks. For the initial Rasch calibration, the Rasch analyses were conducted separately by grade-level cluster and domain; therefore, the parameters for each grade-level cluster and domain were expressed on a unique logit scale. In the later stages of the psychometric analysis, the step or threshold parameters were constrained to be equal across grade-level clusters by domain through an anchoring process in order to put the task parameters across grade-level clusters by domain on the same logit scale. The Grade 3-5 step or threshold parameters were then used as the common step values, primarily because more Grade 3-5 students participated in the field test, therefore producing more stable parameters than other grade-level clusters. For each domain, the Grades 1-2, 6-8, and 9-12 rating scale threshold parameters were anchored to the Grade 3-5 domain values using Winsteps. The difficulty parameters for Grades 1-2, 6-8, and 9-12 were unanchored and thus were calibrated in the runs. All task parameters including the difficulty and threshold parameters were placed on the same logit scale across grade-level clusters by domain through this process. The logit scales were then transformed to the common reporting scale.

Developing the Logit Scale: A calibration of the ability of the students and items using Rasch procedures was applied to the scored student responses, putting the difficulty of the items or tasks and the ability of the students onto one common interval linear scale. The units of this scale are called logits, and by default the scale is usually centered at 0 (representing the average item difficulty for the ACCESS for ELLs items being calibrated). Theoretically, the logit scale runs from minus infinity to plus infinity, although in practice most tests run from about -4 logits to +4 logits.

Transforming the Logit Scale to the Reporting Scale: The logit scale has both negative numbers and decimals, which makes it confusing for many users. Therefore, scores on the logit scale were then transformed onto a reporting scale by means of a linear transformation of the Alternate ACCESS score scale. There is a separate scale for each of the four domains: Listening, Reading, Writing, and Speaking.

Creating the Composite Scores: The scores on the four reporting scales were then combined, in predetermined proportions, to create four composite scores: an Oral Language score (based on performances in Listening and Speaking), a Literacy score (based on performances in Reading and Writing), a Comprehension score (based on performances in Listening and Reading), and an Overall score (based on performances in all four domains).

1.3.4. Standard Setting

The goal of the Standard Setting Study was to interpret performances on the Alternate ACCESS operational field test form in terms of the WIDA ELD Standards, AMPIs, and the WIDA Alternate ELP levels. As discussed in 1.3.3., because the test blueprints across grade-level clusters by domain are the same, and the Alternate ELP levels and AMPIs for the test tasks across grade-level clusters pose nearly identical linguistic challenges and differ only in the topics presented, common cut scores were set across grade-level clusters by domain. The study was held in Arlington, VA, on October 9-10, 2012.

The *Angoff Yes/No* methodology was used for all four domains because this method is thought to simplify the cognitive tasks that panelists are asked to perform (Cizek & Bunch, 2007). Having a straightforward cognitive task was important in this study as panelists had to examine many tasks to set four cut scores (A1/A2, A2/A3, A3/P1, and P1/P2) across the four domains (Listening, Speaking, Reading, and Writing).

The *Angoff Yes/No* method was designed for multiple choice and dichotomously scored tasks. This method asks the panelists to consider a student currently functioning at the borderline between two adjacent levels and then to review each question on the test, judging each task as either: a) *Yes, the borderline student is more likely than not to meet expectations for this task*; or b) *No, the borderline student is not more likely than not to meet expectations for this task*. Under this method, the average of the panelists' *Yes* decisions represents an estimated proportion of the target borderline group who would correctly answer the task.

Some modifications were made to the typical *Angoff Yes/No* methodology. First, for the two tasks in Writing Part C, which are scored using a rubric, panelists were shown various writing samples from all score points and asked to make the decision whether *Yes, the borderline student is more likely than not to have produced this sample*, or *No, the borderline student is not more likely than not to have produced this sample*. This approach to addressing the two rubric-scored tasks meant that the same judging procedures that the panelists used on all other tasks could also be used for these two tasks. The second modification was that the *Yes/No* judgment data collected from the panelists was analyzed using a logistic regression procedure to determine cuts. Logistic regression is a statistical technique for relating a continuous variable (i.e., the difficulty of the assessment tasks) to a dichotomous outcome (i.e., the *Yes/No* decisions made by the panelists). This approach was used to avoid limitations in the traditional summation approach of calculating final cut scores with the *Angoff Yes/No* method, which systematically makes lower cuts easier and higher cuts more difficult as compared to the typical *Angoff* method.

Standards were set on Writing Parts A and B and Speaking using the following procedure. Starting with a student at the lowest borderline within the WIDA Alternate ELP levels (i.e., between A1 and A2), panelists independently indicated whether that borderline student would be more likely than not to meet the expectation for the task. If their decision was *No*, panelists then went on to consider a borderline student at the next higher borderline on that same task (i.e., between A2 and A3). This process was continued, considering students at progressively higher levels of proficiency until they reached the highest borderline OR until they indicated *Yes*, that the borderline student would be more likely than not able to meet expectations for that task. Once a decision of *Yes* was made, then all higher borderlines would also necessarily be *Yes* and did not need to be individually considered. This aspect of the procedure greatly simplified the panelists' task.

After panelists considered the borderlines for one task, they then examined the next task and began again by considering a student at the lowest borderline. This process continued until panelists had considered all the borderlines on all the tasks. The test tasks were considered in the same order as they are presented in the Alternate ACCESS test booklets. Each panelist completed these evaluations independently. After the first round of evaluations, results for each task were tallied, allowing the panelists to see the 'average' borderline student (e.g., A2/A3) at which the group had determined the task to be more likely than not be answered correctly.

Writing Part C consisted of two writing tasks that were scored using a five-point rubric ('No Response,' 'Approaches,' 'Meets 1,' 'Meets 2,' and 'Meets 3') and therefore required a slightly different approach. Sample student responses to the two writing tasks were presented to panelists. Panelists were asked to determine whether a student at each borderline would be more likely than not able to have produced each writing sample.

For Listening and Reading, the prompts for the assessment tasks are repeated to students with increasing levels of support, allowing students multiple opportunities to respond. The repeated prompts are labeled as: CUE A: Initial Prompt; CUE B: Simplified Prompt; CUE C: Simplified Prompt & Answer. A response meeting expectations at CUE A (i.e., with minimal support) is interpreted as demonstrating a higher level of proficiency than a response meeting expectations at CUE B, and a response meeting expectations at CUE B exhibits higher proficiency than one at CUE C. For Listening and Reading, the panelists' task was the same as for Writing Parts A and B and Speaking, except that before moving on to the next task they first considered all borderlines on the first task at CUE A, then all borderlines on that task at CUE B, and, finally, all borderlines on that task at CUE C.

For all tasks across all four domains, panelists provided *Yes/No* decisions in a two-round process. In Round 1, panelists independently made their decisions. Staff members then typed the decisions into a specially prepared Excel spreadsheet which tallied the results by the total number of *Yes* and *No* responses. The tallied *Yes/No* decisions across panelists in the group were then revealed to all panelists on a screen with an LCD projector, at which point the panelists had the opportunity to comment on the tallies. Following this discussion, empirical data on student performances on the tasks were presented to the panelists. Using the results from the first round and this new information, the panelists then made a second round of independent *Yes/No* decisions. The Round 2 decisions were again entered and shared with the entire group. A brief opportunity was given to anyone who wanted to comment on the group results before moving on to the next language domain. At the conclusion of the study, researchers used the percentage of *Yes* decisions across panelists from Round 2 to derive the cut scores.

To derive the final cut scores by domain, a series of logistic regression analyses were conducted. A logistic regression analysis was conducted for each cut for each domain (e.g., the A3/P1 cut for Listening) using the panelists' *Yes/No* decisions across test tasks and grade clusters in that domain. The logistic function was used to find the location along the underlying ability continuum at which 50% of the panelists thought that the borderline student is more likely than not to meet the task expectations. This point became the cut point between the two adjacent proficiency levels being analyzed.

For more details regarding the Standard Setting Study, please refer to the *Alternate ACCESS for ELLs Standard Setting Study: Technical Brief* (CAL, 2012a).

1.4. Reporting of Results

1.4.1. Scale Scores

Alternate ACCESS scores are reported as both scale scores and proficiency level scores. Scores are given for all four language domains. In addition, four composite scores are given: Oral Language (based on performances in Listening and Speaking), Literacy (based on performances in Reading and Writing), Comprehension (based on performances in Listening and Reading), and Overall (based on performances in all four domains).

Raw scores are converted to scale scores through processes called scaling (see section 1.3.3 for details). These processes allow scores to be reported on a standard scale that is familiar to test users and that remains constant across test forms and grade-level clusters. Scale scores range from 910 to 960.

In determining the Oral Language and Literacy composite scores, equal weight is given to each domain. However, in determining the Comprehension and Overall composite scores, more weight is given to literacy skills than to oral skills. The scores are weighted as follows:

Comprehension = 70% Reading + 30% Listening

Overall = 35% Reading + 35% Writing + 15% Listening + 15% Speaking

1.4.2. Language Proficiency Level Scores

In addition to the scale scores, users of Alternate ACCESS also receive proficiency level scores. These scores are *interpretive*; that is, they interpret a student's scale score in terms of the results of the Standard Setting Study. The cut scores between proficiency levels are presented in Table 1.4.2A.

Table 1.4.2A

Cut Scores by Domain and Composite

Domain	A1/A2	A2/A3	A3/P1	P1/P2
Listening	925	932	937	942
Reading	924	932	937	942
Speaking	925	930	939	945
Writing	923	931	938	947
Oral Composite	925	931	938	944
Literacy Composite	924	932	938	945
Comprehension Composite	924	932	937	942
Overall Composite	924	931	938	944

1.5. Test Administration

1.5.1. Test Administrator Training

Test administrators for Alternate ACCESS are required to take the appropriate steps to prepare themselves for test administration. The training steps included reading through the Alternate ACCESS Test Administration Manual (TAM) (WIDA, 2012a) and the Alternate ACCESS Test Administration Tutorial (available on the WIDA website). Test administrators are instructed to internalize the Writing and Speaking rubrics which are essential to consistent scoring across test administrations. For the Writing section, in addition to these materials, the Writing Scoring Guidance document provides sample student papers that help calibrate scoring for the Writing Section.

1.5.2. Test Security

Every effort is made to keep the test secure at all levels of development and administration. CAL and Data Recognition Corporation (DRC) follow policies and procedures regarding the security of the test, and every individual involved in the administration of the test from the district to the classroom level is trained in issues of test security.

1.5.3. Test Accommodations

Alternate ACCESS was designed for a population of students with a wide range of physical and cognitive disabilities. As such, the test design and layout reflect built-in features that aim to provide accessibility and are included as available accommodations on standardized tests for the general population. However, there are many situations where test administrators would need to modify the test administration in order to accommodate student-specific needs. In such cases, the criteria for implementation of any accommodation is determined primarily by the following: guidance in a student's Individual Education Plan (IEP), state accommodation policies, and the WIDA guidelines for appropriate test accommodations specified in the Alternate ACCESS TAM.

1.6. Scoring

All domains (Listening, Reading, Writing and Speaking) are scored locally by test administrators in individual Student Response Booklets. Test administrators must prepare for the scoring of each of the sections by following guidance provided in the TAM. Additional materials for ensuring that test administrators understand the correct scoring guidelines include the Alternate ACCESS Test Administration Video Tutorial and Writing Scoring Guidance document available through the WIDA website at <http://www.wida.wisc.edu>. Once a school has finished testing, all test booklets are returned to DRC, where they are electronically scanned and recorded in an electronic database in preparation for data analysis.

1.6.1. Listening and Reading

As with all sections of the Alternate ACCESS test, the Listening and Reading sections are scored by the test administrator. The Listening and Reading tests are identical in administration procedures and consist of selected-response items that provide students with multiple opportunities to demonstrate their knowledge. It is helpful to understand the administration guidelines for the Listening and Reading tasks in order to understand the scoring procedures. The following steps are used to administer each task in the Listening and the Reading sections:

1. Administer CUE A (initial prompt and question for the task).
2. If the student does not respond, the test administrator must repeat CUE A again, as indicated in the test administrator's script.
3. If the student answers incorrectly or does not respond to CUE A, the test administrator will read CUE B. CUE B simplifies the initial prompt and asks the question again.
4. If the student responds incorrectly, or does not respond at all after the test administrator reads CUE B, the test administrator will administer CUE C. This cue provides the answer to the question, restates the prompt, and asks the question again.

Based on these administration guidelines for Listening and Reading, a student has a maximum of four opportunities to respond to each task (CUE A – 2, CUE B – 1, CUE C – 1). If a student responds correctly to the task at CUE A (including if the teacher repeated CUE A) the test administrator will score the task as **Correct at CUE A**. If after the two possible attempts at CUE A the test administrator moves on to CUE B and the student answers correctly, they will be scored as **Correct at CUE B**. Likewise, if the student has reached CUE C and answers correctly, they will be scored as **Correct at CUE C**. Finally, if after the four possible chances to answer the task the student has not selected the correct answer, the teacher will mark the task as **Incorrect**. If the student did not respond to any of the four opportunities, the task will be marked as **'No Response.'** Test administrators record all student responses in a Student Response Booklet.

1.6.2. Writing

As mentioned earlier, the Writing section is also scored by locally by the test administrator. It is important to understand the design and administration procedures of the Writing test in order to understand the scoring procedures.

The Writing section has three thematic folders, Parts A, B, and C.

- Part A of the Writing section has tasks at levels A1-P1.
- Part B of the Writing section has tasks at levels A1 –P1.
- Part C provides the student with tasks at Levels P1 – P3; a student is only administered Part C if s/he scores 'Meets' on seven of the eight tasks in Parts A and B.

In Parts A and B of the Writing section, the script is designed for the test administrator to model each task for the student. This provides students the opportunity to observe the test administrator perform the task before trying it. For example, in the first task of the Writing section, the test administrator’s script will instruct the test administrator to draw a circle around an image before asking the student to do the same. Similar to the Speaking section, each task in the Writing section provides the student with multiple opportunities for the student to produce a response. If the student produces a response that is appropriate for the task, a score of ‘Meets’ is assigned, and if the student does not produce a response that meets task expectations, a score of ‘Approaches’ is assigned. If the student does not respond during the task administration, ‘No Response’ is assigned to the task. The TAM instructs teachers to score the Writing section using scoring guidance provided in a column of the Writing score sheet termed the ‘Expect’ box. For each task in Parts A and B, the ‘Expect’ box provides the test administrator with a description of a response that would meet the task expectations (e.g., copy or write a word related to the task). The scoring guidelines in the ‘Expect’ boxes parallel the Writing rubric available in the TAM and the Student Response Booklet. Part C is scored based on the Writing rubric. Student performances can receive a score of ‘Meets 1,’ ‘Meets 2,’ ‘Meets 3,’ ‘Approaches,’ or ‘No Response.’ A score of ‘Meets’ 1, 2 or 3 corresponds to performances described in the Writing rubric for PL 1, 2, or 3. Test administrators are trained to follow the WIDA Consortium’s Writing Rubric for Alternate ACCESS and have access to Writing training materials through the WIDA website (www.wida.wisc.edu). Table 1.6.2A presents the Writing Rubric.

Table 1.6.2A

Writing Rubric for Alternate ACCESS

Level	Text Features
3-Developing	One or more simple and expanded sentences. Words in the sentence(s) may be original or adapted from model or source text. Generally comprehensible. Comprehensibility may be impeded from time by errors when text becomes more complex. Text is related to the task.
2-Emerging	One or more simple phrases. Text is original or adapted from model or source text. Comprehensible when text is adapted from model or source text. Comprehensibility may be impeded by errors in original text. Text is related to the task.
1-Entering	One or more general content words. Text is original or adapted from the model or source text. Generally comprehensible when text is adapted from model or source text. Comprehensibility may be significantly impeded in original text. Text is related to the task.
A3-Engaging	Single words and numbers. All or part of text is copied. If original text is present, it is not related to the task. Comprehensibility of the text may be significantly impeded by imprecise letter, symbol, or number formation. Text may or may not be related to the task.
A2-Exploring	Common single-digit numbers, letters, symbols, or syllables. All or part of text is copied. Comprehensibility of the text may be significantly impeded by imprecise letter, symbol, or number formation. Text may or may not be related to the task.
A1-Initiating	Pictorial representations and imprecise, but intentional markings such as drawing and scribbles. Representations may or may not be related to the task.

1.6.3. Speaking

The Speaking section is also scored by the test administrator. As with other sections of the test, it is helpful to understand the design and administration guidelines for the Speaking section in order to understand the scoring criteria for the Speaking section.

The Speaking section has two thematic folders, Parts A and B. Thematic folders are a set of tasks based on a common setting or story (e.g., students in the library). The graphic(s) and character(s) often remain the same for all the tasks in a thematic folder.

- Part A of the Speaking section has tasks at levels A1 - A3.
- Part B of the Speaking section has tasks at levels A1 - P2.
- The script for all tasks includes three questions (Question 1, 2, and 3), which offers multiple opportunities for the student to provide a response at a given task level.

In the Speaking section, the student is given up to six opportunities to respond. This provides students with multiple opportunities to respond appropriately to the task in English. For each task, the test administrator reads Question 1 and prompts the student to respond. If the student does not score 'Meets,' the test administrator must repeat the task again. If the student still does not score 'Meets' after the repetition, the test administrator must ask Question 2, which simplifies the prompt and, in some tasks, models the expected response. If the student again does not score 'Meets,' Question 2 must be repeated. If the student does not score 'Meets' after that repetition, the test administrator must administer Question 3. Again, if the student does not score 'Meets,' this question is repeated once. The possibility of repetition for all three questions provides the student with six opportunities to produce a response in each Speaking task. If the student produces an appropriate response to the task at any point within the six provided opportunities, the task is scored as 'Meets.' If the student is not able at any point to produce a response that meets task expectations, a score of 'Approaches' is assigned. If the student does not make any attempt to respond to the task, a score of 'No Response' is assigned. The TAM instructs teachers to score the Speaking section using scoring guidance provided in a column of the Speaking score sheet termed the 'Expect' box. For each task, the 'Expect' box provides the test administrator with a description of a response that would meet the task expectations (e.g., repeat a word or produce a phrase related to the task). The scoring guidelines in the 'Expect' boxes parallel the Speaking rubric shown in Table 1.6.3A.

Table 1.6.3A

Alternate ACCESS Speaking Rubric

Level	Text Features
2-Emerging	<p>Phrases or short sentences.</p> <p>General language related to the task; groping for vocabulary when going beyond the highly familiar is evident.</p> <p>When using simple discourse, is generally comprehensible and fluent; communication may be impeded by groping for language structures or by phonological, syntactic, or semantic errors when going beyond phrases and short, simple sentences.</p>
1-Entering	<p>Single words or chunks of memorized oral language.</p> <p>General vocabulary from school setting and related to task.</p> <p>When using memorized language, is generally comprehensible; communication may be significantly impeded when going beyond the highly familiar.</p>
A3-Engaging	<p>Single words or chunks of mimicked oral language.</p> <p>Mimicked high frequency vocabulary words related to the task.</p> <p>When using mimicked language, is generally comprehensible; communication may be significantly impeded when going beyond mimicked language.</p>
A2-Exploring	<p>Single syllables or syllables of single words; speech is mimicked.</p> <p>Mimicked sounds and syllables of high frequency vocabulary words related to the task.</p> <p>Language is minimal.</p>
A1-Initating	<p>Communicative vocalizations, which may be imitated (e.g., grunts). Indiscriminant sounds and syllables.</p>

2 An Assessment Use Argument for Alternate ACCESS for ELLs: Focus on Assessment Records

Validity is “the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], 2014, p. 11). Evaluations of test validity assess whether there is evidence that supports the appropriateness and adequacy of the interpretations and decisions made about test takers on the basis of their performance on a test. This chapter contextualizes the information presented in this Annual Technical Report within an argument-based approach to addressing validity (Bachman & Palmer, 2010; Chapelle, Enright, & Jamieson, 2008; Kane, 2002, 2013; Mislevy, Almond, & Lukas, 2004) for Alternate ACCESS for ELLs.

A fully developed validation framework, including an Assessment Use Argument (AUA) (Bachman & Palmer, 2010), consists of several steps (described in Section 2.1 below) that connect test design and administration to intended and actual score interpretation and consequences. This chapter begins the process of developing a complete validation framework for Alternate ACCESS for ELLs. This argument-based structure organizes the information in this Annual Technical Report to support claims about Assessment Records (i.e., test scores and proficiency level descriptions collected via Alternate ACCESS for ELLs). Specifically, tables and figures from this report are explicitly linked to questions related assessment data. Chapelle, Enright, & Jamieson (2010) support using such a structure to present information to assessment users because “based on an analysis of four points of comparison—framing the intended score interpretation, outlining the essential research, structuring research results into a validity argument, and challenging the validity argument—we conclude that an argument-based approach to validity introduces some new and useful concepts and practices” (p.3). A larger, though yet undocumented (as of 2014), validity argument for the complete assessment from its inception to its consequences is currently under development by WIDA.

The complete validity argument that will be employed to support the use of Alternate ACCESS for ELLs will show the path from test design to test taker performance to the uses and interpretations of test scores and the subsequent consequences of test use. This framework is structured around assertions, or claims, about the assessment. The claims are presented as a series of statements that connect some aspect of the assessment process to the intended purposes of the assessment.

Evidence for each claim is then organized by the action that is used to ensure each claim, and it includes results from analyses of test data, outside documentation, and other resources. In the complete validation argument, this process of identifying evidence to support claims will encompass the entire testing process, from the commencement of the test design to the consequences of test use (Bachman & Palmer, 2010; Llosa, 2008); Figure 2A shows the process by which evidence supports validation actions, which are used to establish larger claims about Alternate ACCESS for ELLs.

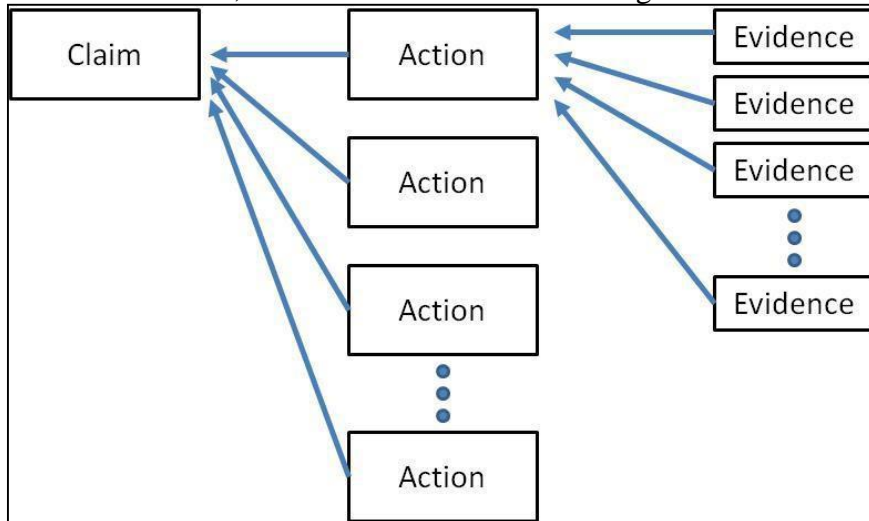


Figure 2A: General Argument Structure for Assessment Validation

2.1 The Generic Validation Framework for Alternate ACCESS

The generic validation framework that will be applied to the entire Alternate ACCESS for ELLs testing process was developed at the Center for Applied Linguistics (CAL) and is hereafter referred to as CAL’s Validation Framework. CAL’s Validation Framework, shown in Figure 2.1A, combines models for both test development (i.e., Evidence-Centered Design [Mislevy, Almond, & Lukas, 2004]) and assessment validation (i.e., Bachman and Palmer’s (2010) AUA) to cover the assessment development and implementation process from initial conceptualization to the score interpretations and consequences of using the assessment. This framework constantly looks both forward and backward; for example, during the initial *Plan* step (Step 7), test developers state the anticipated decisions and consequences of implementing the assessment program, which are investigated in the *Decisions* step (Step 2) and *Consequences* step (Step 1). Because each subsequent step depends upon the strength of the step below it, the steps are numbered from 7 to 1, with *Consequences* being the culmination of the previous steps. This structure highlights the fact that any weakness in a lower step affects the steps above it.

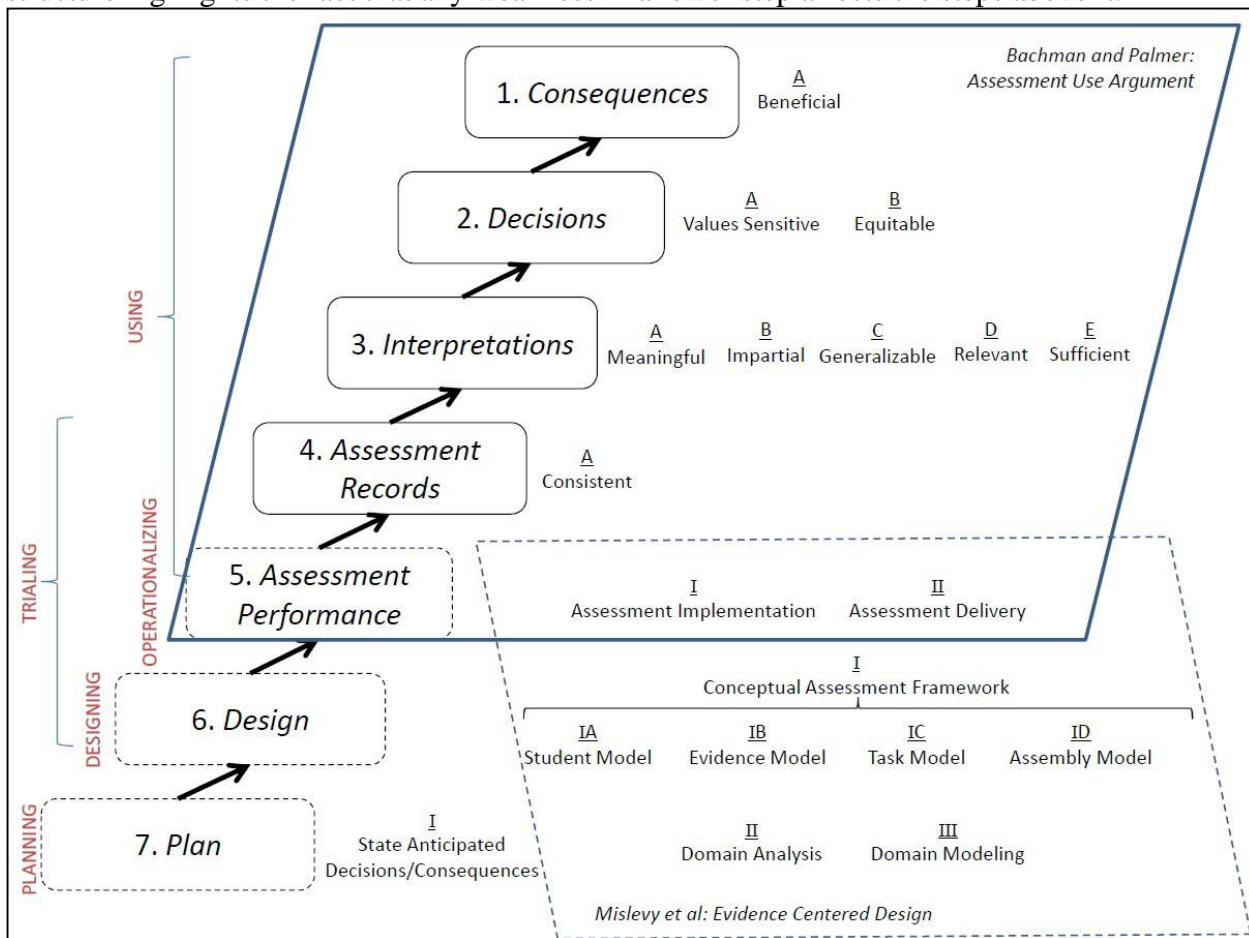


Figure 2.1A: CAL’s Validation Framework (based on Bachman & Palmer, 2010; Mislevy, Almond, & Lukas, 2004)

In CAL’s Validation Framework, the *Plan* step involves an examination of possible decisions states might make and consequences that might result from the assessment. This leads to the consideration of several models during the *Design* step, where specifications that answer such critical questions as “What are we measuring?” and “How do we measure it?” are developed (Mislevy, Almond, & Lukas, 2004). The subsequent steps of the validation framework highlight the trialing, implementation, and use of the assessment results, beginning with test takers’ performance on the assessment (*Assessment Performance*) and continuing through the collection of test scores (*Assessment Records*), interpretations of those test scores (*Interpretations*), decisions made based on the test scores (*Decisions*), and the consequences of test use (*Consequences*).

The WIDA Consortium is using CAL’s Validation Framework to present a complete validity argument, which will be updated as needed, for Alternate ACCESS for ELLs. To date, information related to Step 4, Assessment Records, has been explored and is found in this chapter.

2.2 Focus on Assessment Records

Although the complete validation framework for Alternate ACCESS for ELLs contains seven steps (see Figure 2.1A), the data presented in this document cover the Assessment Records step, which is part of Bachman and Palmer’s (2010) AUA. By focusing on Assessment Records (i.e., test scores and proficiency level descriptions), the information in the Annual Technical Report will be used to support claims related to the quality and consistency of the assessment data gathered and analyzed using Alternate ACCESS for ELLs. The claims in this step of the AUA all pertain to the general question “How do we know that the reported language domain scores and composite scores on Alternate ACCESS for ELLs are consistent and dependable?” Other questions about the development, administration, and outcomes of Alternate ACCESS for ELLs will be evaluated in a forthcoming document, currently in development by WIDA.

The diagram in Figure 2.2A shows a visual representation of an argument-based approach for supporting claims related to Assessment Records. The figure shows how the Assessment Records step, Step 4 of the complete validation framework, will fit in the generic validation framework and be expanded into a series of claims and corresponding actions in this chapter of the Annual Technical Report. Evidence in the form of data from this report or other sources will be presented to support these claims as they relate to ACCESS for ELLs.

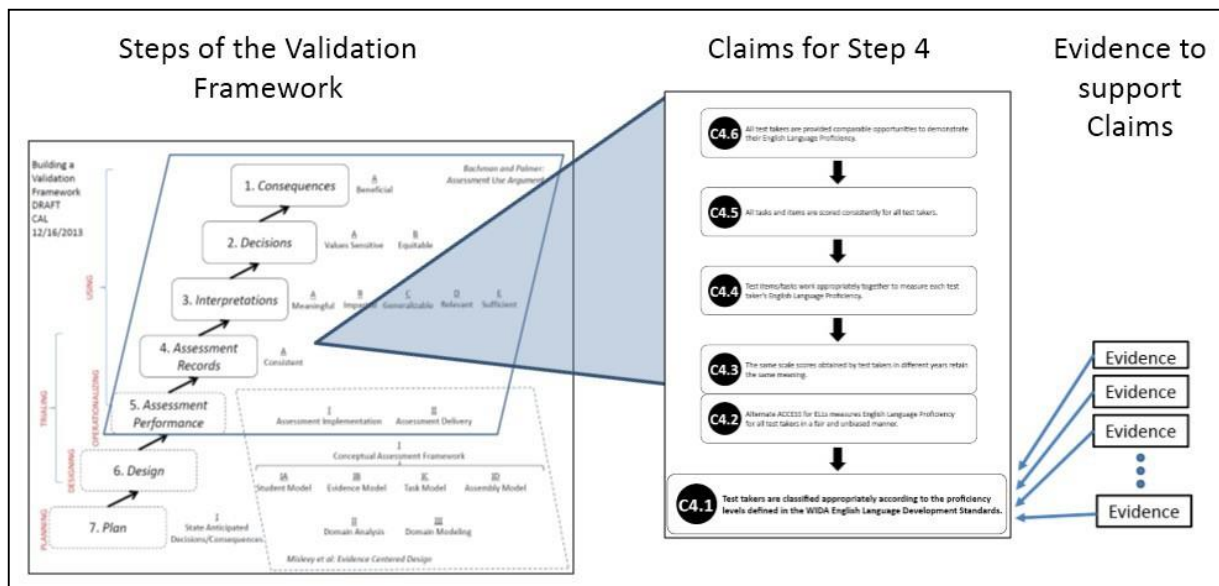


Figure 2.2A: Structure of the Argument-Based Approach Supporting Step 4 Contained in this Chapter

2.2.1 Breakdown of Claims for the Assessment Records Produced in the Alternate ACCESS for ELLs Assessment Program

The general *Assessment Records* step, Step 4 of the full Alternate ACCESS for ELLs validation framework, is broken down into the following six claims:

- C4.6. All test takers are provided comparable opportunities to demonstrate their English Language Proficiency.
- C4.5. All tasks and items are scored consistently for all test takers.
- C4.4. Test items/tasks work appropriately together to measure each test taker’s English Language Proficiency.
- C4.3. The same scale scores obtained by test takers in different years retain the same meaning.
- C4.2. Alternate ACCESS for ELLs measures English Language Proficiency for all test takers in a fair and unbiased manner.
- C4.1. Test takers are classified appropriately according to the Alternate English Proficiency Levels defined in the WIDA English Language Development Standards.

As shown in Figure 2.2.1A, these claims depend upon each other, again moving from (4.6) up to (4.1). Within this organizational structure, each successive claim builds upon the previous one(s) (e.g., ratings are only useful to test developers and stakeholders if all test takers are provided comparable opportunities to demonstrate their proficiency). In the next section, these claims are broken down even further into actions that are taken to ensure the consistency and reliability of the assessment records.

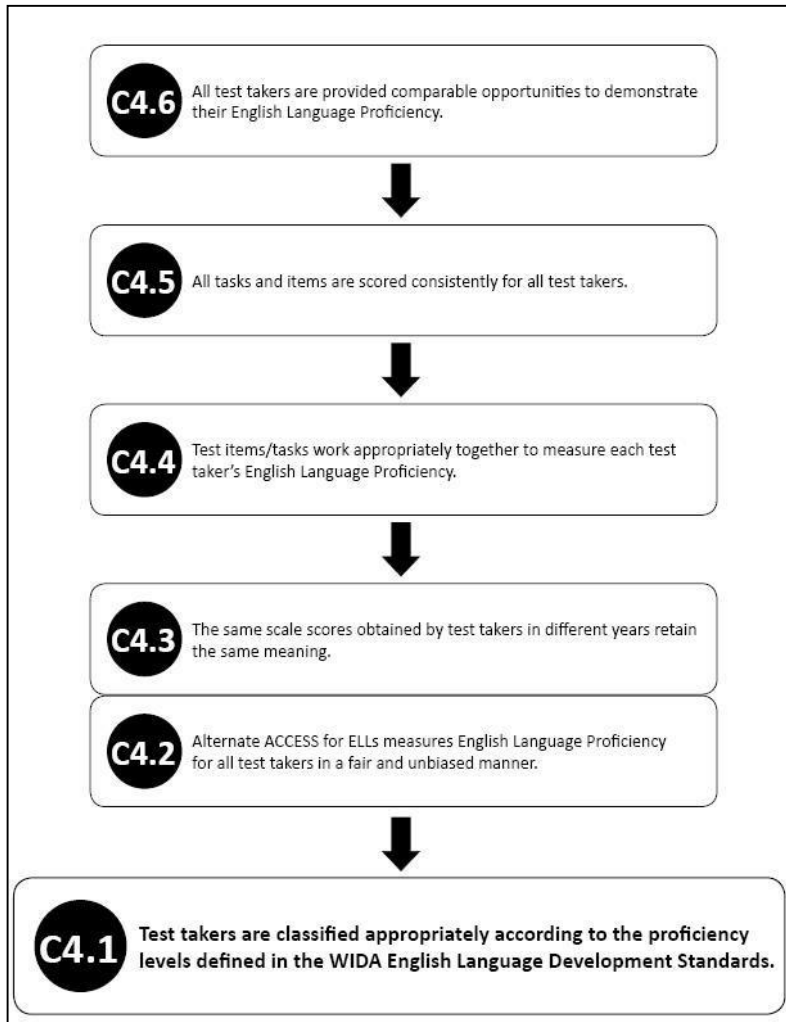


Figure 2.2.1A: Progression of Claims for Step 4: Assessment Records

2.3 Evidence for Assessment Records Claims of Alternate ACCESS for ELLs

In this section, evidence in the form of data or other sources (e.g., Test Administration Manuals, the technical brief of the Alternate ACCESS for ELLs standard setting study, the technical brief of the Alternate ACCESS for ELLs Series 100 development and operational field Test, and other information within this report, etc.) is connected to each of the *Assessment Records* claims via the actions taken to ensure those claims. This section denotes the tables, figures, and external sources that provide evidence related to each action. A summary table of the information presented in this section, including hyperlinks to the detailed description of each table or figure in Chapter 5 of this Annual Technical Report, is contained in Section 2.4. Information on how to navigate the tables and figures throughout this report is presented in Section 2.5.

Because these claims relate to Step 4 of the overall validation framework, their numbering begins with 4. The second number (after the decimal) denotes the level of the claim within Step 4. This numbering system is used in anticipation of the development of more complete documentation of a validity argument for Alternate ACCESS for ELLs, which will be completed by WIDA. Individual actions to ensure each claim are denoted by the final letter (a, b, c, and so on).

Claim 4.6 - All test takers are provided comparable opportunities to demonstrate their English Language Proficiency.

Action 4.6.a: The students that take Alternate ACCESS for ELLs have been identified as English language learners and participate in an alternate curriculum that aligns with the test.

Evidence: Exclusionary criteria and participation guidelines are closely followed by local test administrators (see Table 4.10.1 Participation by Disability, S502).

Action 4.6b: All test takers are given equal opportunities to demonstrate their English language proficiency.

Evidence: The Test Administration Manual provides clear guidance on the use of supporting features of Alternate ACCESS for ELLs, including repetition of questions, availability of cues, etc. (WIDA, 2013). If necessary, further accommodations for test takers are taken following the principles in the test administration manual.

Action 4.6c: Well-specified procedures were developed for test administrators so that they are able to administer the test consistently.

Evidence: Procedures for administering the test, stopping the test, and producing reported scores are documented in the Alternate ACCESS for ELLs Test Administration Manual (WIDA, 2013).

Action 4.6d: Test administrators document and report any irregularities that may occur so that appropriate action may be taken.

Evidence: Alternate ACCESS student response booklets contain a section for reporting irregular cases, such as invalid administration, absent student, or declined assessment. Test administration procedures are documented in the Alternate ACCESS for ELLs Test Administration Manual (WIDA, 2013).

Claim 4.5 – All items and tasks are scored consistently for all test takers.

Action 4.5a: A clear scoring design facilitates the task rating process for Test Administrators.

Evidence: The scoring procedures are clearly stated in the test administrator’s script and the Student Response Booklet is designed to match the scoring procedures and to avoid any scoring ambiguity.

Action 4.5b: Test Administrators undergo training so that they know how to score appropriately.

Evidence: Section 1.6 of this report specifies the scoring procedure for Alternate ACCESS for ELLs. Since all sections of Alternate ACCESS are scored locally, Test Administrators are provided with adequate training materials through an online program on the WIDA website to make sure they follow the test administration script and scoring rubrics for the Speaking and Writing sections. The scoring rubrics for Speaking and Writing are in the Test Administration Manual (WIDA, 2013).

Claim 4.4 - Test items/tasks work appropriately together to measure each test taker’s English Language Proficiency.

Action 4.4a: For each *test form* (e.g., Reading 6–8), item and task analyses are performed and psychometric properties of the items and tasks are evaluated to confirm that scores are internally consistent.

Evidence: Reliability information based on Classical Test Theory is calculated for each test form. This information includes Cronbach’s alpha, which is a measure of internal consistency. Cronbach’s coefficient alpha is widely used as an estimate of reliability and expresses how well the items on a test appear to work together to measure the same construct (see Table 6E).

Action 4.4b: For each *domain and composite score*, item and task analyses are performed and psychometric properties of the items and tasks are evaluated to confirm that scores are internally consistent.

Evidence: A single reliability estimate, a stratified Cronbach’s alpha (Cronbach, Schonemann, & McKie, 1965), is calculated by grade-level cluster for each domain and composite score. Cronbach’s alpha indicates the extent to which test items are consistent with each other. The stratified Cronbach’s alpha is an average reliability, and it is used when test takers are administered several related subtests but are then evaluated based on a composite of those subtest scores. Table 6E presents the data used to calculate an estimate of the reliability of the composite scores using a stratified Cronbach’s alpha.

Action 4.4c: Analyses of Rasch model fit statistics are conducted to show that individual tasks perform appropriately.

Evidence: The Complete Items Analysis table includes information on the Rasch fit statistics for each test item (see Table 6G). These statistics, called outfit mean square and infit mean square statistics, measure how well an item is measuring the same construct as other items on the test. Infit and outfit statistics indicate any consistently unusual performance in relation to the item’s difficulty measure by measuring the degree to which examinees’ responses to items deviate from expected responses. Both statistics have an expected value of 1.0. Items with infit and outfit mean square statistics between 0.5 and 1.5 are considered “productive for measurement” (Linacre, 2002). Values between 1.5 and 2.0 are “unproductive for construction of measurement, but not degrading.” Values greater than 2.0 might “distort or degrade the measurement system.” Values below 0.5 are “less productive for measurement, but not degrading.” Infit helps ensure that test takers within range of the targeted proficiency level perform as expected. It is not as sensitive to outliers as Outfit. Outfit can be skewed if test takers with extreme (i.e., high-level or low-level) proficiency do not perform as expected. High infit is a bigger threat to validity, but is more difficult to explain than high outfit (Linacre, 2002). The infit and outfit mean square statistics are part of the evaluation criteria used to select the items and tasks that appear on the final operational forms. Alternate ACCESS for ELLs test items with infit or outfit values between 1.2 and 1.3 are reviewed and items with values greater than 1.3 are not used on operational forms of the test.

Claim 4.3 - The same scale scores obtained by test takers in different years retain the same meaning.

Action 4.3a: All test items and tasks have been field tested and anchored using items from the operational field test (Series 100) to maintain a consistent scale from year to year.

Evidence: These retained “anchor items” ensure that performances on the newer form may be interpreted in the same frame of reference as the previous year. Table 6G displays information on the anchor items for each test form.

Action 4.3b: The same scaling equation is applied from year to year to ensure that scale scores are obtained consistently over time.

Evidence: The scaling equation table is used to convert a test taker’s ability measure, which is calculated based on test performance using Rasch modeling, into an Alternate ACCESS for ELLs scale score (see Table 6H). The same equation is used across grade-level clusters within each domain.

Claim 4.2 – Alternate ACCESS for ELLs measures English Language Proficiency for all test takers in a fair and unbiased manner.

Action 4.2a: Differential Item Functioning (DIF) analyses are conducted to determine whether any items or tasks may be biased against certain subgroups in terms of gender and ethnicity.

Evidence: The Item Analysis Summary provides a summary of the findings of the differential item functioning (DIF) analyses, which look for measurement bias in test items (see Table 6F). Analyses search for bias in contrasting groups based on gender (male versus female) and ethnicity (Hispanic versus non-Hispanic). This table shows the number of items that favored one group or the other at all levels of DIF.

The Complete Items Analysis table includes more detailed information on the DIF analyses, showing the degree of measurement bias for each item and which group is favored (Table 6G). Each item is categorized into three levels of DIF: A, B, or C (Zieky, 1993). An item exhibiting A level DIF shows little or no evidence of bias toward a particular group, an item exhibiting B level DIF displays a moderate amount of bias, and an item exhibiting C level DIF is considered to display considerable evidence for potential bias and should be closely examined by test developers to identify any construct irrelevant factors that may contribute to DIF.

Action 4.2b: Items that show evidence of DIF are carefully reviewed so that any that indicate bias are not used for scoring and are removed from future test forms.

Evidence: As described in Chapter 5.1.4 (DIF Items), ethnicity and gender DIF analyses are conducted using all test taker data.

Claim 4.1 - Test takers are classified appropriately according to the Alternate proficiency levels defined in the WIDA English Language Development Standards.

Action 4.1a: Distributions of scale scores and proficiency levels for each domain are analyzed to confirm that Alternate ACCESS for ELLs effectively measures the performance of test takers across the range of Alternate English Language Proficiency levels as defined by the WIDA English Language Development (ELD) Standards.

Evidence: The distribution of test takers' raw scores on Alternate ACCESS for ELLs, organized by individual test form (e.g., Reading 3–5), shows the extent to which Alternate ACCESS for ELLs effectively measures the performance of test takers across the range of ELD abilities that each form was designed to assess (see Table 6A; see Figure 6A).

The distribution of test takers' scale scores on Alternate ACCESS for ELLs, organized by test form (e.g., Reading 3–5), shows that Alternate ACCESS for ELLs effectively measures the performance of test takers across the range of ELD abilities that each form was designed to assess (see Table 6B; see Figure 6B).

The proficiency level distribution of test takers' scores on Alternate ACCESS for ELLs, organized by individual test form (e.g., Reading 3–5), shows that Alternate ACCESS for ELLs effectively measures the performance of test takers across the range of proficiency levels that each form was designed to assess (see Table 6C; see Figure 6C).

The Raw Score to Proficiency Level Score table shows the interpretive proficiency level score associated with each raw score (see Table 6I). This distribution of scores shows that Alternate ACCESS for ELLs effectively measures the performance of test takers across the range of proficiency levels that each form was designed to assess.

The Test Characteristic Curve for each test form graphically shows the relationship between test takers' ability measure (which is calculated based on test performance using Rasch modeling) on the horizontal axis and the expected raw scores on the vertical axis (see Figure 6D). Four vertical lines indicate the four cut scores for the highest grade in the cluster, dividing the figure into five sections for each of the five WIDA proficiency levels. The curve shows that higher expected raw scores are required to be placed into higher language proficiency levels.

Action 4.1b: Distributions of scale scores and proficiency levels, organized by grade-level cluster, are analyzed to confirm that Alternate ACCESS for ELLs effectively measures the performance of test takers across the range of Alternate English Language Proficiency levels as defined by the WIDA ELD Standards.

Evidence: The distribution of test takers' scale scores on Alternate ACCESS for ELLs, organized by grade-level cluster, shows that Alternate ACCESS for ELLs effectively measures the performance of test takers across the range of abilities as described by the WIDA ELD Standards (see Table 6B; see Figure 6B).

The proficiency level distribution of test takers' scores on Alternate ACCESS for ELLs, organized by grade-level cluster, shows that Alternate ACCESS for ELLs effectively measures the performance of test takers across the range of Alternate proficiency levels as defined by the WIDA ELD Standards (see Table 6C; see Figure 6C).

The Test Characteristic Curve reflects test takers' mean raw scores by domain on Alternate ACCESS for ELLs across the entire test for each grade-level cluster (except for the Kindergarten level) (see Figure 6D).

Action 4.1c: For each test form, analyses are run to confirm that English Language Proficiency is measured with high precision at the cut points.

Evidence: The Test Information Function graphically shows how well the test is measuring across the ability measure spectrum, which is calculated based on test performance using Rasch modeling (see Figure 6E). High values indicate more accuracy in measurement.

In the Raw Score to Proficiency Level Conversion Chart, the proficiency level associated with each raw score shows the distribution of proficiency level scores associated with each raw score for each grade in the cluster, along with the percentage of test takers in that grade who scored at that raw score/proficiency level score (see Table 6I). The Raw Score to Scale Score Conversion Chart (Table 6H) presents the conditional standard error for each scale score, along with the upper and lower bound of the scale scores within this standard error of measurement. This value indicates how accurately or precisely the test is measuring test takers at a particular ability level by estimating the error measurement at each score point. Because there is usually more information about test takers with scores in the middle of the score distribution on each form, the conditional standard error values are usually smallest and scores are more reliable in that region of the score distribution.

Action 4.1d: Classification and accuracy analyses are conducted by grade level to confirm that proficiency level classifications are reliable for all domain and composite scores.

Evidence: Information related to the accuracy of test takers' proficiency-level classifications is presented in multiple ways (see Table 6J). A separate table is provided for each grade level in a cluster. The table provides overall indices related to the accuracy and consistency of classification. These indices indicate the percent of all test takers who would be classified into the same language proficiency level by both the administered test and either the true score distribution (accuracy) or a parallel test (consistency). Cohen's kappa, which is a statistical measure of interrater agreement between two raters that takes chance agreement between raters into account, is also presented. A kappa value of 1 indicates complete agreement between the two raters, while a kappa value of 0 indicates no agreement other than what would be expected by chance. Table 6J also shows accuracy and consistency information conditional on level and provides indices of classification accuracy and consistency at the cut points.

2.4 Summary of Assessment Records Claims, Actions, and Evidence

Table 2.4A

Summary of Assessment Records Claims, Actions, and Evidence

Claim	Actions	Evidence
6. All test takers are provided comparable opportunities to demonstrate their English Language Proficiency	a. The students that take Alternate ACCESS have been identified as English language learners and participate in an alternate curriculum that aligns with the test.	a. Test Administration Manual Table 4.10.1 (<i>Participation by Disability</i>)
	b. All test takers are given supported opportunities to demonstrate their English language proficiency.	b. Test Administration Manual
	c. Well-specified procedures were developed for test administrators so that they are able to administer the test consistently.	c. Test Administration Manual
	d. Test administrators document and report any irregularities that may occur so that appropriate action may be taken	d. Test Administration Manual
5. All items and tasks are scored consistently for all test takers.	a. A clear scoring design facilitates the task rating process for Test Administrators.	a. Test Administration Manual; Student Response Booklets
	b. Raters of performance-based tasks undergo thorough training so that they know how to score appropriately.	b. Chapter 1.6
4. Test items/tasks work appropriately together to measure each test taker's English Language Proficiency.	a. For each test form (e.g., Reading 6-8), item and task analyses are performed and psychometric properties of the items and tasks are evaluated to confirm that scores are internally consistent.	a. Table 6E (<i>Reliability</i>)
	b. For each domain and composite score, item and task analyses are performed and psychometric properties of the items and tasks are evaluated to confirm that scores are internally consistent.	b. Table 6E (<i>Reliability</i>)
	c. Analyses of Rasch model fit statistics are conducted to show that individual tasks perform appropriately.	c. Table 6G (<i>Complete Item Analysis</i>)
3. The same scale scores obtained by test takers in	a. All the items and tasks have been field tested and are used as anchor items from the operational field test (Series 100) to maintain a consistent scale from year to year.	a. Table 6D (<i>Equating Summary</i>)

different years retain the same meaning.	b. The same scaling equation is applied from year to year to ensure that scale scores are obtained consistently overtime.	b. Table 6H (Raw Score to Scale Score Conversation Chart)
2. Alternate ACCESS for ELLs measures English Language Proficiency for all test takers in a fair and unbiased manner.	<p>a. Differential Item Functioning (DIF) analyses are conducted to determine whether any items or tasks are biased against certain subgroups in terms of gender and ethnicity.</p> <p>b. Items that show evidence of DIF are carefully reviewed so that any that indicate bias are not used for scoring and are removed from future test forms.</p>	<p>a. Table 6F (<i>Item Analysis Summary</i>); Table 6G (<i>Complete Item Analysis</i>)</p> <p>b. Chapter 5.1.4 (<i>DIF Items</i>)</p>
1. Test takers are classified appropriately according to the Alternate proficiency levels defined in the WIDA English Language Development (ELD) Standards.	<p>a. Distributions of scale scores and proficiency levels for each domain are analyzed to confirm that Alternate ACCESS for ELLs effectively measures the performance of test takers across the range of Alternate English Language Proficiency levels as defined by the WIDA ELD Standards.</p> <p>b. Distributions of scale scores and proficiency levels, organized by grade-level cluster, are analyzed to confirm that Alternate ACCESS for ELLs effectively measures the performance of test takers across the range of Alternate English Language Proficiency levels as defined by the WIDA ELD Standards</p> <p>c. For each test form, analyses are run to confirm that English Language Proficiency is measured with high precision at the pertinent cutpoints.</p> <p>d. Classification and accuracy analyses are conducted by grade-level to confirm that proficiency level classifications are reliable for all domain and composite scores.</p>	<p>a. Figure 6A (<i>Raw Scores</i>) & Table 6A (<i>Raw Score Descriptive Statistics</i>); Figure 6B (<i>Scale Scores</i>) & Table 6B (<i>Scale Score Descriptive Statistics</i>); Figure 6C (<i>Proficiency Level</i>) & Table 6C (<i>Proficiency Level Distribution</i>); Table 6I (<i>Raw Score to Proficiency Level Score Conversion Chart</i>); Figure 6D (<i>Test Characteristic Curve</i>)</p> <p>b. Figure 6B (<i>Scale Scores</i>) & Table 6B (<i>Scale Score Descriptive Statistics</i>); Figure 6C (<i>Proficiency Level</i>) & Table 6C (<i>Proficiency Level Distribution</i>); Figure 6D (<i>Test Characteristic Curve</i>)</p> <p>c. Figure 6E (<i>Test Information Function</i>); Table 6H (<i>Raw Score to Scale Score Conversion Chart</i>)</p> <p>d. Table 6J (<i>Accuracy and Consistency of Classification Indices</i>)</p>

2.5 Visual Guide to Tables and Figures

This section provides navigational support for the tables and figures contained in the Alternate ACCESS for ELLs Annual Technical Report. The Visual Guide to Tables and Figures, shown in

Figures 2.5.1 and 2.5.2, serves as a resource to quickly identify which table and/or figure to look for when seeking specific information based on grade, grade-level cluster, and demographic characteristics, such as state, gender, disability type, and ethnicity and race, as well as domains and domain composites.

To use the Visual Guide to Tables and Figures as a navigational tool, click on the links in Figures through 2.5.3 to navigate to the selected tables and figures in the Annual Technical Report. A link is provided at the end of each section in Chapters 4 and 6. Detailed descriptions of the information in each of the tables and figures is included in the preceding chapters (e.g., Chapter 5 contains information on tables and figures in Chapter 6). These descriptions may be accessed through links in Table 2.4A *Summary of Assessment Records Claims, Actions, and Evidence*.

Figure 2.5.1 displays the tables in Chapter 4 that provide information on participation, scale score, and proficiency level results, as well as results by standard. The key in the upper left corner of the figure describes the tables contained in each section of the chapter. For example, tables in Section 4.1 contain information about participation. To find specific information in Chapter 4, select the Grade or Grade Cluster tab, and then the Domain tab, and then choose from three categories: Demographic Characteristics, Domain Composites, or Domains. Within each of these categories, several additional options organize information so that individual tables can be accessed. For example, to find a table that displays information on the number of female Grade 2 students who completed the Speaking section, refer to Figure 2.5.1 and complete the following steps: one, select Grade; two, select Domains; three, select Demographic Characteristics; four, select Gender. The information is found in Table 4.2.2.2. Click on 4.2.2.2 to go to the appropriate table in Chapter 4.

Figure 2.5.2 displays the sections in Chapter 6 that contains analyses for each Alternate ACCESS for ELLs test form by grade-level cluster and domain. The key above the figure describes specific information in each table and figure. For example, to find the Reliability table for Grade-level Cluster 9–12 in the Reading domain, refer to Figure 2.5.2 and complete the following steps: one, select Grade Cluster 9–12; two, select; three, select Reading under Domains. Information for 9–12 Reading is shown in section 6.5.2.3. Finally, look at the key that explains that reliability information is located in table F. The result is Table 6.5.2.3F. Click on 6.5.2.3 to go to the appropriate section, and then locate Table F.

2.5.1 Chapter 4 Visual Guide to Tables and Figures

		Test Form Characteristics			
		Grade		Grade-Level Cluster	
		Domain		Domain	
Demographic Characteristics	State		4.3.1		4.2.1
	Gender	4.7.1	4.3.2	4.6.1	4.2.2
	Ethnicity and Race	4.7.2	4.3.3	4.6.2	4.2.3
Domain Composites	Overall		4.9.2 H		4.9.1 H
	Oral Language		4.9.2 E		4.9.1 E
	Literacy		4.9.2 F		4.9.1 F
	Comprehension		4.9.2 G		4.9.1 G
Domains	Across All Domains		4.4.2 4.5.1		4.4.1 4.8 4.5.1
	Listening		4.9.2 A		4.9.1 A
	Reading		4.9.2 B		4.9.1 B
	Writing		4.9.2 D		4.9.1 D
	Speaking		4.9.2 C		4.9.1 C

Figure 2.5.1 Chapter 4 Visual Guide to Tables and Figures

2.5.2 Chapter 6 Visual Guide to Tables and Figures

Table A and Figure A	Raw Score Descriptive Statistics
Table B and Figure B	Scale Score Descriptive Statistics
Table C and Figure C	Proficiency Level Distribution
Table D	Equating Summary
Figure D	Test Characteristic Curve
Table E	Reliability
Figure E	Test Information Function
Table F	Item Analysis Summary
Table G	Complete Item Analysis
Table H	Raw Score to Scale Score Conversion
Table I	Raw Score to Proficiency Level Conversion
Table J	Accuracy and Consistency of Classification Indices

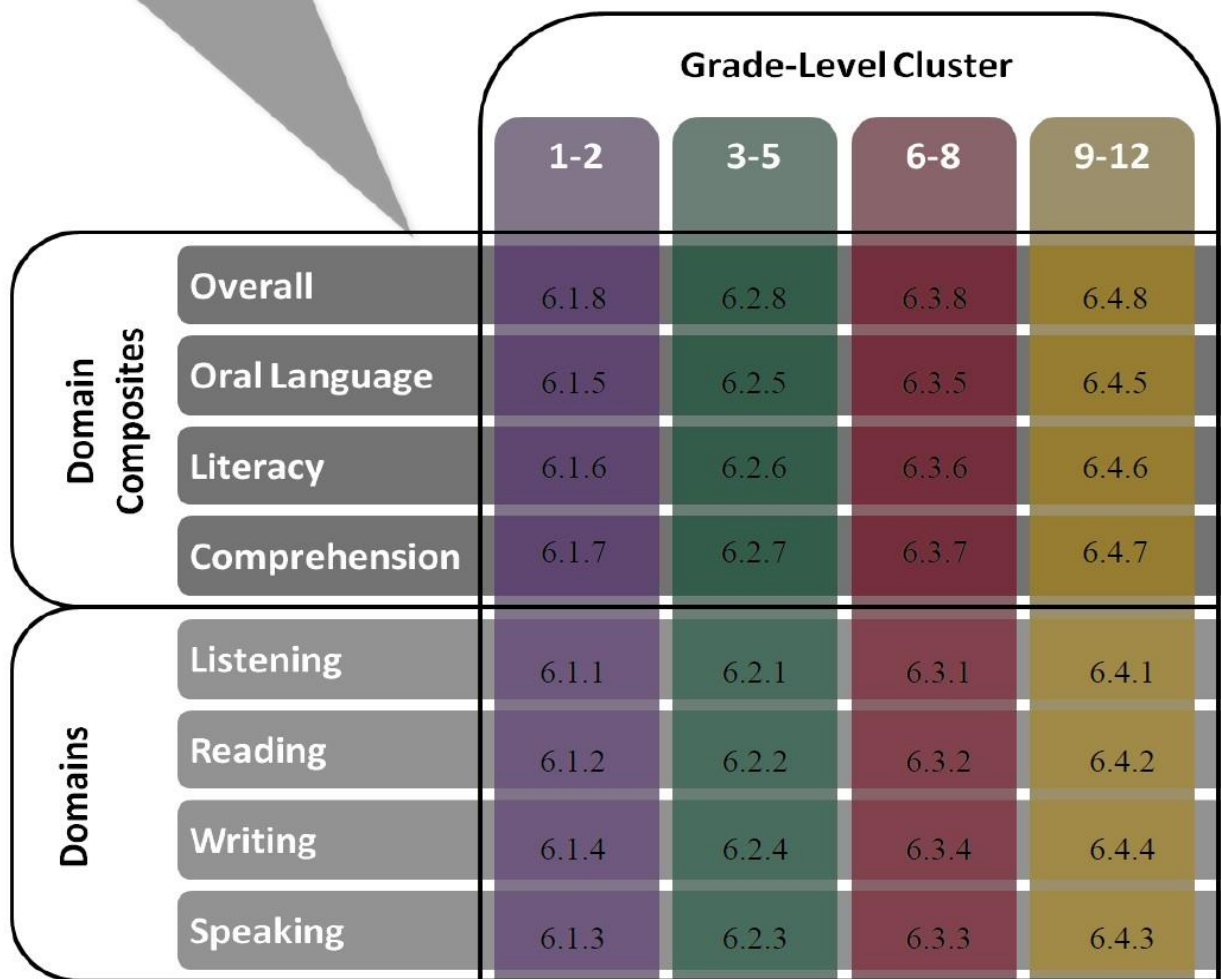


Figure 2.5.2 Chapter 6 Visual Guide to Tables and Figures

3. Descriptions of Student Results

Chapter 3 provides a description of the Chapter 4 tables summarizing students' participation, scale scores, and proficiency levels; results are further subdivided by grade, grade-level cluster, state, domain, domain and composite scores, gender, ethnicity/race, and disability. The 38 WIDA Consortium states/territories participated in the 2020-2021 Alternate ACCESS operational administration. The data used for analyses of the 502 Alternate ACCESS technical report is 14,688, which was drawn in September 2021. The final number of students that have taken the 502 Alternate ACCESS tests is 18,034 as of January 2022.

3.1 Participation

Table 4.1.1—Students Excluded from Analysis

In some circumstances there was a mismatch between a student's reported grade and the grade-level cluster (i.e., 1-2, 3-5, 6-8, or 9-12) actually administered (e.g., a student reported to be in Grade 1 who was administered a test intended for students in the 3-5 grade-level cluster). In all, 16 students were administered a test form not intended for their grade-level cluster. See Table 4.1.1 for a breakdown of the incorrect test forms assigned, by grade. The data from these 16 students were eliminated from all subsequent analyses in this report.

Section 4.2—Grade-Level Cluster, Gender, Ethnicity

Section 4.2 provides a breakdown of participation by *grade-level cluster* as a function of state (Table 4.2.1), gender (Table 4.2.2) and ethnicity (Table 4.2.3). For each of the 38 WIDA states who participated in the 2019-2020 operational testing program, Table 4.2.1 provides the number of test takers by grade-level cluster as well as total counts by state (final column) and grade-level cluster across all states (final row). For each grade-level cluster, Table 4.2.2 provides the distribution of test takers by gender (Female, Male, or Missing). Table 4.2.3 provides a similar breakdown of grade-level cluster by ethnicity (Hispanic or Non-Hispanic).

Section 4.3—Grade, Gender, Ethnicity

Section 4.3 duplicates the information provided by Section 4.2, but further breaks down the distribution of test takers by *grade* (Grades 1 to 12), instead of grade-level cluster. For each state, Table 4.3.1 provides the distribution of test takers by grade; for each grade, Table 4.3.2 provides the distribution of test takers by gender; for each grade, Table 4.3.3 provides the distribution of test takers by ethnicity.

Section 4.4—Domain, Grade-Level Cluster, Grade

Section 4.4 provides a breakdown of test taker counts by *domain* (Listening, Reading, Speaking, and Writing), with Table 4.4.1 summarizing the distribution by grade-level cluster and Table 4.4.2 summarizing the distribution by grade.

3.2 Scale Score Results

3.2.1 Mean Scale Scores Across Domain and Composite Scores

Overview of Sections 4.5 – 4.7

Sections 4.5 through 4.7 display the mean scale scores (Mean), standard deviation (Std. Dev.) and counts (N) by **grade and/or grade-level cluster** across the eight scores awarded on Alternate ACCESS for ELLs, first for each of the four domains (Listening, Reading, Speaking, and Writing) and then for each of the four composites (Oral Language, Literacy, Comprehension, and Overall). Sections 4.6 and 4.7 include gender and ethnicity information.

Section 4.5–Grade and Grade-Level Cluster

For each of the four grade-level clusters, Tables 4.5.1A through 4.5.1D display the mean scale scores for each domain and composite — first separately by grades within each cluster and then by the grade-level cluster overall (as the final column).

Section 4.6–Grade-Level Cluster, Gender, Ethnicity and Race

For each of the four grade-level clusters, Tables 4.6.1A through 4.6.1D display the mean scale scores for each domain and composite by gender. Correspondingly, Tables 4.6.2A through 4.6.2.D provide the mean scale score information by ethnicity and race. (Note that for the 4.6.1 Table series Domain is the row variable, and for the 4.6.2 table series Domain is the column variable.)

Section 4.7–Grade, Gender, Ethnicity and Race

For each of the 12 grades, Tables 4.7.1A through 4.7.1L display the mean scale scores for each domain and composite. Correspondingly, Tables 4.7.2.A through 4.7.2L display the mean scale scores by ethnicity and race.

3.2.2 Correlations

For each of the four grade-level clusters, Tables 4.8.1 through 4.8.4 display the Pearson correlations between scale scores on the four domains.

3.3 Proficiency Level Results

Section 3.3, Proficiency Level Results, displays the distribution of students' language proficiency level³ by grade-level cluster (Tables 4.9.1A-H) and grade (Tables 4.9.2A-H), with each sub-table presenting results by domain/composite:

- A. Listening
- B. Reading

³ The WIDA Alternate ELD Standards has six levels (A1-A3; P1; P2; P3). P3 was not part of the current analysis.

- C. Speaking
- D. Writing
- E. Oral Language Composite
- F. Literacy Composite
- G. Comprehension Composite
- H. Overall Composite

3.4 Participation by Disability

Table 4.10.1 displays the distribution of test takers as function of primary and secondary disability, each with 15 categories:

- No Primary Disability recorded (NPD)
- No Secondary Disability recorded (NSD)
- Autism Spectrum Disorder (AS)
- Deaf-blindness (DB)
- Developmental Delay (DD)
- Hearing Impairment, including Deafness (HI)
- Infant/Toddler with a Disability (ITD)
- Intellectual Disability (ID)
- Multiple Disability (MD)
- Orthopedic Impairment (OI)
- Other Health Impairment (OHI)
- Serious Emotional Disability (SED)
- Specific Learning Disability (SLD)
- Speech or Language Impairment (SLI)
- Traumatic Brain Injury (TBI)
- Visual Impairment, including Blindness (VI)

The accompanying *Acronyms for Table 4.10.1* table matches each disability category with its acronym to aid in interpretation.

4. Student Results

4.1 Students excluded from Analysis

4.1.1 Out-of-grade-level Test Administration

Table 4.1.1
Out-of-grade-level Test Administrations

Grade	Cluster				Total
	1-2	3-5	6-8	9-12	
1		0	0	0	0
2		2	0	0	2
3	3		0	0	3
4	0		1	0	1
5	0		4	0	4
6	0	2		0	2
7	0	0		0	0
8	0	0		1	1
9	0	0	3		3
10	0	0	0		0
11	0	0	0		0
12	0	0	0		0
Total	3	5	8	1	16

4.2 Participation by Grade-level Cluster

4.2.1 Participation by Grade-level Cluster by State

Table 4.2.1

Participation by Cluster by State

State	Cluster				Total
	1-2	3-5	6-8	9-12	
AK	3	6	13	13	35
AL	60	99	78	47	284
BI	2	1	1	.	4
CO	121	227	178	152	678
DD	.	12	1	8	21
DE	14	3	4	15	36
FL	335	382	174	116	1,007
GA	270	351	322	213	1,156
HI	34	67	53	50	204
ID	23	52	46	27	148
IL	404	517	384	506	1,811
IN	173	222	302	390	1,087
KY	63	69	63	91	286
MA	284	345	232	203	1,064
MD	3	9	16	11	39
ME	10	8	6	8	32
MI	133	167	131	169	600
MN	200	216	133	114	663
MO	42	51	46	40	179
MP	.	1	.	.	1
MT	4	5	5	4	18
NC	213	404	341	351	1,309
ND	4	5	6	8	23
NH	6	11	7	6	30
NJ	129	73	57	28	287
NM	5	5	4	3	17
NV	36	89	79	88	292
OK	150	222	173	145	690
PA	93	143	96	103	435
RI	19	41	38	40	138
SC	88	110	72	75	345
SD	8	14	12	17	51
TN	83	108	103	84	378
UT	87	135	113	138	473
VA	117	140	90	68	415
VT	9	10	3	5	27
WI	51	117	110	121	399
WY	5	10	3	8	26
Total	3,281	4,447	3,495	3,465	14,688

4.2.2 Participation by Grade-level Cluster by Gender

Table 4.2.2

Participation by Cluster by Gender

Cluster	Gender						Total
	Female		Male		Missing		
	Count	% within Cluster	Count	% within Cluster	Count	% withi	
1-2	920	28.04	2,310	70.41	51	1.55	3,281
3-5	1,444	32.47	2,947	66.27	56	1.26	4,447
6-8	1,236	35.36	2,192	62.72	67	1.92	3,495
9-12	1,236	35.67	2,164	62.45	65	1.88	3,465
Total	4,836	32.92	9,613	65.45	239	1.63	14,688

4.2.3 Participation by Grade-level Cluster by Ethnicity

Table 4.2.3

Participation by Cluster by Ethnicity

Cluster	Hispanic/Non-Hispanic						Total
	Hispanic		Non-Hispanic		Missing		
	Count	% within Cluster	Count	% within Cluster	Count	% within Cluster	
1-2	1,919	58.49	1,164	35.48	198	6.03	3,281
3-5	2,830	63.64	1,351	30.38	266	5.98	4,447
6-8	2,336	66.84	970	27.75	189	5.41	3,495
9-12	2,253	65.02	982	28.34	230	6.64	3,465
Total	9,338	63.58	4,467	30.41	883	6.01	14,688

4.3 Participation by Grade

4.3.1 Participation by Grade by State

Table 4.3.1
Participation by Grade by State

State	Grade												Total
	1	2	3	4	5	6	7	8	9	10	11	12	
AK	.	3	3	2	1	3	5	5	4	1	3	5	35
AL	34	26	39	31	29	24	28	26	14	11	10	12	284
BI	2	.	.	.	1	.	1	4
CO	47	74	75	76	76	68	54	56	45	38	28	41	678
DD	.	.	5	4	3	.	.	1	1	3	.	4	21
DE	10	4	.	2	1	3	1	.	8	2	1	4	36
FL	176	159	119	154	109	76	51	47	30	29	24	33	1,007
GA	122	148	121	119	111	115	98	109	59	60	43	51	1,156
HI	12	22	20	25	22	15	15	23	12	9	5	24	204
ID	9	14	12	21	19	15	14	17	4	8	13	2	148
IL	212	192	178	169	170	146	122	116	132	107	90	177	1,811
IN	91	82	77	79	66	114	92	96	102	86	79	123	1,087
KY	37	26	23	26	20	20	22	21	29	23	21	18	286
MA	124	160	126	112	107	77	84	71	40	63	40	60	1,064
MD	.	3	3	2	4	5	7	4	2	3	2	4	39
ME	5	5	4	2	2	1	4	1	1	2	3	2	32
MI	56	77	62	62	43	51	42	38	42	49	44	34	600
MN	100	100	82	65	69	49	35	49	27	36	19	32	663
MO	23	19	15	11	25	19	16	11	13	7	14	6	179
MP	.	.	1	1
MT	.	4	.	4	1	5	.	.	3	.	1	.	18
NC	89	124	119	141	144	114	117	110	77	76	80	118	1,309
ND	1	3	2	2	1	1	3	2	2	2	2	2	23
NH	3	3	3	4	4	2	4	1	4	.	.	2	30
NJ	75	54	27	27	19	23	17	17	8	12	5	3	287
NM	.	5	.	5	.	1	.	3	.	2	1	.	17
NV	19	17	26	27	36	25	29	25	20	24	25	19	292
OK	78	72	84	73	65	69	59	45	31	37	37	40	690
PA	49	44	54	50	39	46	31	19	21	20	20	42	435
RI	5	14	10	15	16	10	15	13	10	2	14	14	138
SC	38	50	47	34	29	24	18	30	21	18	18	18	345
SD	3	5	5	5	4	4	1	7	1	6	4	6	51
TN	46	37	34	40	34	34	40	29	28	21	18	17	378
UT	39	48	51	42	42	39	42	32	36	37	25	40	473
VA	63	54	62	41	37	37	27	26	19	17	10	22	415
VT	5	4	7	2	1	.	1	2	2	.	2	1	27
WI	22	29	43	41	33	40	38	32	30	14	20	57	399
WY	3	2	6	3	1	1	1	1	2	4	1	1	26
Total	1,598	1,683	1,545	1,518	1,384	1,276	1,134	1,085	880	829	722	1,034	14,688

4.3.2 Participation by Grade by Gender

Table 4.3.2

Participation by Grade by Gender

Grade	Gender						Total
	Female		Male		Missing		
	Count	% within Grade	Count	% within Grade	Count	% within Grade	
1	457	28.60	1,113	69.65	28	1.75	1,598
2	463	27.51	1,197	71.12	23	1.37	1,683
3	467	30.23	1,043	67.51	35	2.27	1,545
4	522	34.39	983	64.76	13	0.86	1,518
5	455	32.88	921	66.55	8	0.58	1,384
6	441	34.56	802	62.85	33	2.59	1,276
7	411	36.24	703	61.99	20	1.76	1,134
8	384	35.39	687	63.32	14	1.29	1,085
9	290	32.95	553	62.84	37	4.20	880
10	301	36.31	518	62.48	10	1.21	829
11	255	35.32	458	63.43	9	1.25	722
12	390	37.72	635	61.41	9	0.87	1,034
Total	4,836	32.92	9,613	65.45	239	1.63	14,688

4.3.3 Participation by Grade by Ethnicity

Table 4.3.3

Participation by Grade by Ethnicity

Grade	Hispanic/Non-Hispanic						Total
	Hispanic		Non-Hispanic		Missing		
	Count	% within Grade	Count	% within Grade	Count	% within Grade	
1	921	57.63	573	35.86	104	6.51	1,598
2	998	59.30	591	35.12	94	5.59	1,683
3	962	62.27	472	30.55	111	7.18	1,545
4	982	64.69	453	29.84	83	5.47	1,518
5	886	64.02	426	30.78	72	5.20	1,384
6	822	64.42	364	28.53	90	7.05	1,276
7	775	68.34	309	27.25	50	4.41	1,134
8	739	68.11	297	27.37	49	4.52	1,085
9	562	63.86	238	27.05	80	9.09	880
10	556	67.07	237	28.59	36	4.34	829
11	494	68.42	189	26.18	39	5.40	722
12	641	61.99	318	30.75	75	7.25	1,034
Total	9,338	63.58	4,467	30.41	883	6.01	14,688

4.4 Participation by Domain

4.4.1 Participation by Grade-level Cluster by Domain

Table 4.4.1
Participation by Cluster by Domain

Cluster	Domain			
	Listening	Reading	Speaking	Writing
1-2	3,232	3,207	3,193	3,194
3-5	4,425	4,413	4,377	4,354
6-8	3,473	3,468	3,429	3,411
9-12	3,439	3,430	3,394	3,366
Total	14,569	14,518	14,393	14,325

4.4.2 Participation by Grade by Domain

Table 4.4.2
Participation by Grade by Domain

Grade	Domain			
	Listening	Reading	Speaking	Writing
1	1,567	1,552	1,556	1,561
2	1,665	1,655	1,637	1,633
3	1,539	1,532	1,519	1,513
4	1,510	1,510	1,494	1,493
5	1,376	1,371	1,364	1,348
6	1,270	1,266	1,255	1,243
7	1,127	1,123	1,114	1,107
8	1,076	1,079	1,060	1,061
9	870	869	857	849
10	824	824	815	809
11	716	712	708	703
12	1,029	1,025	1,014	1,005
Total	14,569	14,518	14,393	14,325

4.5 Scale Scores by Domain and Composite

4.5.1 Mean Scale Scores by Domain and Composite

Table 4.5.1 A
Mean Scale Scores: 1-2

	Grade 1			Grade 2			Cluster 1-2		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	931.21	11.24	1,567	933.14	10.84	1,665	932.20	11.07	3,232
Reading	931.53	13.07	1,552	933.98	12.93	1,655	932.79	13.05	3,207
Speaking	932.28	14.56	1,556	934.69	14.13	1,637	933.51	14.39	3,193
Writing	927.05	11.04	1,561	929.54	11.35	1,633	928.32	11.27	3,194
Oral	932.08	12.02	1,546	934.31	11.68	1,626	933.22	11.90	3,172
Literacy	929.59	11.12	1,529	932.13	11.27	1,617	930.90	11.26	3,146
Comprehension	931.56	12.15	1,548	933.86	11.93	1,648	932.74	12.09	3,196
Overall	930.12	10.96	1,522	932.61	10.91	1,596	931.40	11.01	3,118

Table 4.5.1 B
Mean Scale Scores: 3-5

	Grade 3			Grade 4			Grade 5			Cluster 3-5		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	935.75	10.28	1,539	937.59	10.02	1,510	938.58	9.46	1,376	937.26	10.01	4,425
Reading	934.59	10.51	1,532	936.84	10.35	1,510	937.96	9.82	1,371	936.41	10.34	4,413
Speaking	935.45	13.25	1,519	937.61	12.60	1,494	938.12	12.36	1,364	937.02	12.81	4,377
Writing	930.84	10.76	1,513	933.14	11.29	1,493	933.75	11.28	1,348	932.53	11.17	4,354
Oral	935.81	10.82	1,516	937.77	10.48	1,487	938.56	10.10	1,360	937.33	10.54	4,363
Literacy	933.05	9.96	1,507	935.31	10.20	1,488	936.23	9.87	1,342	934.81	10.10	4,337
Comprehension	935.02	10.04	1,530	937.13	9.91	1,505	938.22	9.36	1,365	936.73	9.88	4,400
Overall	933.67	9.86	1,497	935.92	9.89	1,473	936.76	9.54	1,334	935.40	9.86	4,304

Table 4.5.1 C
Mean Scale Scores: 6-8

	Grade 6			Grade 7			Grade 8			Cluster 6-8		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	937.67	10.00	1,270	938.45	9.82	1,127	938.33	10.12	1,076	938.13	9.98	3,473
Reading	938.04	11.21	1,266	938.79	11.22	1,123	938.94	11.69	1,079	938.56	11.37	3,468
Speaking	937.16	12.55	1,255	937.49	12.47	1,114	937.66	12.82	1,060	937.42	12.61	3,429
Writing	932.98	10.30	1,243	933.38	10.36	1,107	934.13	10.40	1,061	933.46	10.36	3,411
Oral	937.88	10.68	1,254	938.46	10.60	1,109	938.46	10.87	1,059	938.25	10.71	3,422
Literacy	935.84	10.00	1,241	936.39	10.15	1,103	936.89	10.34	1,058	936.34	10.16	3,402
Comprehension	937.97	10.48	1,262	938.74	10.51	1,121	938.83	10.90	1,072	938.49	10.63	3,455
Overall	936.28	9.79	1,234	936.82	9.92	1,095	937.20	10.10	1,045	936.74	9.93	3,374

Table 4.5.1 D
Mean Scale Scores: 9-12

	Grade 9			Grade 10			Grade 11			Grade 12			Cluster 9-12		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	938.31	10.14	870	938.92	9.68	824	939.72	9.09	716	939.28	9.70	1,029	939.04	9.69	3,439
Reading	938.52	10.51	869	939.41	9.73	824	939.85	9.36	712	939.12	10.17	1,025	939.19	10.00	3,430
Speaking	937.19	11.58	857	937.45	11.20	815	937.76	11.24	708	937.28	11.57	1,014	937.40	11.41	3,394
Writing	934.67	10.71	849	935.48	10.62	809	934.99	10.24	703	934.86	10.79	1,005	934.99	10.62	3,366
Oral	938.11	9.89	853	938.43	9.65	815	939.00	9.35	704	938.54	9.86	1,012	938.50	9.71	3,384
Literacy	936.96	9.83	844	937.75	9.47	807	937.66	9.15	698	937.28	9.83	1,002	937.39	9.60	3,351
Comprehension	938.55	10.11	863	939.37	9.45	819	939.96	8.99	709	939.29	9.78	1,022	939.26	9.63	3,413
Overall	937.12	9.55	839	937.79	9.18	802	937.95	8.83	691	937.51	9.54	996	937.57	9.31	3,328

4.6 Scale Scores by Grade-level Cluster

4.6.1 Mean Scale Scores by Gender

Table 4.6.1 A
Mean Scale Scores by Gender: 1-2

	Female			Male			Missing		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	931.81	10.99	910	932.32	11.13	2,271	934.08	10.02	51
Reading	931.86	12.2	897	933.17	13.35	2,259	932.29	13.92	51
Speaking	932.41	14.62	893	933.95	14.26	2,249	933.82	15.26	51
Writing	926.7	10.72	887	928.96	11.39	2,257	928.52	12.28	50
Oral	932.49	11.95	889	933.49	11.86	2,232	934.27	12.23	51
Literacy	929.67	10.6	875	931.38	11.46	2,221	930.7	12.47	50
Comprehension	931.99	11.45	895	933.04	12.32	2,250	932.92	12.42	51
Overall	930.3	10.59	871	931.83	11.12	2,197	931.38	12.13	50

Table 4.6.1 B
Mean Scale Scores by Gender: 3-5

	Female			Male			Missing		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	937.55	9.55	1,434	937.21	10.15	2,935	932.66	12.67	56
Reading	936.44	9.80	1,433	936.47	10.53	2,925	932.55	12.7	55
Speaking	937.33	12.56	1,415	936.93	12.88	2,908	933.74	15.00	54
Writing	932.01	10.94	1,412	932.87	11.22	2,888	927.98	12.99	54
Oral	937.67	10.16	1,407	937.24	10.66	2,902	933.59	13.14	54
Literacy	934.6	9.71	1,405	934.99	10.22	2,878	930.57	12.45	54
Comprehension	936.86	9.35	1,427	936.75	10.06	2,918	932.75	12.32	55
Overall	935.36	9.46	1,389	935.49	9.98	2,861	931.3	12.44	54

Table 4.6.1 C
Mean Scale Scores by Gender: 6-8

	Female			Male			Missing		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	938.25	9.84	1,228	938.11	10.02	2,179	936.61	11.42	66
Reading	938.34	11.10	1,227	938.68	11.50	2,174	938.55	12.05	67
Speaking	937.28	12.72	1,217	937.52	12.55	2,146	936.89	12.49	66
Writing	933.12	10.40	1,212	933.65	10.34	2,134	933.63	10.20	65
Oral	938.21	10.71	1,214	938.3	10.69	2,142	937.08	11.70	66
Literacy	936.05	10.07	1,207	936.5	10.19	2,130	936.42	10.75	65
Comprehension	938.37	10.40	1,224	938.57	10.73	2,165	938.05	11.59	66
Overall	936.54	9.84	1,197	936.87	9.96	2,113	936.52	10.77	64

Table 4.6.1 D
Mean Scale Scores by Gender: 9-12

	Female			Male			Missing		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	939.04	9.66	1,226	939.09	9.68	2,148	937.37	10.60	65
Reading	939.37	9.52	1,223	939.15	10.18	2,143	936.92	12.36	64
Speaking	937.48	11.20	1,204	937.4	11.49	2,126	935.78	12.76	64
Writing	934.78	10.48	1,196	935.14	10.68	2,108	933.87	11.04	62
Oral	938.59	9.50	1,200	938.5	9.80	2,120	936.86	10.74	64
Literacy	937.38	9.32	1,192	937.45	9.72	2,097	935.48	11.00	62
Comprehension	939.39	9.30	1,215	939.25	9.76	2,134	937.19	11.40	64
Overall	937.58	9.07	1,183	937.62	9.41	2,083	935.77	10.63	62

4.6.2 Mean Scale Scores by Ethnicity

Table 4.6.2 A

Mean Scale Scores by Ethnicity: 1-2

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	931.35	932.78	932.95	929.13	932.40	931.21	932.45	931.38
	Std. Dev.	11.36	13.69	14.52	11.86	12.09	11.81	12.64	11.41
	N	493	490	492	492	490	485	489	482
Non-Hispanic Pacific Islander	Mean	930.24	928.07	930.68	924.52	930.75	926.82	928.86	927.68
	Std. Dev.	11.86	11.18	14.38	11.01	13.11	10.84	11.21	10.95
	N	29	29	28	29	28	28	29	28
Non-Hispanic Black	Mean	931.53	933.02	934.48	928.63	933.53	931.36	932.67	931.88
	Std. Dev.	10.39	13.26	13.50	11.16	10.89	11.11	12.00	10.49
	N	269	266	270	262	265	255	264	252
Hispanic (Of Any Race)	Mean	932.59	932.92	933.60	928.19	933.47	930.88	932.98	931.46
	Std. Dev.	11.07	12.81	14.46	10.93	11.93	10.99	11.91	10.83
	N	1,887	1,871	1,856	1,870	1,843	1,839	1,864	1,818
Non-Hispanic American Indian	Mean	931.18	932.00	934.00	930.60	932.82	932.00	931.73	932.10
	Std. Dev.	12.35	14.75	14.22	7.50	12.43	10.86	13.57	11.07
	N	11	11	11	10	11	10	11	10
Non-Hispanic Multi-racial	Mean	929.95	931.32	930.58	926.74	930.53	929.26	930.95	929.42
	Std. Dev.	13.69	16.44	16.37	13.67	14.59	14.33	15.27	14.05
	N	19	19	19	19	19	19	19	19
Non-Hispanic White	Mean	931.66	931.96	932.90	927.24	932.60	929.95	931.95	930.53
	Std. Dev.	10.97	12.81	14.56	12.06	11.98	11.58	11.91	11.25
	N	327	324	322	318	321	316	324	316
Missing	Mean	932.98	933.48	934.47	929.57	934.12	931.91	933.49	932.29
	Std. Dev.	10.93	13.61	14.12	11.45	11.82	11.86	12.43	11.50
	N	197	197	195	194	195	194	196	193

Table 4.6.2 B
Mean Scale Scores by Ethnicity: 3-5

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	935.85	935.84	935.91	933.72	936.08	935.10	935.93	935.22
	Std. Dev.	10.31	10.99	13.34	11.80	10.88	10.72	10.47	10.48
	N	491	493	486	491	485	490	490	484
Non-Hispanic Pacific Islander	Mean	936.47	935.12	937.91	930.73	937.48	933.45	935.56	934.45
	Std. Dev.	9.90	10.55	11.86	10.25	10.10	9.14	9.98	8.87
	N	34	34	33	33	33	33	34	33
Non-Hispanic Black	Mean	936.59	935.63	936.94	932.09	936.99	934.26	936.00	934.93
	Std. Dev.	10.87	10.57	13.09	11.25	11.20	10.20	10.20	10.03
	N	352	352	351	346	349	346	350	344
Hispanic (Of Any Race)	Mean	937.57	936.53	937.03	932.26	937.49	934.73	936.91	935.38
	Std. Dev.	9.78	10.20	12.72	10.87	10.38	9.87	9.72	9.65
	N	2,815	2,804	2,783	2,761	2,772	2,747	2,797	2,725
Non-Hispanic American Indian	Mean	936.65	936.00	936.20	933.10	936.50	934.95	936.20	935.25
	Std. Dev.	12.92	13.68	15.95	14.28	14.17	13.99	13.17	13.76
	N	20	20	20	21	20	20	20	20
Non-Hispanic Multi-racial	Mean	937.80	936.34	937.11	932.62	937.51	934.62	936.77	935.26
	Std. Dev.	7.97	9.24	11.96	11.33	9.11	9.83	8.31	9.11
	N	35	35	35	34	35	34	35	34
Non-Hispanic White	Mean	936.69	936.33	937.20	932.17	937.14	934.60	936.48	935.23
	Std. Dev.	9.97	9.88	12.55	11.75	10.44	10.21	9.58	9.85
	N	413	412	408	406	408	405	412	403
Missing	Mean	938.42	937.57	938.68	934.54	938.83	936.38	937.96	936.90
	Std. Dev.	10.51	10.65	12.65	11.69	10.74	10.73	10.25	10.4
	N	265	263	261	262	261	262	262	261

Table 4.6.2 C
Mean Scale Scores by Ethnicity: 6-8

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	936.93	937.94	936.31	933.77	937.06	936.25	937.69	936.41
	Std. Dev.	10.51	11.81	13.24	10.90	11.25	10.52	11.05	10.26
	N	382	384	377	377	377	377	381	372
Non-Hispanic Pacific Islander	Mean	938.79	940.62	939.86	933.41	940.18	937.24	940.03	938.54
	Std. Dev.	8.28	10.75	9.14	9.80	6.95	9.62	9.69	7.54
	N	29	29	28	29	28	29	29	28
Non-Hispanic Black	Mean	937.63	937.78	936.55	932.30	937.51	935.31	937.78	935.75
	Std. Dev.	10.30	12.24	13.57	10.61	11.38	10.90	11.37	10.71
	N	242	240	237	234	237	234	240	233
Hispanic (Of Any Race)	Mean	938.30	938.66	937.32	933.38	938.27	936.35	938.60	936.75
	Std. Dev.	9.83	11.22	12.61	10.21	10.66	9.99	10.49	9.82
	N	2,320	2,319	2,294	2,279	2,288	2,272	2,310	2,253
Non-Hispanic American Indian	Mean	943.70	943.30	941.30	934.30	943.00	939.00	943.40	940.00
	Std. Dev.	4.11	3.89	10.81	10.78	7.35	6.57	3.37	6.46
	N	10	10	10	10	10	10	10	10
Non-Hispanic Multi-racial	Mean	935.84	939.16	939.68	932.12	938.12	935.84	938.08	936.32
	Std. Dev.	11.40	12.03	9.83	8.92	9.85	9.89	11.56	9.68
	N	25	25	25	25	25	25	25	25
Non-Hispanic White	Mean	937.88	937.99	938.24	933.57	938.66	936.09	938.07	936.64
	Std. Dev.	10.64	11.87	12.39	11.03	10.89	10.83	11.12	10.48
	N	279	274	271	270	271	270	274	269
Missing	Mean	939.41	939.82	939.94	935.33	940.11	937.91	939.76	938.37
	Std. Dev.	9.30	10.50	10.64	9.82	9.51	9.70	9.85	9.26
	N	186	187	187	187	186	185	186	184

Table 4.6.2 D
Mean Scale Scores by Ethnicity: 9-12

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	938.62	938.80	936.49	935.27	937.80	937.26	938.84	937.24
	Std. Dev.	9.79	10.26	11.97	10.76	10.07	9.92	9.86	9.68
	N	369	369	362	364	362	364	367	359
Non-Hispanic Pacific Islander	Mean	940.55	941.86	938.77	937.36	939.73	939.77	941.59	939.59
	Std. Dev.	11.04	10.80	12.72	10.77	11.14	10.27	10.80	10.29
	N	22	22	22	22	22	22	22	22
Non-Hispanic Black	Mean	937.99	937.35	938.03	933.46	938.25	935.66	937.67	936.25
	Std. Dev.	9.43	9.93	9.96	10.07	9.03	9.46	9.47	8.97
	N	219	218	216	213	216	213	216	211
Hispanic (Of Any Race)	Mean	939.45	939.67	937.48	935.12	938.74	937.72	939.71	937.88
	Std. Dev.	9.38	9.65	11.43	10.45	9.54	9.31	9.31	9.04
	N	2,233	2,231	2,212	2,197	2,202	2,182	2,220	2,168
Non-Hispanic American Indian	Mean	938.47	939.74	935.84	933.68	937.32	936.89	939.42	936.89
	Std. Dev.	11.06	11.72	13.23	13.04	11.60	11.87	11.50	11.66
	N	19	19	19	19	19	19	19	19
Non-Hispanic Multi-racial	Mean	939.32	939.71	939.62	938.67	939.95	939.38	939.95	939.38
	Std. Dev.	8.21	10.47	7.13	10.25	6.92	9.34	9.23	8.41
	N	22	21	21	21	21	21	21	21
Non-Hispanic White	Mean	937.55	937.52	936.86	933.87	937.59	936.08	937.70	936.39
	Std. Dev.	11.46	11.25	11.75	11.65	10.76	10.72	10.94	10.42
	N	327	322	316	307	316	307	322	307
Missing	Mean	938.83	938.88	937.98	935.77	938.62	937.50	938.95	937.62
	Std. Dev.	9.73	10.44	11.24	10.67	9.84	9.92	9.97	9.68
	N	228	228	226	223	226	223	226	221

4.7 Scale Scores by Grade

4.7.1 Mean Scale Scores by Gender

Table 4.7.1 A

Mean Scale Scores by Gender: Grade 1

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	930.73	11.13	449	931.35	11.30	1,090	933.50	10.70	28	931.21	11.24	1,567
Reading	930.57	12.34	445	931.91	13.27	1,079	931.82	16.04	28	931.53	13.07	1,552
Speaking	931.36	14.61	445	932.69	14.48	1,083	930.93	16.67	28	932.28	14.56	1,556
Writing	925.75	10.35	443	927.60	11.24	1,090	926.32	12.57	28	927.05	11.04	1,561
Oral	931.37	11.96	442	932.36	12.01	1,076	932.50	13.22	28	932.08	12.02	1,546
Literacy	928.53	10.51	436	930.02	11.27	1,065	929.36	13.57	28	929.59	11.12	1,529
Comprehension	930.75	11.56	444	931.87	12.32	1,076	932.46	14.09	28	931.56	12.15	1,548
Overall	929.17	10.54	435	930.52	11.06	1,059	929.96	13.14	28	930.12	10.96	1,522

Table 4.7.1 B

Mean Scale Scores by Gender: Grade 2

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	932.87	10.75	461	933.21	10.90	1,181	934.78	9.31	23	933.14	10.84	1,665
Reading	933.13	11.93	452	934.32	13.32	1,180	932.87	11.13	23	933.98	12.93	1,655
Speaking	933.45	14.57	448	935.11	13.96	1,166	937.35	12.83	23	934.69	14.13	1,637
Writing	927.65	11.00	444	930.23	11.40	1,167	931.32	11.58	22	929.54	11.35	1,633
Oral	933.60	11.85	447	934.54	11.63	1,156	936.43	10.81	23	934.31	11.68	1,626
Literacy	930.79	10.59	439	932.64	11.49	1,156	932.41	10.99	22	932.13	11.27	1,617
Comprehension	933.20	11.22	451	934.12	12.22	1,174	933.48	10.31	23	933.86	11.93	1,648
Overall	931.43	10.53	436	933.05	11.03	1,138	933.18	10.74	22	932.61	10.91	1,596

Table 4.7.1 C
Mean Scale Scores by Gender: Grade 3

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	936.08	9.97	463	935.75	10.28	1,041	931.23	13.04	35	935.75	10.28	1,539
Reading	934.54	10.20	463	934.72	10.57	1,035	931.41	12.55	34	934.59	10.51	1,532
Speaking	935.31	13.27	461	935.61	13.16	1,025	932.27	15.45	33	935.45	13.25	1,519
Writing	930.49	10.85	459	931.13	10.62	1,021	926.73	13.00	33	930.84	10.76	1,513
Oral	935.90	10.70	458	935.88	10.77	1,025	932.24	13.43	33	935.81	10.82	1,516
Literacy	932.89	9.90	456	933.25	9.88	1,018	929.39	12.51	33	933.05	9.96	1,507
Comprehension	935.11	9.75	461	935.10	10.08	1,035	931.59	12.25	34	935.02	10.04	1,530
Overall	933.54	9.77	453	933.85	9.78	1,011	930.06	12.63	33	933.67	9.86	1,497

Table 4.7.1 D
Mean Scale Scores by Gender: Grade 4

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	938.12	9.18	518	937.38	10.38	979	932.54	13.00	13	937.59	10.02	1,510
Reading	937.05	9.51	520	936.78	10.72	977	932.38	13.48	13	936.84	10.35	1,510
Speaking	938.32	12.04	507	937.28	12.84	974	934.38	15.32	13	937.61	12.60	1,494
Writing	932.28	10.69	512	933.71	11.53	968	924.92	11.14	13	933.14	11.29	1,493
Oral	938.45	9.68	504	937.48	10.82	970	933.69	13.63	13	937.77	10.48	1,487
Literacy	935.02	9.41	510	935.55	10.55	965	928.92	11.77	13	935.31	10.20	1,488
Comprehension	937.48	9.02	517	937.01	10.30	975	932.46	13.03	13	937.13	9.91	1,505
Overall	935.98	9.10	500	935.97	10.24	960	930.08	11.93	13	935.92	9.89	1,473

Table 4.7.1 E
Mean Scale Scores by Gender: Grade 5

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	938.39	9.35	453	938.67	9.52	915	939.13	9.40	8	938.58	9.46	1,376
Reading	937.67	9.45	450	938.11	9.98	913	937.63	12.41	8	937.96	9.82	1,371
Speaking	938.27	12.14	447	938.04	12.48	909	938.75	13.10	8	938.12	12.36	1,364
Writing	933.28	11.17	441	933.95	11.33	899	938.13	12.21	8	933.75	11.28	1,348
Oral	938.62	9.90	445	938.52	10.20	907	939.00	11.06	8	938.56	10.10	1,360
Literacy	935.89	9.63	439	936.38	9.97	895	938.13	11.96	8	936.23	9.87	1,342
Comprehension	937.94	9.07	449	938.35	9.49	908	938.13	11.51	8	938.22	9.36	1,365
Overall	936.56	9.26	436	936.85	9.66	890	938.38	11.49	8	936.76	9.54	1,334

Table 4.7.1 F
Mean Scale Scores by Gender: Grade 6

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	937.86	9.86	439	937.68	9.92	798	935.06	13.21	33	937.67	10.00	1,270
Reading	937.59	10.63	439	938.34	11.41	794	936.70	13.51	33	938.04	11.21	1,266
Speaking	936.94	12.70	435	937.39	12.41	787	934.48	14.04	33	937.16	12.55	1,255
Writing	932.40	10.06	433	933.36	10.35	778	931.44	11.91	32	932.98	10.30	1,243
Oral	937.85	10.64	435	938.01	10.57	786	935.06	13.41	33	937.88	10.68	1,254
Literacy	935.27	9.60	432	936.22	10.09	777	934.16	12.49	32	935.84	10.00	1,241
Comprehension	937.70	10.07	438	938.19	10.59	791	936.27	13.22	33	937.97	10.48	1,262
Overall	935.87	9.51	429	936.60	9.82	773	934.28	12.45	32	936.28	9.79	1,234

Table 4.7.1 G
Mean Scale Scores by Gender: Grade 7

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	938.67	9.62	408	938.29	9.97	700	939.32	8.61	19	938.45	9.82	1,127
Reading	938.64	11.13	407	938.85	11.27	696	939.30	11.97	20	938.79	11.22	1,123
Speaking	937.41	12.72	406	937.48	12.37	689	939.74	10.89	19	937.49	12.47	1,114
Writing	933.04	10.72	404	933.52	10.20	684	935.32	8.08	19	933.38	10.36	1,107
Oral	938.48	10.70	404	938.40	10.58	686	939.95	9.61	19	938.46	10.60	1,109
Literacy	936.13	10.33	402	936.49	10.07	682	938.00	9.37	19	936.39	10.15	1,103
Comprehension	938.66	10.39	407	938.77	10.58	695	939.53	10.80	19	938.74	10.51	1,121
Overall	936.68	9.99	399	936.85	9.90	678	938.83	9.31	18	936.82	9.92	1,095

Table 4.7.1 H
Mean Scale Scores by Gender: Grade 8

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	938.24	10.06	381	938.42	10.17	681	936.57	10.14	14	938.33	10.12	1,076
Reading	938.89	11.57	381	938.90	11.83	684	941.86	7.53	14	938.94	11.69	1,079
Speaking	937.53	12.77	376	937.70	12.91	670	938.71	10.00	14	937.66	12.82	1,060
Writing	934.04	10.37	375	934.13	10.47	672	936.36	7.72	14	934.13	10.40	1,061
Oral	938.33	10.81	375	938.54	10.93	670	937.93	9.56	14	938.46	10.87	1,059
Literacy	936.88	10.29	373	936.84	10.43	671	939.43	7.01	14	936.89	10.34	1,058
Comprehension	938.82	10.75	379	938.80	11.04	679	940.21	7.99	14	938.83	10.90	1,072
Overall	937.16	10.04	369	937.20	10.19	662	938.64	7.42	14	937.20	10.10	1,045

Table 4.7.1 I
Mean Scale Scores by Gender: Grade 9

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	938.81	9.50	287	938.22	10.41	546	935.89	10.86	37	938.31	10.14	870
Reading	938.94	9.06	286	938.43	11.01	547	936.47	13.22	36	938.52	10.51	869
Speaking	936.50	11.78	282	937.59	11.43	539	936.69	12.37	36	937.19	11.58	857
Writing	934.46	10.38	278	934.94	10.86	536	932.20	11.07	35	934.67	10.71	849
Oral	938.09	9.43	280	938.22	10.06	537	936.61	10.90	36	938.11	9.89	853
Literacy	937.03	8.95	277	937.11	10.14	532	934.31	11.34	35	936.96	9.83	844
Comprehension	939.00	8.87	284	938.46	10.58	543	936.44	11.92	36	938.55	10.11	863
Overall	937.22	8.76	276	937.22	9.84	528	934.94	10.94	35	937.12	9.55	839

Table 4.7.1 J
Mean Scale Scores by Gender: Grade 10

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	937.98	10.43	298	939.43	9.23	516	941.00	6.38	10	938.92	9.68	824
Reading	938.75	10.26	299	939.79	9.45	515	939.80	6.58	10	939.41	9.73	824
Speaking	937.35	11.16	292	937.51	11.22	513	937.10	12.93	10	937.45	11.20	815
Writing	934.60	10.87	290	935.92	10.49	510	939.00	8.44	9	935.48	10.62	809
Oral	937.99	9.91	292	938.66	9.54	513	939.30	7.60	10	938.43	9.65	815
Literacy	937.07	9.85	289	938.09	9.27	509	939.89	7.13	9	937.75	9.47	807
Comprehension	938.61	10.05	296	939.79	9.12	513	940.30	6.31	10	939.37	9.45	819
Overall	937.16	9.55	286	938.11	8.98	507	939.56	7.26	9	937.79	9.18	802

Table 4.7.1 K
Mean Scale Scores by Gender: Grade 11

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	939.52	9.07	252	939.86	9.06	455	938.78	11.94	9	939.72	9.09	716
Reading	940.22	9.11	251	939.70	9.44	452	937.33	12.37	9	939.85	9.36	712
Speaking	938.16	10.95	251	937.64	11.35	448	932.56	13.31	9	937.76	11.24	708
Writing	935.09	9.95	250	934.89	10.47	444	936.78	6.38	9	934.99	10.24	703
Oral	939.06	9.29	249	939.03	9.35	446	935.78	11.94	9	939.00	9.35	704
Literacy	937.96	8.89	248	937.51	9.32	441	937.22	8.24	9	937.66	9.15	698
Comprehension	940.14	8.87	249	939.90	9.00	451	937.78	12.15	9	939.96	8.99	709
Overall	938.11	8.74	246	937.88	8.90	436	936.67	9.04	9	937.95	8.83	691

Table 4.7.1 L
Mean Scale Scores by Gender: Grade 12

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	939.73	9.50	389	939.03	9.78	631	938.00	12.25	9	939.28	9.70	1,029
Reading	939.61	9.52	387	938.87	10.48	629	935.11	14.91	9	939.12	10.17	1,025
Speaking	937.87	10.94	379	936.98	11.88	626	933.89	15.03	9	937.28	11.57	1,014
Writing	934.93	10.62	378	934.85	10.83	618	932.33	15.59	9	934.86	10.79	1,005
Oral	939.11	9.35	379	938.24	10.11	624	936.22	13.11	9	938.54	9.86	1,012
Literacy	937.50	9.46	378	937.19	9.98	615	933.89	14.94	9	937.28	9.83	1,002
Comprehension	939.79	9.26	386	939.02	10.01	627	936.11	13.88	9	939.29	9.78	1,022
Overall	937.84	9.13	375	937.36	9.72	612	934.33	13.98	9	937.51	9.54	996

4.7.2 Mean Scale Scores by Ethnicity

Table 4.7.2 A

Mean Scale Scores by Ethnicity: Grade 1

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	930.33	932.01	931.71	928.58	931.22	930.51	931.58	930.48
	Std. Dev.	11.23	13.66	14.53	11.89	11.93	11.57	12.60	11.17
	N	227	224	228	226	227	222	224	222
Non-Hispanic Pacific Islander	Mean	929.46	929.15	927.92	926.08	928.92	928.50	929.31	928.33
	Std. Dev.	12.51	12.52	15.72	11.42	14.20	11.55	12.31	11.80
	N	13	13	12	13	12	12	13	12
Non-Hispanic Black	Mean	930.55	931.26	932.45	926.94	931.91	929.44	931.11	929.97
	Std. Dev.	10.02	12.56	13.72	11.07	10.95	10.89	11.45	10.46
	N	137	134	139	136	137	131	134	131
Hispanic (Of Any Race)	Mean	931.61	931.59	932.62	926.82	932.48	929.48	931.76	930.19
	Std. Dev.	11.35	12.79	14.62	10.45	12.09	10.72	11.97	10.74
	N	899	889	889	899	882	877	886	871
Non-Hispanic American Indian	Mean	926.25	922.50	934.50	927.50	930.75	925.50	923.75	926.25
	Std. Dev.	13.70	10.15	11.56	7.14	12.55	8.10	10.87	9.18
	N	4	4	4	4	4	4	4	4
Non-Hispanic Multi-racial	Mean	930.38	932.88	930.25	923.75	930.63	928.50	932.13	929.00
	Std. Dev.	14.12	16.78	15.06	14.93	14.08	14.87	15.87	14.38
	N	8	8	8	8	8	8	8	8
Non-Hispanic White	Mean	931.18	931.02	931.71	926.15	931.75	928.96	931.11	929.60
	Std. Dev.	11.03	13.07	14.61	12.03	11.99	11.64	12.07	11.29
	N	176	176	174	173	174	173	176	173
Missing	Mean	931.05	931.67	931.98	927.67	931.88	930.10	931.63	930.32
	Std. Dev.	11.80	14.75	15.40	12.10	12.80	12.66	13.54	12.38
	N	103	104	102	102	102	102	103	101

Table 4.7.2 B
Mean Scale Scores by Ethnicity: Grade 2

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	932.23	933.43	934.01	929.61	933.42	931.80	933.18	932.15
	Std. Dev.	11.42	13.71	14.45	11.85	12.16	12.00	12.66	11.58
	N	266	266	264	266	263	263	265	260
Non-Hispanic Pacific Islander	Mean	930.88	927.19	932.75	923.25	932.13	925.56	928.50	927.19
	Std. Dev.	11.67	10.31	13.44	10.87	12.53	10.47	10.64	10.63
	N	16	16	16	16	16	16	16	16
Non-Hispanic Black	Mean	932.55	934.80	936.64	930.46	935.27	933.39	934.27	933.95
	Std. Dev.	10.72	13.75	12.97	11.01	10.60	11.02	12.39	10.17
	N	132	132	131	126	128	124	130	121
Hispanic (Of Any Race)	Mean	933.48	934.13	934.50	929.45	934.39	932.15	934.08	932.63
	Std. Dev.	10.73	12.72	14.27	11.23	11.71	11.07	11.76	10.79
	N	988	982	967	971	961	962	978	947
Non-Hispanic American Indian	Mean	934.00	937.43	933.71	932.67	934.00	936.33	936.29	936.00
	Std. Dev.	11.62	14.72	16.43	7.61	13.20	10.80	13.46	11.15
	N	7	7	7	6	7	6	7	6
Non-Hispanic Multi-racial	Mean	929.64	930.18	930.82	928.91	930.45	929.82	930.09	929.73
	Std. Dev.	14.06	16.91	17.98	12.96	15.64	14.63	15.54	14.51
	N	11	11	11	11	11	11	11	11
Non-Hispanic White	Mean	932.23	933.08	934.30	928.53	933.62	931.15	932.95	931.66
	Std. Dev.	10.91	12.44	14.44	11.99	11.92	11.43	11.68	11.13
	N	151	148	148	145	147	143	148	143
Missing	Mean	935.11	935.49	937.20	931.67	936.57	933.92	935.55	934.46
	Std. Dev.	9.51	11.98	12.07	10.36	10.16	10.62	10.78	10.09
	N	94	93	93	92	93	92	93	92

Table 4.7.2 C
Mean Scale Scores by Ethnicity: Grade 3

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	933.75	933.74	934.08	932.69	934.10	933.53	933.78	933.48
	Std. Dev.	10.44	11.28	13.56	11.92	11.00	10.78	10.60	10.54
	N	175	175	174	175	174	175	175	174
Non-Hispanic Pacific Islander	Mean	933.00	930.56	934.25	925.00	934.00	928.63	931.33	930.00
	Std. Dev.	10.48	11.80	13.40	8.47	11.95	9.61	10.87	9.53
	N	9	9	8	8	8	8	9	8
Non-Hispanic Black	Mean	935.64	933.75	935.03	930.57	935.55	932.55	934.47	933.22
	Std. Dev.	10.69	10.29	13.47	9.79	11.12	9.31	9.90	9.48
	N	118	119	119	116	118	116	118	116
Hispanic (Of Any Race)	Mean	936.00	934.56	935.34	930.20	935.88	932.71	935.06	933.44
	Std. Dev.	10.05	10.35	13.24	10.23	10.68	9.62	9.90	9.59
	N	958	952	943	938	941	933	951	925
Non-Hispanic American Indian	Mean	939.78	938.89	940.78	934.90	940.33	937.56	939.22	938.22
	Std. Dev.	11.51	10.96	12.56	11.93	11.78	11.62	10.79	11.20
	N	9	9	9	10	9	9	9	9
Non-Hispanic Multi-racial	Mean	935.11	934.44	936.67	931.89	935.89	933.33	934.67	933.89
	Std. Dev.	10.37	11.29	12.08	12.11	10.74	11.39	10.82	10.79
	N	9	9	9	9	9	9	9	9
Non-Hispanic White	Mean	935.87	935.93	936.24	931.29	936.19	933.91	935.97	934.50
	Std. Dev.	10.22	10.07	12.95	12.03	10.73	10.39	9.67	10.02
	N	150	150	149	149	149	149	150	148
Missing	Mean	936.58	935.34	937.43	933.06	937.41	934.54	935.93	935.17
	Std. Dev.	11.33	11.21	13.12	11.80	11.24	11.13	10.76	10.91
	N	111	109	108	108	108	108	109	108

Table 4.7.2 D
Mean Scale Scores by Ethnicity: Grade 4

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	936.81	936.50	936.44	933.45	936.76	935.28	936.72	935.60
	Std. Dev.	10.38	10.80	13.55	11.92	11.01	10.86	10.40	10.61
	N	162	164	161	163	160	163	162	159
Non-Hispanic Pacific Islander	Mean	936.30	937.50	937.60	933.30	937.20	935.80	937.10	935.90
	Std. Dev.	9.97	10.70	11.67	10.18	9.83	9.04	10.24	8.82
	N	10	10	10	10	10	10	10	10
Non-Hispanic Black	Mean	936.21	935.69	937.24	932.56	936.79	934.46	935.81	934.91
	Std. Dev.	11.22	10.61	13.21	12.10	11.65	10.60	10.40	10.43
	N	118	119	119	118	118	118	118	117
Hispanic (Of Any Race)	Mean	937.80	936.89	937.68	932.98	937.94	935.27	937.24	935.95
	Std. Dev.	9.89	10.30	12.42	10.98	10.28	9.98	9.81	9.66
	N	977	975	963	964	958	960	973	951
Non-Hispanic American Indian	Mean	933.75	933.13	931.63	930.75	932.75	932.25	933.25	932.13
	Std. Dev.	13.21	15.68	18.11	15.92	15.39	15.73	14.55	15.34
	N	8	8	8	8	8	8	8	8
Non-Hispanic Multi-racial	Mean	940.17	936.42	939.00	933.73	939.67	934.91	937.50	936.18
	Std. Dev.	6.41	10.50	9.06	9.80	6.53	9.94	8.24	8.45
	N	12	12	12	11	12	11	12	11
Non-Hispanic White	Mean	936.70	936.40	937.10	932.88	937.10	934.96	936.55	935.52
	Std. Dev.	9.72	9.93	12.76	11.90	10.51	10.56	9.61	10.13
	N	140	139	138	136	138	135	139	134
Missing	Mean	940.27	939.58	940.75	935.83	940.59	938.00	939.81	938.59
	Std. Dev.	9.21	9.47	11.29	11.17	9.64	9.79	9.23	9.45
	N	83	83	83	83	83	83	83	83

Table 4.7.2 E
Mean Scale Scores by Ethnicity: Grade 5

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	937.22	937.52	937.44	935.17	937.63	936.74	937.56	936.81
	Std. Dev.	9.74	10.55	12.70	11.46	10.31	10.29	10.02	10.02
	N	154	154	151	153	151	152	153	151
Non-Hispanic Pacific Islander	Mean	938.67	936.27	940.07	932.07	939.53	934.47	937.07	935.87
	Std. Dev.	9.57	9.51	11.45	10.59	9.38	8.60	9.17	8.32
	N	15	15	15	15	15	15	15	15
Non-Hispanic Black	Mean	937.94	937.53	938.64	933.17	938.70	935.82	937.79	936.75
	Std. Dev.	10.64	10.58	12.39	11.65	10.65	10.46	10.10	9.92
	N	116	114	113	112	113	112	114	111
Hispanic (Of Any Race)	Mean	939.01	938.26	938.13	933.69	938.72	936.33	938.54	936.87
	Std. Dev.	9.10	9.56	12.29	11.11	9.93	9.67	9.07	9.35
	N	880	877	877	859	873	854	873	849
Non-Hispanic American Indian	Mean	935.00	935.00	934.67	933.33	935.00	934.33	935.00	934.67
	Std. Dev.	19.08	19.08	21.36	21.73	19.92	19.86	19.08	19.86
	N	3	3	3	3	3	3	3	3
Non-Hispanic Multi-racial	Mean	937.50	937.50	935.79	932.21	936.71	935.21	937.50	935.43
	Std. Dev.	7.40	6.90	14.47	12.66	10.17	9.36	6.87	9.07
	N	14	14	14	14	14	14	14	14
Non-Hispanic White	Mean	937.70	936.75	938.50	932.47	938.34	935.04	937.03	935.79
	Std. Dev.	9.92	9.67	11.79	11.24	9.94	9.62	9.48	9.35
	N	123	123	121	121	121	121	123	121
Missing	Mean	939.13	938.63	938.17	935.27	938.93	937.30	938.94	937.56
	Std. Dev.	10.27	10.60	13.30	12.01	11.03	10.87	10.15	10.61
	N	71	71	70	71	70	71	70	70

Table 4.7.2 F
Mean Scale Scores by Ethnicity: Grade 6

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	937.04	937.82	937.19	933.34	937.60	935.90	937.60	936.48
	Std. Dev.	10.05	10.90	11.97	10.34	10.27	9.86	10.30	9.30
	N	153	154	151	152	151	152	153	149
Non-Hispanic Pacific Islander	Mean	939.83	941.67	937.50	931.75	939.08	936.92	941.00	937.33
	Std. Dev.	6.48	8.93	11.20	8.21	8.16	7.87	7.84	7.57
	N	12	12	12	12	12	12	12	12
Non-Hispanic Black	Mean	937.81	938.27	936.86	931.96	937.70	935.36	938.21	935.96
	Std. Dev.	9.98	12.93	13.47	10.54	11.37	11.46	11.57	11.03
	N	74	73	74	73	74	73	73	73
Hispanic (Of Any Race)	Mean	937.64	937.87	936.73	932.76	937.62	935.67	937.82	936.04
	Std. Dev.	10.04	11.24	12.81	10.15	10.83	9.89	10.55	9.81
	N	818	818	810	797	809	796	815	793
Non-Hispanic American Indian	Mean	942.50	942.83	939.50	931.67	941.50	937.33	942.67	938.50
	Std. Dev.	5.05	5.00	13.38	13.41	9.12	8.07	4.27	8.17
	N	6	6	6	6	6	6	6	6
Non-Hispanic Multi-racial	Mean	937.33	941.22	941.44	933.11	939.67	937.33	940.00	937.78
	Std. Dev.	6.61	5.63	6.29	7.15	6.42	5.50	5.61	5.56
	N	9	9	9	9	9	9	9	9
Non-Hispanic White	Mean	937.12	937.41	938.44	933.75	938.52	935.87	937.53	936.55
	Std. Dev.	10.42	11.06	11.86	10.89	10.39	10.20	10.27	9.75
	N	109	105	104	105	104	105	105	104
Missing	Mean	939.07	939.42	939.13	934.48	939.48	937.20	939.38	937.67
	Std. Dev.	9.96	11.14	11.75	11.07	10.54	10.62	10.55	10.27
	N	89	89	89	89	89	88	89	88

Table 4.7.2 G
Mean Scale Scores by Ethnicity: Grade 7

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	937.19	937.98	935.29	933.07	936.73	935.94	937.71	935.98
	Std. Dev.	10.44	12.34	14.28	11.45	11.79	11.04	11.50	10.97
	N	113	112	110	109	110	109	112	108
Non-Hispanic Pacific Islander	Mean	938.63	940.63	942.25	935.13	940.63	938.00	940.00	938.63
	Std. Dev.	6.61	11.72	7.25	8.37	6.35	9.12	9.90	8.09
	N	8	8	8	8	8	8	8	8
Non-Hispanic Black	Mean	936.15	936.92	935.34	931.58	936.31	934.60	936.78	934.85
	Std. Dev.	11.32	12.66	13.88	10.74	11.92	11.03	12.04	11.03
	N	86	85	83	81	83	81	85	81
Hispanic (Of Any Race)	Mean	938.77	939.01	937.61	933.42	938.67	936.49	939.01	936.97
	Std. Dev.	9.54	10.93	12.24	10.22	10.36	9.95	10.21	9.69
	N	770	768	763	760	759	757	767	751
Non-Hispanic American Indian	Mean	945.33	944.00	943.00	937.33	944.67	941.00	944.33	941.67
	Std. Dev.	1.15	2.00	6.93	4.04	4.04	3.00	1.15	1.53
	N	3	3	3	3	3	3	3	3
Non-Hispanic Multi-racial	Mean	936.71	942.14	942.57	935.71	940.00	939.14	940.43	939.29
	Std. Dev.	12.79	9.97	6.43	6.07	9.50	7.99	10.83	8.36
	N	7	7	7	7	7	7	7	7
Non-Hispanic White	Mean	939.17	939.27	939.41	934.01	939.79	936.89	939.30	937.48
	Std. Dev.	10.06	11.01	11.66	10.83	10.55	10.39	10.45	10.05
	N	92	91	91	90	91	90	91	90
Missing	Mean	938.65	938.27	938.92	934.41	939.31	936.83	938.48	937.51
	Std. Dev.	9.29	11.36	11.19	9.50	9.54	10.13	10.45	9.54
	N	48	49	49	49	48	48	48	47

Table 4.7.2 H
Mean Scale Scores by Ethnicity: Grade 8

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	936.52	938.06	936.12	935.01	936.68	937.02	937.78	936.72
	Std. Dev.	11.24	12.51	13.80	11.08	11.99	10.89	11.65	10.83
	N	116	118	116	116	116	116	116	115
Non-Hispanic Pacific Islander	Mean	937.56	939.22	941.00	934.11	941.38	937.00	938.78	940.25
	Std. Dev.	11.86	13.07	7.43	13.20	6.12	12.85	12.47	7.63
	N	9	9	8	9	8	9	9	8
Non-Hispanic Black	Mean	939.01	938.24	937.51	933.34	938.56	935.98	938.45	936.48
	Std. Dev.	9.35	11.24	13.41	10.59	10.83	10.33	10.51	10.13
	N	82	82	80	80	80	80	82	79
Hispanic (Of Any Race)	Mean	938.54	939.16	937.69	934.02	938.59	936.95	939.05	937.31
	Std. Dev.	9.87	11.48	12.77	10.24	10.75	10.13	10.68	9.92
	N	732	733	721	722	720	719	728	709
Non-Hispanic American Indian	Mean	946.00	944.00	947.00	941.00	947.00	943.00	945.00	944.00
	Std. Dev.
	N	1	1	1	1	1	1	1	1
Non-Hispanic Multi-racial	Mean	933.67	934.78	935.67	928.33	935.11	931.78	934.33	932.56
	Std. Dev.	14.66	17.13	13.79	11.47	12.88	13.58	16.03	13.10
	N	9	9	9	9	9	9	9	9
Non-Hispanic White	Mean	937.42	937.28	936.55	932.80	937.50	935.45	937.36	935.76
	Std. Dev.	11.57	13.78	13.85	11.56	11.94	12.20	12.86	11.91
	N	78	78	76	75	76	75	78	75
Missing	Mean	940.80	942.10	942.41	937.80	942.02	940.22	941.71	940.45
	Std. Dev.	8.03	7.94	7.21	7.11	7.10	6.93	7.57	6.51
	N	49	49	49	49	49	49	49	49

Table 4.7.2 I
Mean Scale Scores by Ethnicity: Grade 9

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	938.42	938.66	936.08	933.98	937.49	936.60	938.72	936.73
	Std. Dev.	9.37	9.51	11.15	9.62	9.49	8.79	9.27	8.74
	N	88	88	87	87	87	87	88	86
Non-Hispanic Pacific Islander	Mean	936.43	939.00	940.00	937.14	938.29	938.14	938.29	938.14
	Std. Dev.	13.36	13.30	13.23	12.54	12.88	12.75	13.12	12.71
	N	7	7	7	7	7	7	7	7
Non-Hispanic Black	Mean	936.46	936.54	937.51	932.47	937.31	934.62	936.56	935.11
	Std. Dev.	11.10	11.32	10.72	10.49	10.62	10.50	11.10	10.36
	N	56	56	55	55	55	55	55	54
Hispanic (Of Any Race)	Mean	939.04	939.15	937.40	935.10	938.56	937.53	939.17	937.62
	Std. Dev.	9.54	9.87	11.56	10.50	9.46	9.36	9.51	9.08
	N	553	556	547	544	543	539	551	536
Non-Hispanic American Indian	Mean	942.67	944.83	937.50	936.50	940.33	941.00	944.33	940.67
	Std. Dev.	4.27	3.60	10.52	8.43	6.71	6.03	3.44	5.99
	N	6	6	6	6	6	6	6	6
Non-Hispanic Multi-racial	Mean	940.17	938.00	939.83	941.67	940.17	940.00	938.67	939.83
	Std. Dev.	9.89	15.80	8.30	9.89	8.68	12.46	13.76	11.37
	N	6	6	6	6	6	6	6	6
Non-Hispanic White	Mean	934.05	934.43	934.62	932.87	934.99	934.31	934.60	934.51
	Std. Dev.	13.61	13.80	13.89	13.26	12.85	12.73	13.38	12.40
	N	74	72	71	68	71	68	72	68
Missing	Mean	938.10	938.56	938.63	934.72	938.72	936.82	938.60	937.22
	Std. Dev.	9.81	10.84	10.65	10.97	9.45	10.16	10.03	9.72
	N	80	78	78	76	78	76	78	76

Table 4.7.2 J
Mean Scale Scores by Ethnicity: Grade 10

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	939.92	940.83	938.12	937.45	939.33	939.32	940.69	939.12
	Std. Dev.	8.55	9.20	10.98	10.23	8.85	9.22	8.52	8.80
	N	86	86	84	85	84	85	85	84
Non-Hispanic Pacific Islander	Mean	946.50	946.50	944.25	942.00	945.50	944.50	946.50	944.50
	Std. Dev.	1.00	1.00	1.50	7.62	0.58	3.87	1.00	2.38
	N	4	4	4	4	4	4	4	4
Non-Hispanic Black	Mean	938.28	937.51	938.26	936.56	938.41	937.69	937.72	937.86
	Std. Dev.	9.15	8.78	9.70	7.76	8.33	7.16	8.70	6.78
	N	46	47	46	45	46	45	46	44
Hispanic (Of Any Race)	Mean	938.82	939.53	937.30	935.03	938.28	937.58	939.42	937.64
	Std. Dev.	9.92	9.87	11.40	10.86	9.97	9.65	9.66	9.40
	N	553	553	548	545	548	543	550	540
Non-Hispanic American Indian	Mean	933.67	933.67	932.83	931.83	933.33	932.83	933.67	932.83
	Std. Dev.	18.39	18.57	17.72	18.44	18.11	18.38	18.49	18.09
	N	6	6	6	6	6	6	6	6
Non-Hispanic Multi-racial	Mean	935.63	937.43	935.86	936.86	936.71	937.29	937.57	937.00
	Std. Dev.	8.31	9.20	7.69	11.38	6.82	9.36	7.89	8.10
	N	8	7	7	7	7	7	7	7
Non-Hispanic White	Mean	939.28	938.91	937.93	935.36	938.95	937.48	939.13	937.67
	Std. Dev.	9.48	9.23	10.26	10.49	8.70	9.13	8.93	8.65
	N	85	85	84	83	84	83	85	83
Missing	Mean	938.89	938.39	936.28	936.24	937.75	937.41	938.64	937.24
	Std. Dev.	8.28	9.55	12.74	9.64	9.53	9.12	8.94	9.18
	N	36	36	36	34	36	34	36	34

Table 4.7.2 K
Mean Scale Scores by Ethnicity: Grade 11

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	937.36	938.05	935.14	934.32	936.59	936.47	937.85	936.19
	Std. Dev.	10.17	10.56	13.42	11.10	10.70	10.39	10.20	10.26
	N	61	61	58	59	58	59	60	57
Non-Hispanic Pacific Islander	Mean	943.67	942.67	943.00	941.67	943.33	942.33	943.33	942.33
	Std. Dev.	5.77	5.03	3.46	6.03	4.62	4.62	5.03	4.62
	N	3	3	3	3	3	3	3	3
Non-Hispanic Black	Mean	938.85	938.29	939.61	933.67	939.39	936.27	938.71	937.09
	Std. Dev.	8.70	8.81	9.34	10.54	8.61	9.14	8.40	8.52
	N	46	45	46	45	46	45	45	45
Hispanic (Of Any Race)	Mean	940.39	940.59	938.19	935.52	939.55	938.30	940.68	938.60
	Std. Dev.	8.47	8.84	10.98	9.95	8.90	8.72	8.45	8.35
	N	489	487	486	485	482	480	485	475
Non-Hispanic American Indian	Mean	938.75	939.50	942.00	936.50	940.50	938.25	939.25	938.75
	Std. Dev.	5.19	8.85	4.24	7.33	4.43	7.80	7.89	6.65
	N	4	4	4	4	4	4	4	4
Non-Hispanic Multi-racial	Mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Std. Dev.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	0	0	0	0	0	0	0	0
Non-Hispanic White	Mean	938.23	938.53	936.24	932.41	937.53	935.70	938.63	936.10
	Std. Dev.	10.68	10.37	11.45	10.97	10.14	10.04	9.98	9.87
	N	74	73	72	69	72	69	73	69
Missing	Mean	938.77	937.54	936.10	934.74	937.56	936.26	937.95	936.37
	Std. Dev.	11.62	11.70	12.73	10.90	11.83	10.47	11.55	10.53
	N	39	39	39	38	39	38	39	38

Table 4.7.2 L
Mean Scale Scores by Ethnicity: Grade 12

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	938.48	937.94	936.32	935.14	937.56	936.71	938.20	936.83
	Std. Dev.	10.62	11.13	12.44	11.53	10.85	10.72	10.78	10.45
	N	134	134	133	133	133	133	134	132
Non-Hispanic Pacific Islander	Mean	940.00	941.75	933.38	933.63	936.75	937.88	941.38	937.38
	Std. Dev.	12.68	13.02	16.47	11.92	13.68	11.93	12.99	12.08
	N	8	8	8	8	8	8	8	8
Non-Hispanic Black	Mean	938.44	937.29	937.25	932.09	938.14	934.76	937.83	935.56
	Std. Dev.	8.66	10.29	10.01	10.50	8.45	10.06	9.33	9.29
	N	71	70	69	68	69	68	70	68
Hispanic (Of Any Race)	Mean	939.62	939.53	937.17	934.90	938.69	937.56	939.69	937.77
	Std. Dev.	9.38	9.83	11.67	10.43	9.68	9.42	9.43	9.19
	N	638	635	631	623	629	620	634	617
Non-Hispanic American Indian	Mean	939.33	942.00	930.33	928.00	935.00	935.00	941.33	935.00
	Std. Dev.	6.81	5.29	18.18	18.00	12.12	11.53	5.86	12.00
	N	3	3	3	3	3	3	3	3
Non-Hispanic Multi-racial	Mean	942.38	943.00	942.75	938.00	942.63	940.75	943.00	941.13
	Std. Dev.	6.02	6.50	4.46	10.39	4.93	7.52	6.09	6.62
	N	8	8	8	8	8	8	8	8
Non-Hispanic White	Mean	938.19	937.84	938.15	934.39	938.43	936.44	938.07	936.87
	Std. Dev.	11.42	11.13	11.36	11.88	11.00	10.91	10.95	10.65
	N	94	92	89	87	89	87	92	87
Missing	Mean	939.63	940.13	939.14	937.13	939.52	938.85	940.01	938.88
	Std. Dev.	9.34	9.81	10.21	10.73	9.33	9.78	9.59	9.48
	N	73	75	73	75	73	75	73	73

4.8 Correlations among Scale Scores by Grade-level Cluster

4.8.1 Correlations among Scale Scores: Grade-level Cluster 1-2

Table 4.8.1
Correlations Among Scale Scores: 1-2

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.849	0.652	0.735
	N	3,232	3,196	3,150	3,172
Reading	Pearson Correlation		1	0.705	0.721
	N		3,207	3,146	3,163
Writing	Pearson Correlation			1	0.692
	N			3,194	3,146
Speaking	Pearson Correlation				1
	N				3,193

** . Correlation is significant at the 0.05 level (2-tailed).

4.8.2 Correlations among Scale Scores: Grade-level Cluster 3-5

Table 4.8.2
Correlations Among Scale Scores: 3-5

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.861	0.671	0.747
	N	4,425	4,400	4,334	4,363
Reading	Pearson Correlation		1	0.759	0.764
	N		4,413	4,337	4,362
Writing	Pearson Correlation			1	0.709
	N			4,354	4,323
Speaking	Pearson Correlation				1
	N				4,377

** . Correlation is significant at the 0.05 level (2-tailed).

4.8.3 Correlations among Scale Scores: Grade-level Cluster 6-8

Table 4.8.3
Correlations Among Scale Scores: 6-8

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.863	0.703	0.772
	N	3,473	3,455	3,391	3,422
Reading	Pearson Correlation		1	0.750	0.770
	N		3,468	3,402	3,417
Writing	Pearson Correlation			1	0.724
	N			3,411	3,381
Speaking	Pearson Correlation				1
	N				3,429

** . Correlation is significant at the 0.05 level (2-tailed).

4.8.4 Correlations among Scale Scores: Grade-level Cluster 9-12

Table 4.8.4
Correlations Among Scale Scores: 9-12

		Listening	Reading	Writing	Speaking
Listening	Pearson Correlation	1	0.871	0.713	0.746
	N	3,439	3,413	3,342	3,384
Reading	Pearson Correlation		1	0.756	0.754
	N		3,430	3,351	3,382
Writing	Pearson Correlation			1	0.699
	N			3,366	3,339
Speaking	Pearson Correlation				1
	N				3,394

** . Correlation is significant at the 0.05 level (2-tailed).

4.9 Proficiency Levels

4.9.1 Proficiency Level by Grade-level Cluster

Table 4.9.1 A

Proficiency Level by Cluster: Listening

Cluster	Listening Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1-2	721	22.31	418	12.93	590	18.25	669	20.70	834	25.80	3,232
3-5	542	12.25	419	9.47	588	13.29	1,017	22.98	1,859	42.01	4,425
6-8	369	10.62	315	9.07	404	11.63	482	13.88	1,903	54.79	3,473
9-12	314	9.13	217	6.31	425	12.36	727	21.14	1,756	51.06	3,439
Total	1,946	13.36	1,369	9.40	2,007	13.78	2,895	19.87	6,352	43.60	14,569

Table 4.9.1 B

Proficiency Level by Cluster: Reading

Cluster	Reading Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1-2	773	24.10	475	14.81	582	18.15	632	19.71	745	23.23	3,207
3-5	551	12.49	562	12.74	633	14.34	1,042	23.61	1,625	36.82	4,413
6-8	414	11.94	246	7.09	388	11.19	588	16.96	1,832	52.83	3,468
9-12	311	9.07	261	7.61	364	10.61	631	18.40	1,863	54.31	3,430
Total	2,049	14.11	1,544	10.64	1,967	13.55	2,893	19.93	6,065	41.78	14,518

Table 4.9.1 C

Proficiency Level by Cluster: Speaking

Cluster	Speaking Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1-2	902	28.25	129	4.04	391	12.25	1,086	34.01	685	21.45	3,193
3-5	820	18.73	208	4.75	333	7.61	1,376	31.44	1,640	37.47	4,377
6-8	602	17.56	101	2.95	368	10.73	945	27.56	1,413	41.21	3,429
9-12	517	15.23	111	3.27	341	10.05	854	25.16	1,571	46.29	3,394
Total	2,841	19.74	549	3.81	1,433	9.96	4,261	29.60	5,309	36.89	14,393

Table 4.9.1 D

Proficiency Level by Cluster: Writing

Cluster	Writing Proficiency Range												Total
	A1		A2		A3		P1		P2		P3		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1-2	986	30.87	705	22.07	838	26.24	567	17.75	74	2.32	24	0.75	3,194
3-5	825	18.95	930	21.36	1,152	26.46	857	19.68	483	11.09	107	2.46	4,354
6-8	522	15.30	851	24.95	663	19.44	1,181	34.62	62	1.82	132	3.87	3,411
9-12	428	12.72	757	22.49	570	16.93	1,315	39.07	90	2.67	206	6.12	3,366
Total	2,761	19.27	3,243	22.64	3,223	22.50	3,920	27.36	709	4.95	469	3.27	14,325

Table 4.9.1 E

Proficiency Level by Cluster: Oral

Cluster	Oral Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1-2	811	25.57	282	8.89	454	14.31	879	27.71	746	23.52	3,172
3-5	680	15.59	258	5.91	566	12.97	1,161	26.61	1,698	38.92	4,363
6-8	491	14.35	216	6.31	393	11.48	713	20.84	1,609	47.02	3,422
9-12	395	11.67	169	4.99	410	12.12	947	27.98	1,463	43.23	3,384
Total	2,377	16.57	925	6.45	1,823	12.71	3,700	25.80	5,516	38.46	14,341

Table 4.9.1 F

Proficiency Level by Cluster: Literacy

Cluster	Literacy Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1-2	830	26.38	621	19.74	793	25.21	551	17.51	351	11.16	3,146
3-5	630	14.53	758	17.48	966	22.27	1,111	25.62	872	20.11	4,337
6-8	436	12.82	384	11.29	730	21.46	1,085	31.89	767	22.55	3,402
9-12	324	9.67	376	11.22	720	21.49	1,014	30.26	917	27.36	3,351
Total	2,220	15.59	2,139	15.03	3,209	22.54	3,761	26.42	2,907	20.42	14,236

Table 4.9.1 G

Proficiency Level by Cluster: Comprehension

Cluster	Comprehension Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1-2	732	22.90	445	13.92	577	18.05	776	24.28	666	20.84	3,196
3-5	516	11.73	531	12.07	591	13.43	977	22.20	1,785	40.57	4,400
6-8	401	11.61	245	7.09	345	9.99	676	19.57	1,788	51.75	3,455
9-12	295	8.64	239	7.00	351	10.28	700	20.51	1,828	53.56	3,413
Total	1,944	13.44	1,460	10.09	1,864	12.89	3,129	21.63	6,067	41.95	14,464

Table 4.9.1 H

Proficiency Level by Cluster: Overall

Cluster	Overall Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1-2	773	24.79	444	14.24	847	27.16	669	21.46	385	12.35	3,118
3-5	588	13.66	521	12.11	966	22.44	1,248	29.00	981	22.79	4,304
6-8	417	12.36	294	8.71	638	18.91	1,089	32.28	936	27.74	3,374
9-12	318	9.56	270	8.11	717	21.54	945	28.40	1,078	32.39	3,328
Total	2,096	14.84	1,529	10.83	3,168	22.43	3,951	27.97	3,380	23.93	14,124

4.9.2 Proficiency Level by Grade

Table 4.9.2 A

Proficiency Level by Grade: Listening

Grade	Listening Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	386	24.63	223	14.23	300	19.14	329	21.00	329	21.00	1,567
2	335	20.12	195	11.71	290	17.42	340	20.42	505	30.33	1,665
3	230	14.94	178	11.57	249	16.18	352	22.87	530	34.44	1,539
4	181	11.99	124	8.21	187	12.38	349	23.11	669	44.30	1,510
5	131	9.52	117	8.50	152	11.05	316	22.97	660	47.97	1,376
6	139	10.94	128	10.08	171	13.46	174	13.70	658	51.81	1,270
7	114	10.12	103	9.14	116	10.29	154	13.66	640	56.79	1,127
8	116	10.78	84	7.81	117	10.87	154	14.31	605	56.23	1,076
9	93	10.69	65	7.47	97	11.15	196	22.53	419	48.16	870
10	72	8.74	52	6.31	127	15.41	157	19.05	416	50.49	824
11	58	8.10	39	5.45	86	12.01	153	21.37	380	53.07	716
12	91	8.84	61	5.93	115	11.18	221	21.48	541	52.58	1,029
Total	1,946	13.36	1,369	9.40	2,007	13.78	2,895	19.87	6,352	43.60	14,569

Table 4.9.2 B

Proficiency Level by Grade: Reading

Grade	Reading Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	415	26.74	240	15.46	306	19.72	291	18.75	300	19.33	1,552
2	358	21.63	235	14.20	276	16.68	341	20.60	445	26.89	1,655
3	231	15.08	250	16.32	245	15.99	384	25.07	422	27.55	1,532
4	182	12.05	170	11.26	216	14.30	354	23.44	588	38.94	1,510
5	138	10.07	142	10.36	172	12.55	304	22.17	615	44.86	1,371
6	151	11.93	90	7.11	167	13.19	241	19.04	617	48.74	1,266
7	129	11.49	82	7.30	117	10.42	186	16.56	609	54.23	1,123
8	134	12.42	74	6.86	104	9.64	161	14.92	606	56.16	1,079
9	88	10.13	76	8.75	92	10.59	159	18.30	454	52.24	869
10	65	7.89	64	7.77	96	11.65	144	17.48	455	55.22	824
11	61	8.57	51	7.16	68	9.55	127	17.84	405	56.88	712
12	97	9.46	70	6.83	108	10.54	201	19.61	549	53.56	1,025
Total	2,049	14.11	1,544	10.64	1,967	13.55	2,893	19.93	6,065	41.78	14,518

Table 4.9.2 C

Proficiency Level by Grade: Speaking

Grade	Speaking Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	490	31.49	63	4.05	188	12.08	539	34.64	276	17.74	1,556
2	412	25.17	66	4.03	203	12.40	547	33.41	409	24.98	1,637
3	332	21.86	87	5.73	133	8.76	511	33.64	456	30.02	1,519
4	271	18.14	60	4.02	104	6.96	452	30.25	607	40.63	1,494
5	217	15.91	61	4.47	96	7.04	413	30.28	577	42.30	1,364
6	230	18.33	28	2.23	150	11.95	357	28.45	490	39.04	1,255
7	189	16.97	45	4.04	106	9.52	324	29.08	450	40.39	1,114
8	183	17.26	28	2.64	112	10.57	264	24.91	473	44.62	1,060
9	139	16.22	25	2.92	76	8.87	244	28.47	373	43.52	857
10	121	14.85	28	3.44	89	10.92	207	25.40	370	45.40	815
11	100	14.12	26	3.67	73	10.31	155	21.89	354	50.00	708
12	157	15.48	32	3.16	103	10.16	248	24.46	474	46.75	1,014
Total	2,841	19.74	549	3.81	1,433	9.96	4,261	29.60	5,309	36.89	14,393

Table 4.9.2 D

Proficiency Level by Grade: Writing

Grade	Writing Proficiency Range												Total
	A1		A2		A3		P1		P2		P3		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	544	34.85	362	23.19	394	25.24	231	14.80	20	1.28	10	0.64	1,561
2	442	27.07	343	21.00	444	27.19	336	20.58	54	3.31	14	0.86	1,633
3	329	21.74	376	24.85	415	27.43	254	16.79	120	7.93	19	1.26	1,513
4	273	18.29	301	20.16	378	25.32	313	20.96	188	12.59	40	2.68	1,493
5	223	16.54	253	18.77	359	26.63	290	21.51	175	12.98	48	3.56	1,348
6	205	16.49	317	25.50	244	19.63	414	33.31	18	1.45	45	3.62	1,243
7	163	14.72	278	25.11	232	20.96	371	33.51	23	2.08	40	3.61	1,107
8	154	14.51	256	24.13	187	17.62	396	37.32	21	1.98	47	4.43	1,061
9	111	13.07	188	22.14	153	18.02	323	38.04	22	2.59	52	6.12	849
10	100	12.36	170	21.01	137	16.93	323	39.93	25	3.09	54	6.67	809
11	81	11.52	169	24.04	129	18.35	262	37.27	18	2.56	44	6.26	703
12	136	13.53	230	22.89	151	15.02	407	40.50	25	2.49	56	5.57	1,005
Total	2,761	19.27	3,243	22.64	3,223	22.50	3,920	27.36	709	4.95	469	3.27	14,325

Table 4.9.2 E

Proficiency Level by Grade: Oral

Grade	Oral Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	439	28.40	146	9.44	231	14.94	445	28.78	285	18.43	1,546
2	372	22.88	136	8.36	223	13.71	434	26.69	461	28.35	1,626
3	275	18.14	117	7.72	235	15.50	420	27.70	469	30.94	1,516
4	227	15.27	76	5.11	179	12.04	387	26.03	618	41.56	1,487
5	178	13.09	65	4.78	152	11.18	354	26.03	611	44.93	1,360
6	183	14.59	82	6.54	160	12.76	274	21.85	555	44.26	1,254
7	153	13.80	75	6.76	111	10.01	249	22.45	521	46.98	1,109
8	155	14.64	59	5.57	122	11.52	190	17.94	533	50.33	1,059
9	106	12.43	43	5.04	104	12.19	259	30.36	341	39.98	853
10	97	11.90	38	4.66	107	13.13	219	26.87	354	43.44	815
11	74	10.51	35	4.97	82	11.65	193	27.41	320	45.45	704
12	118	11.66	53	5.24	117	11.56	276	27.27	448	44.27	1,012
Total	2,377	16.57	925	6.45	1,823	12.71	3,700	25.80	5,516	38.46	14,341

Table 4.9.2 F

Proficiency Level by Grade: Literacy

Grade	Literacy Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	451	29.50	311	20.34	400	26.16	239	15.63	128	8.37	1,529
2	379	23.44	310	19.17	393	24.30	312	19.29	223	13.79	1,617
3	267	17.72	306	20.31	367	24.35	374	24.82	193	12.81	1,507
4	211	14.18	239	16.06	322	21.64	368	24.73	348	23.39	1,488
5	152	11.33	213	15.87	277	20.64	369	27.50	331	24.66	1,342
6	168	13.54	147	11.85	287	23.13	399	32.15	240	19.34	1,241
7	136	12.33	134	12.15	226	20.49	362	32.82	245	22.21	1,103
8	132	12.48	103	9.74	217	20.51	324	30.62	282	26.65	1,058
9	88	10.43	93	11.02	193	22.87	266	31.52	204	24.17	844
10	70	8.67	96	11.90	168	20.82	250	30.98	223	27.63	807
11	65	9.31	74	10.60	153	21.92	204	29.23	202	28.94	698
12	101	10.08	113	11.28	206	20.56	294	29.34	288	28.74	1,002
Total	2,220	15.59	2,139	15.03	3,209	22.54	3,761	26.42	2,907	20.42	14,236

Table 4.9.2 G

Proficiency Level by Grade: Comprehension

Grade	Comprehension Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	389	25.13	231	14.92	303	19.57	366	23.64	259	16.73	1,548
2	343	20.81	214	12.99	274	16.63	410	24.88	407	24.70	1,648
3	213	13.92	241	15.75	235	15.36	362	23.66	479	31.31	1,530
4	176	11.69	162	10.76	190	12.62	340	22.59	637	42.33	1,505
5	127	9.30	128	9.38	166	12.16	275	20.15	669	49.01	1,365
6	146	11.57	97	7.69	140	11.09	280	22.19	599	47.46	1,262
7	124	11.06	79	7.05	110	9.81	205	18.29	603	53.79	1,121
8	131	12.22	69	6.44	95	8.86	191	17.82	586	54.66	1,072
9	83	9.62	68	7.88	93	10.78	181	20.97	438	50.75	863
10	64	7.81	62	7.57	89	10.87	155	18.93	449	54.82	819
11	54	7.62	46	6.49	68	9.59	141	19.89	400	56.42	709
12	94	9.20	63	6.16	101	9.88	223	21.82	541	52.94	1,022
Total	1,944	13.44	1,460	10.09	1,864	12.89	3,129	21.63	6,067	41.95	14,464

Table 4.9.2 H

Proficiency Level by Grade: Overall

Grade	Overall Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	423	27.79	230	15.11	442	29.04	289	18.99	138	9.07	1,522
2	350	21.93	214	13.41	405	25.38	380	23.81	247	15.48	1,596
3	249	16.63	215	14.36	385	25.72	433	28.92	215	14.36	1,497
4	201	13.65	151	10.25	323	21.93	416	28.24	382	25.93	1,473
5	138	10.34	155	11.62	258	19.34	399	29.91	384	28.79	1,334
6	158	12.80	115	9.32	251	20.34	412	33.39	298	24.15	1,234
7	133	12.15	95	8.68	204	18.63	355	32.42	308	28.13	1,095
8	126	12.06	84	8.04	183	17.51	322	30.81	330	31.58	1,045
9	90	10.73	69	8.22	181	21.57	255	30.39	244	29.08	839
10	70	8.73	71	8.85	171	21.32	225	28.05	265	33.04	802
11	61	8.83	49	7.09	151	21.85	196	28.36	234	33.86	691
12	97	9.74	81	8.13	214	21.49	269	27.01	335	33.63	996
Total	2,096	14.84	1,529	10.83	3,168	22.43	3,951	27.97	3,380	23.93	14,124

4.10 Participation by Disability

4.10.1 Participation by Disability

Table 4.10.1
Participation by Disability

		Secondary Disability														Total
		AS	DB	DD	HI	ID	MD	NSD	OHI	OI	SED	SLD	SLI	TBI	VI	
Primary Disability	AS	5	.	19	12	297	18	2,852	69	5	7	54	1,114	.	6	4,458
	DB	1	.	1	.	2	.	5	1	.	.	.	2	.	.	12
	DD	21	1	.	9	17	5	522	23	9	.	14	256	.	11	888
	HI	2	1	1	1	17	2	37	2	3	1	2	21	.	.	90
	ID	125	14	52	63	11	50	3,347	272	105	30	61	1,535	5	60	5,730
	MD	10	.	3	2	25	35	212	20	7	1	3	118	1	3	440
	NPD	3	.	2	2	.	6	1,668	1	.	.	.	16	.	1	1,699
	OHI	12	3	7	14	54	8	363	2	10	1	8	150	2	8	642
	OI	1	1	1	1	21	2	38	5	.	.	1	21	.	1	93
	SED	.	.	.	1	6	.	12	3	.	.	.	3	.	.	25
	SLD	2	.	3	1	5	1	226	5	.	3	2	83	1	1	333
	SLI	6	1	3	.	17	1	127	6	1	1	6	6	1	.	176
TBI	2	.	.	.	9	.	40	2	1	.	.	16	1	4	75	
VI	1	.	3	1	6	1	9	4	.	.	.	1	.	1	27	
Total		191	21	95	107	487	129	9,458	415	141	44	151	3,342	11	96	14,688

Acronyms for Table 4.10.1

Acronym	Category Name
NPD	No Primary Disability Recorded
NSD	No Secondary Disability Recorded
AS	Autism Spectrum Disorder
DB	Deaf-blindness
DD	Developmental Delay
HI	Hearing Impairment, including Deafness
ID	Intellectual Disability
MD	Multiple Disability
OI	Orthopedic Impairment
OHI	Other Health Impairment
SED	Serious Emotional Disability
SLD	Specific Learning Disability
SLI	Speech or Language Impairment
TBI	Traumatic Brain Injury
VI	Visual Impairment, including Blindness

5. Analyses of Test Forms: Overview

This chapter contains two parts. The first part provides some background on the technical measurement and statistical tools used to analyze Alternate ACCESS for ELLs. The second part explains the results that are presented for each test form in Chapter 6.

5.1 Background

5.1.1 Measurement Models Used

The measurement model that forms the basis of the analysis for the development of Alternate ACCESS for ELLs is the Rasch measurement model (Wright and Stone, 1979). Additional information on its use in the development of the test is available in WIDA Technical Report 1, *Alternate ACCESS for ELLs™, Series 100 Development and Operational Field Test: Technical Report*. The test was developed using Rasch measurement principles, and in that sense the Rasch model guided all decisions throughout the development of the assessment and was not just a tool for the statistical analysis of the data. For example, data based on Rasch fit statistics guided the inclusion, revision, or deletion of items during the development and field testing of the test forms and will continue to guide the refinement and further development of the test.

For all domains, a Rasch Rating Scale model was used. Mathematically, this can be represented as

$$\log\left(\frac{P_{nik}}{P_{nik-1}}\right) = B_n - D_i - F_k$$

where

P_{nik} = probability of person “n” on task “i” receiving a rating at level “k” on the rating scale

P_{nik-1} = probability of person “n” on task “i” receiving a rating at level “k - 1” on the rating scale (i.e., the next lowest rating)

B_n = ability of person “n”

D_i = difficulty of task “i”

F_k = calibration of step “k” on the rating scale

All Rasch analyses were conducted using the Rasch measurement software program *Winsteps* (Linacre, 2006). When speaking of the measure of examinee ability, we use the term “ability measure” (rather than *theta*, which is used commonly when discussing models based on Item Response Theory [IRT]). When speaking of the measure of how hard an item was, we use the term “item difficulty measure” (rather than the term *b parameter*, which is used commonly when discussing models based on IRT). “Step measures” refer to the calibration of the steps in the Rasch Rating Scale model presented above. All three measures (ability, difficulty, and step) are expressed in terms of Rasch logits, which then are converted into scores on the Alternate ACCESS for ELLs score scale for reporting purposes (see WIDA Technical Report 1 for more details).

Rasch model standard errors also appear in the tables. These are an indication of the precision with which the measures have been estimated. Unlike the standard error of measurement (SEM) based on classical test theory, which posits the same SEM for all persons regardless of their position on the ability distribution, Rasch model standard errors are conditional on the individual's ability measure. All things being equal, if a person gets few items correct or few items incorrect, the standard error of that person's measure will be greater than if a person gets a moderate number of items correct. In addition, for ability measures, standard errors are a function of the number of items on a test form as well as the distribution and quality of the items (i.e., their fit to the Rasch model).

Fit statistics for the Rasch model are provided in Chapter 6. These statistics are calculated by comparing the observed empirical data with the data that would be expected to be produced by the Rasch model. Of the several statistics available, the mean square fit statistics were used to flag items in the development of Alternate ACCESS for ELLs that needed to be deleted or revised. Outfit mean square statistics are more sensitive to outliers. For example, a difficult item that some low ability examinees get correct will have a high outfit mean square statistic that indicates that the item may not be measuring the same thing as other items on the test. Infit mean square statistics are influenced by more aberrant response patterns and generally indicate a more serious measurement problem. The expectation for both of these statistics is 1.00 and values near are not of great concern. Values less than 1.00 indicate that the observations are too predictable and thus redundant, but are not of great concern. High values are more of a concern.

According to Linacre (2002):

values greater than 2.0 “distort or degrade the measurement system”

values between 1.5 and 2.0 are “unproductive for construction of measurement, but not

degrading” values between 0.5 and 1.5 should be considered “productive for measurement”

values below 0.5 are considered “less productive for measurement, but not degrading”

Because conservative guidelines were followed in the development of Alternate ACCESS for ELLs, the vast majority of items and tasks on the test forms have mean square fit statistics in the range of 0.75 and 1.25 and therefore fall within the range that is “productive for measurement” according to the guidelines above.

5.1.2 Sampling

The results presented in most of the tables in Chapter 6 are based on the full data set of all students who were administered operational Series 502 of Alternate ACCESS for ELLs in the academic year 2019-2020. The item analysis summary tables (Table F), the complete item analysis tables (Table G), and the raw score to scale score conversion tables (Table H) use item difficulties from this calibration.

5.1.3 Scaling

Complete information on the horizontal and vertical scaling of Alternate ACCESS for ELLs scores is provided in Technical Report 1, *Alternate Access for ELLs™ Series 100 Development and Operational Field Test: Technical Report*. In brief, this scaling was accomplished during the field test based on an elaborate common item design, across grade-level clusters, which spanned two series of complete test forms. Concurrent calibration was used to determine item difficulty measures. These item difficulty measures were used to create the Alternate ACCESS for ELLs scale scores used for reporting results on the test.

Table 5.1.3A provides the scaling equation for each domain. This equation is used to convert an examinee’s ability measure into the scale score. Since Alternate ACCESS for ELLs is vertically equated, though each domain has its own equation, the same equation is used across all grade- level clusters within each domain.

Table 5.1.3A

Scaling Equation for each Domain

Domain	Scale Score
Listening	$(\text{Ability Measure in Logits} * 7.913) + 925.056$
Reading	$(\text{Ability Measure in Logits} * 6.026) + 925.788$
Speaking	$(\text{Ability Measure in Logits} * 4.433) + 924.531$
Writing	$(\text{Ability Measure in Logits} * 2.4) + 926.408$

5.1.4 DIF Analyses

Differential item analyses (DIF) attempt to investigate whether performances on items or tasks were influenced by factors extraneous to English language proficiency (i.e., the construct being measured on the test). In other words, it attempts to find items or tasks that may be functioning differently for different groups based on criteria irrelevant to what is being tested. The performance of students on the Alternate ACCESS for ELLs tasks was compared by dividing students into two different groupings: first, males versus females; second, students of Hispanic ethnic background versus students of non-Hispanic ethnic background (For both analyses, students for whom test scores and gender or ethnicity was missing were excluded). The underlying assumption of DIF analysis is that students who performed similarly overall on the test should perform similarly on the individual tasks. To test this assumption, students are initially placed into groups based on their total raw scores by domain. Then, student performance on a task of interest within that domain, the studied item, is compared between groups.

The Mantel Chi-square statistic and the standardized P-DIF (i.e., the DIF procedure used for polytomous items) or the standardized mean difference (SMD) procedures developed by the Education Testing Service (ETS) (Zwick, Donoghue, & Grima, 1993; Allen, Carlson, & Zalanak, 1999) for polytomous items were used for identifying tasks that exhibit DIF. JMetrik (Meyer, 2014), an open source computer program for psychometric analysis, was used in conducting the analyses. The procedures first calculate the Mantel statistic and determine its probability of significance. This statistic gives an indication of the probability that observed differences are the result of chance but does not indicate how significant that difference is. To indicate how significant the difference is, the SMD between the performances of the two groups being compared is calculated. The SMD compares the means of the two groups, adjusting for differences in the distribution of the two groups being compared across the values of the total raw scores. To standardize the outcome, this difference is divided by the standard deviation (SD) of the task for the total group. The ratio of SMD over SD serves as an effect size measure for the Mantel Chi-square statistic. Since this effect size measure can be positive or negative which may present some challenges when interpreting them, it is divided by the item score range in JMetrik (Meyer, 2014) such that the range of the rescaled effect size (called standardized P-DIF* on the JMetrik DIF output) is restricted to 0 and 1. The effect size flagging criterion for polytomous items, proposed by ETS (Allen, Carlson, & Zalanak, 1999) was also rescaled to the standardized P-DIF* metric (Meyer, 2014).

Following guidance proposed by ETS for NAEP assessment (Allen, Carlson, & Zalanak, 1999), Alternate ACCESS for ELLs tasks are classified into three DIF levels as follows:

- AA (no DIF), when the Mantel Chi-square statistic is not significant or when it is significant and standardized P-DIF* is less than 0.05
- BB (weak DIF), when the Mantel Chi-square statistic is significant and standardized P-DIF* is greater than or equal to 0.05 but less than 0.10
- CC (strong DIF), when the Mantel Chi-square statistic is significant and standardized P-DIF* is greater than or equal to 0.10

5.1.5 Reliability of Composites

Four composite scores are reported for Alternate ACCESS: Oral Language Composite (oral), Literacy Composite (litr), Comprehension Composite (cphn), and Overall Composite (over). To estimate the reliability of these composite scores, a stratified Cronbach's alpha coefficient (e.g., Kamata, Turhan, & Darandari, 2003; April, Kane, & Case, 2004; Rudner, 2001) is computed, weighted by the contribution of each domain score into the composite. Specifically, the formula is

$$\alpha_c = 1 - \frac{\sum_{j=1}^k w_j^2 \sigma_j^2 (1 - \rho_j)}{\sigma_c^2}$$

Where

k = number of components j

w_j = domain weight of component j

σ_j^2 = variance of component j

σ_c^2 = variance of composite

ρ_j = reliability coefficient of component j .

The data to compute the stratified Cronbach's alpha is provided in the appropriate tables in Chapter 6.

5.1.6 Accuracy and Consistency of Classification

For each domain across grade-level clusters, as well as for the four composite scores, tables were produced that indicate estimates of the accuracy and consistency of classification of examinees into the Alternate ACCESS for ELLs language proficiency levels based on their performances on the test. It is important to know the reliability of any student's test score and the degree of precision with which it has been measured (i.e., the estimate of the invariant standard error of measure [SEM] of classical test theory and the estimate of the variable conditional standard error of the Rasch measurement model). However, because decisions about students are ultimately made on the basis of their classification into language proficiency levels on the basis of their performance on Alternate ACCESS for ELLs[®], it is important to know how well these classifications are made. The analyses that we employed make use of the methods outlined and implemented in Livingston and Lewis (1995) and Young and Yoon (1998) as implemented in the software program BB-CLASS (Brennan, 2004) (cf. also Lee, Hanson, & Brennan, 2002).

In the approach of Livingston and Lewis (1995), the accuracy of a decision is the extent to which decisions made on the basis of the administered test (i.e., the observed scores) would agree with the decisions that would be made if each student could somehow be tested with all possible parallel forms of the assessments; that is, decisions based on the examinees' "true score." On the other hand, the consistency of a decision is the extent to which decisions made on the basis of the administered test would agree with the decisions that would be made if the students had taken a different but parallel form of the test. Thus, in every analysis of classification, two parallel analyses are made: accuracy (that is, vis-à-vis "true scores") and consistency (that is, vis-à-vis a second form).

In terms of classifications around a single cut point, students can be misclassified in one of two ways. Students who were below the proficiency cut score (based on their "true score"), but were classified on the basis of the assessment as being above the cut score, are considered to be false positives. Students who were above the proficiency cut score (based on their "true score"), but were classified as being below a cut score, are considered to be false negatives. All other students are considered to be accurately placed either above or below the cut score.

Since a 'true score' is a theoretical construct, it is unknown for any given student. The approach taken by Livingston and Lewis (1995) and implemented here *to model true scores* uses information about the reliability of the test, the cut scores, and the observed distribution of scores. Then, using a four-parameter beta distribution, we modeled the distribution of the true scores and of scores on a parallel form. Overall accuracy and consistency indices are produced by comparing the percentage of students classified across all categories the same way by both the observed distribution and modeled distribution. These indices indicate the percent of all students who would be classified into the same language proficiency level by both the administered test and either the true score distribution (accuracy) or a parallel test (consistency). Our tables also provide an estimate of Cohen's kappa statistic, which is a very conservative estimate of the overall classification since it corrects for chance.

We also look at accuracy and consistency conditional on the language proficiency level. These indices examine the percent of students classified by both tests into a level divided by all students classified into that level according either to the true score distribution (accuracy) or based on a parallel test (consistency).

Finally, we look at what may be the most important set of indices, which are the indices at the cut points. That is, at every cut point, using the true score distribution (e.g., accuracy), we provide the percent of students who are consistently placed above and below the cut score, as well as those who are false positives and false negatives. For consistency, only the percent of students classified consistently above and below the cut score is calculated. Thus, for example, to evaluate the degree of confidence that one can have in a decision made based on the Overall Composite score as to whether students are being accurately classified into Alternate WIDA language proficiency level P2 (“Beginning”) or not, one can look at the accuracy index provided in the table for the cut score P1/P2.

5.2 Descriptions

The following paragraphs describe the tables and figures that appear in Chapter 6. Each description applies to each test form in each domain. Information on raw and scale score descriptive statistics, proficiency level distribution, and the equating summary, are displayed in tables/figures A-D. Reliability, item analysis summary, complete item analysis, raw score to scale score conversion, and raw score to proficiency level conversion tables are provided in tables E-I. These tables are organized by: grade, grade-level cluster, domain, domain and composite scores.

Note that because the composite scores do not have raw scores associated with them, any table or figure that draws on raw scores is not included for the composite scores. This includes Table A, Table D, Table F, Table G, Table H and Table I, and Figure A, Figure D and Figure E.

5.2.1 Raw Score Information (Figure A and Table A)

Figure A and Table A relate to the raw scores on each test form (the raw score to proficiency level conversion table for each test form is displayed in Table I in each section). All domains were scored polytomously. The highest possible score for Listening and Reading is 36 (4 points per item for 9 items). The highest possible score for Speaking is 16 (2 points per item for 8 items). The highest possible score for Writing is 24 (Writing parts A & B: 2 points per item for 8 items; Writing part C: 4 points per item for 2 items). For each test form, Figure A shows the distribution of the raw scores. The horizontal axis shows the raw scores. The vertical axis shows the number of students (count). Each bar shows how many students were awarded each raw score.

Table A shows the following information, by each grade in the cluster and by total for the cluster:

- The number of students in the analyses (the number of students who were not absent, invalid, refused, exempt, or in the wrong cluster)
- The minimum observed raw score
- The maximum observed raw score
- The mean (average) raw score
- The standard deviation (std. dev.) of the raw scores

5.2.2 Scale Score Information (Figure B and Table B)

Figure B and Table B relate to the *scale scores* on each test form. For each test form, raw scores were converted to vertically-equated scale scores. The raw score to scale score conversion table for each test form is displayed in Table H in each section. Thus, for each test form, Figure B shows the distribution of the scale scores. The horizontal axis shows the scale scores based on performances on the test form. The vertical axis shows the number of students (count). Each bar shows how many students were awarded each scale score.

Table B shows the following information, by each grade in the cluster and by total for the cluster:

- Number of students in the analyses
- The minimum observed scale score
- The maximum observed scale score
- The mean (average) scale score
- The standard deviation (std. dev.) of the scale scores

5.2.3 Proficiency Level Information (Figure C and Table C)

Figure C and Table C provide information on the proficiency level distribution of the students who took the test form based on their performance. Thus, for each test form, Figure C shows the information graphically for the cluster as a whole. The horizontal axis shows five out of six Alternate WIDA proficiency levels.⁴ The vertical axis shows the percent of students. Each bar shows the percent of students who were placed into each proficiency level in the domain being tested on this test form.

Table C shows the following information, by each grade in the cluster and by total for the cluster:

- The Alternate WIDA proficiency level designation (A1-A3;P1-P2)
- The number of students (count) whose performance on the test form placed them into that proficiency level in the domain being tested
- The percent of students, out of the total number of students taking the form (by grade or by total for the cluster), who were placed into that proficiency level in the domain being tested

5.2.4 Equating Summary Table (Table D)

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the Series 100 field test. Thus, the results from the original field test of Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the *Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report* (2013).

5.2.5 Reliability (Table E)

Table E presents reliability information based on Classical Test Theory and shows the following information:

- The number of students
- The number of items

⁴ In Series 502, only the Alternate WIDA proficiency levels A1, A2, A3, P1 and P2 were reported. In Series 102, the proficiency level P3 will be reported as well.

- Cronbach’s coefficient alpha (as a measure of internal consistency)
- The classical standard error of measurement (SEM) in terms of *raw scores*

Cronbach’s coefficient alpha is widely used as an estimate of reliability, particularly of the internal consistency of test items. It expresses how well the items on a test appear to measure the same construct. Conceptually, it may be thought of as the correlation obtained between performances on two halves of the test, if every possibility of dividing the test items in two were attempted. Thus, Cronbach’s alpha may be low if some items are measuring something other than what the majority of the items are measuring. As with any reliability index, it is affected by the number of test items (or test score points that may be awarded). That is, all things being equal, the greater the number of items, the higher the reliability.

Cronbach’s alpha is also affected by the distribution of ability within the group of students tested. All things being equal, the greater the heterogeneity of abilities within the group of students tested (i.e., the more widely the scores are distributed), the higher the reliability. In this sense, Cronbach’s alpha is sample dependent. It is widely recognized that reliability can be as much a function of the test as of the sample of students tested. That is, the exact same test can produce widely disparate reliability indices based on ability distribution of the group of students tested.

The formula for Cronbach’s alpha is

$$\alpha = \frac{n}{n-1} \left[1 - \frac{\sum_{i=1}^n \sigma_i^2}{\sigma_t^2} \right]$$

where

n = number of items i

σ_i^2 = variance of score on item i

σ_t^2 = variance of total score

Table E also presents the *standard error of measurement* (SEM) based on classical test theory. Unlike IRT, in this approach, SEM is seen as a constant across the spread of test scores (ability continuum). Thus, it is *not* conditional on ability being measured. It is, however, a function of two statistics: the reliability of the test and the (observed) standard deviation of the test scores. It is calculated as

$$SEM = SD \sqrt{1 - reliability}$$

Traditionally, SEM has been used to create a band around an examinee’s observed score. The assertion in the view of classical test theory is that the examinee’s true score (i.e., what the examinee’s score would be if it could be measured without error) would lie with a certain degree of probability within this band. Therefore, the statistical expectation is that an examinee’s true score has a 68% probability of lying within the band, extending from the observed score minus 1 SEM to the observed score plus 1 SEM.

5.2.6 Test Characteristic Curve (Figure D)

For each test form, Figure D graphically shows the relationship between the ability measure (in logits) on the horizontal axis and the expected raw score on the vertical axis. Four vertical lines indicate the four cut scores, dividing the figure into five sections for each of the WIDA proficiency levels (A1-A3; P1-P2) for the domain being tested. As would be expected, higher raw scores are required to be placed into higher language proficiency levels. The relative width of each section between the cut score lines, however, gives an indication of how many points must be earned to be placed into a WIDA language proficiency level.

5.2.7 Test Information Function (Figure E)

With the Rasch measurement model, as with any measurement model following Item Response Theory (IRT), the relationship between the ability measure (in logits) and the accuracy of test scores can be modeled. It is recognized that tests measure most accurately when the abilities of the examinees and the difficulty of the items are most appropriate for each other. If a test is too difficult for an examinee (i.e., the examinee scores close to zero), or if the test is too easy for an examinee (i.e., the examinee “tops out”), accurate measurement of the examinee’s ability cannot be made. The test information function shows graphically how well the test is measuring across the ability measure spectrum. High values indicate more accuracy in measurement. Thus, for each test form, Figure E shows the relationship between the ability measure (in logits) on the horizontal axis and measurement accuracy, represented as the Fisher information value (which is the inverse squared of the standard error), on the vertical axis. The test information function, then, reflects the conditional standard error of measurement.

Again, as in Figure D, four vertical lines in Figure E indicate the four cut scores, dividing the figure into five sections for each of the WIDA language proficiency levels (A1-A3:P1-P2) for the domain being tested. It is important that each test form measure most accurately in the areas for which it is primarily used to make classification decisions. In other words, optimally the test information function should be high for the cuts between A1/A2, A2/A3, A3/P1, and P1/P2.

5.2.8 Item Analysis Summary (Table F)

Table F provides a summary of the analyses of the items. This table is divided into two parts: one, the item summary; two, the DIF summary. The upper half of the table displays the item summary. The first column in this part states the type of item (MOSR for multiple opportunities for selected response or CR for constructed response). The next columns show the number of items on the test form and average item or task difficulty value in logits, respectively. The following column displays the average percentage of maximum possible score points across items. The last two columns give information on the Rasch model fit statistics (see 5.1.1). The first is the average infit mean square statistic; the second is the average outfit mean square statistic. Optimally, these values should be close to 1.00.

The lower half of Table F provides a summary of the findings of the DIF analyses (see 5.1.4). The first column gives the DIF level: AA, BB, or CC. The next major columns show the contrasting groups in the DIF analyses: either male versus female (M/F) or Hispanic versus other ethnicities (H/O). Even though DIF may be negligible (category AA), this table shows the number of items that were favoring one group or the other at all levels of DIF. Optimally, even when items are all in category AA, there should be roughly an even number of items favoring each of the two groups to ensure that there is no systematic biasing test effect across items.

5.2.9 Complete Item Analysis Table (Table G)

Table G presents results of the analyses of all of the items or tasks on the test form. The first column provides a descriptive name of the item. The item names vary slightly across domains, consisting of characters that represent the domain (e.g., “R” for Reading), the language proficiency level targeted (e.g., “P2”), and the test series (e.g., 502).

The second column in Table G presents the item difficulty in logits, while the third column indicates whether that item served as a common item, anchoring the measurement scale to the results of the field test. The next column shows the percent of maximum possible score points (PMPS). This is obtained by dividing the average score by the maximum possible score point for that task, then multiplying by 100. It is basically a rescaling of the average score. The percentage of maximum possible score points is a common measure used to indicate the task difficulty for a polytomously scored task, with a higher value indicating an easier task. The next two columns show the Rasch fit statistics (see 5.1.1) for the item. The next column provides the point biserial correlation, a measure of the degree to which performance on an item corresponds with performance on the entire test form. In other words, it is a measure of how useful the item is at distinguishing between high-scoring and low-scoring test-takers. The following columns show the results of the two DIF analyses (see 5.2.8) for that item. These last columns are interpreted just as in Table F.

5.2.10 Complete Raw Score to Scale Score Conversion Chart (Table H)

Table H presents the raw score to scale score conversion for the test form. The first column shows all possible raw scores. The next column shows the corresponding scale score for the grade-level cluster.

The next column shows the *conditional* standard error (i.e., from the Rasch analysis) in the metric of the scale score. The last two columns show a lower bound (i.e., the scale score minus one standard error) and an upper bound (i.e., the scale score plus one standard error) around the scale score. In some cases the resulting lower bound or upper bound is below 910, which has been set as the lowest score on the scale.

All domains were adjusted for an end-of-scale effect by allowing the top scale scores to increase only at the same rate as the preceding scale scores. If they were not adjusted, their effect in the composite scores might be excessive.

Thus, if the scale scores towards the high end of the raw score scale were increasing with each raw score by 9 scale points before the group of adjusted scores, then each of the adjusted scores would increase by only 9 scale points each. Because the lower and upper bounds were calculated based on the original logit scores, these adjusted scores do not fall in the middle of the range; they fall toward the lower end of the range, but they always fall *within* the range. In other words, the adjusted scale score is a very possible observed score for that number of raw score points obtained.

In addition, at the lower end of the raw score scale, scale scores are truncated when necessary so that the lowest scale score given is the scale score corresponding to a proficiency level score of A1.

5.2.11 Raw Score to Proficiency Level Score Conversion Table (Table I)

Table I shows the interpretive proficiency level score associated with each raw score. The first column in Table I shows the raw score. The remaining columns show the proficiency level score associated with each raw score/scale score for each grade in the cluster, the percentage of students in that grade who scored at that raw score/scale score/proficiency level score, and the cumulative percentage of students in that grade who scored up to that raw score/scale score/proficiency level score.

There are two things to note about this table. First, unlike scale scores, which are determined psychometrically and have a one-to-one correspondence to raw scores regardless of the grade level of the student, proficiency level scores are interpretations of the scale score. Second, for Alternate ACCESS, cut scores between proficiency levels were determined by domain and do not change by grade level.

In students with severe cognitive disabilities, the cognitive abilities that support language proficiency development are not expected to increase dramatically from one grade level to the next. At this point in the understanding of the development of ELP in such students, it appears appropriate to use the same cut scores for all grade clusters (from grades 1 to 12) by domain. In this way, it becomes easier to detect growth in ELP from year to year for this population of English learners.

5.2.12 Accuracy and Consistency of Classification Table (Table J)

Table J presents three rows of information related to the accuracy and consistency of placement into proficiency categories based on Alternate ACCESS (see above). The first row provides overall indices related to the accuracy and consistency of classification, as well as Cohen's kappa. The second row of information shows accuracy and consistency information conditional on level. The third provides indices of classification accuracy and consistency at the cut points. These indices are perhaps the most important of all when using any of these as an absolute cut-point for placement decisions. Note that the consistency is generally higher at the cut points than over the levels. For practical purposes, the primary score used for such decisions are the Overall Composite scores. In general, the reliability and the accuracy and consistency of classification of the Overall Composite are very high for Alternate ACCESS for ELLs.

5.2.13 Conditional Standard Error of Measurement for Composite Figure (Figure F)

Figure F presents conditional standard error of measurement (CSEM) for composite score. CSEM is measurement errors computed by applying weights of individual domain scale scores in each composite score. The CSEM curves are presented by each proficiency levels in composite scores. This figure informs amount of error variability on scale score level. Higher CSEM informs more measurement error and lower CSEM indicates more reliability.

6. Analyses of Test Forms: Results

6.1 Grades: 1-2

6.1.1 Listening 1-2

Figure 6.1.1A

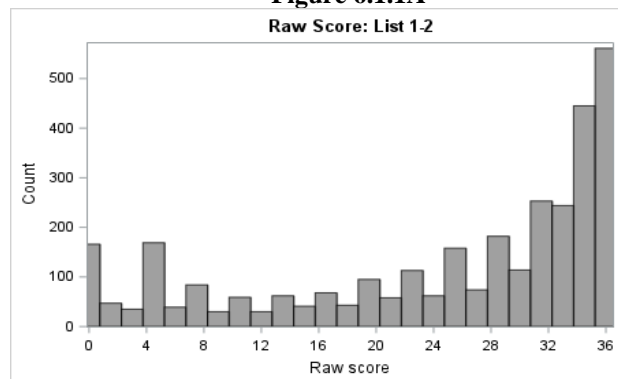


Table 6.1.1A

Raw Score Descriptive Statistics: List 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,567	0	36	23.62	11.96
2	1,665	0	36	25.65	11.55
Total	3,232	0	36	24.66	11.79

Figure 6.1.1B

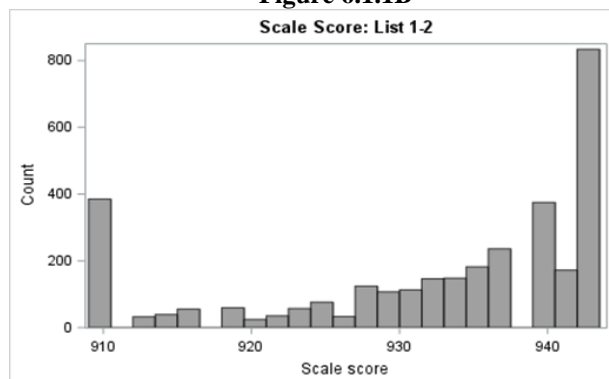


Table 6.1.1B

Scale Score Descriptive Statistics: List 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,567	910	943	931.21	11.24
2	1,665	910	943	933.14	10.84
Total	3,232	910	943	932.20	11.07

Figure 6.1.1C

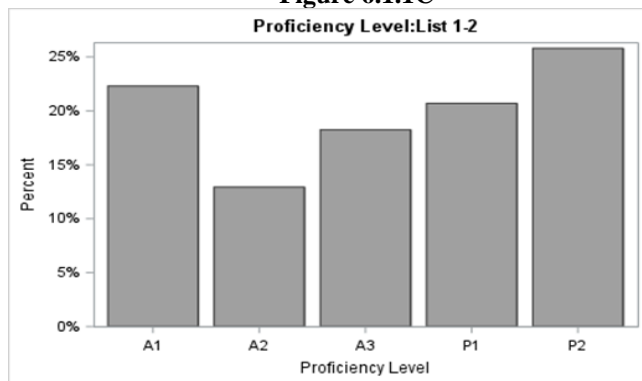


Table 6.1.1C

Proficiency Level Distribution: List 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	386	24.63	335	20.12	721	22.31
A2	223	14.23	195	11.71	418	12.93
A3	300	19.14	290	17.42	590	18.25
P1	329	21.00	340	20.42	669	20.70
P2	329	21.00	505	30.33	834	25.80
Total	1,567	100.00	1,665	100.00	3,232	100.00

Table 6.1.1D

Equating Summary: List 1-2

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 102. Thus, the results from the S102 of the Alternate ACCESS were used to determine raw-to-scale score conversions.

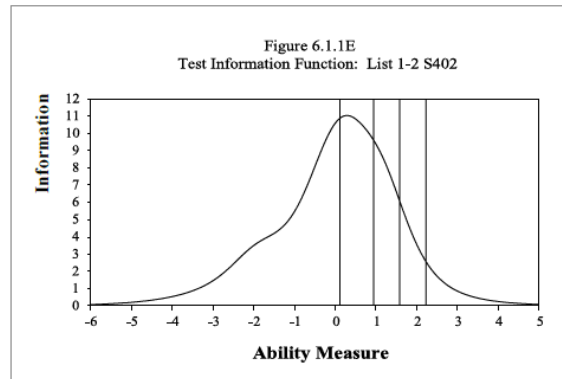
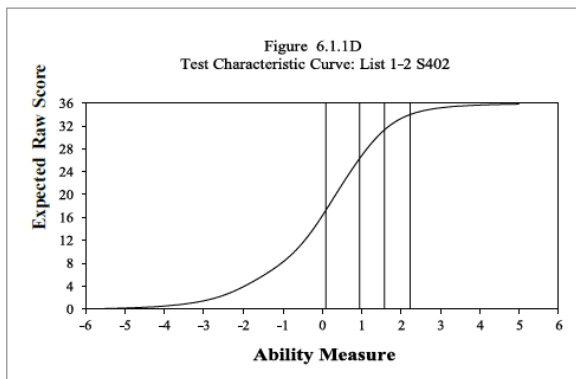


Table 6.1.1E

Reliability: List 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,232	9	0.942	2.885

Table 6.1.1F

Item Analysis Summary: List 1-2

Note: Item information is removed in the redacted version.

Table 6.1.1G

Complete Item Analysis: List 1-2

Note: Item information is removed in the redacted version.

Table 6.1.1H

Raw Score to Scale Score Conversion: List 1-2

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	14.80	910.00^	910.00^
1	910^	8.23	910.00^	910.00^
2	910^	5.70	910.00^	910.00^
3	910^	4.67	910.00^	911.68
4	910^	4.19	910.00^	913.58
5	912	3.96	910.00^	915.48
6	914	3.88	910.00^	917.38
7	915	3.72	911.60	919.04
8	917	3.48	913.50	920.47
9	918	3.24	915.16	921.65
10	920	3.09	916.59	922.76
11	921	2.85	917.93	923.63
12	922	2.77	919.04	924.58
13	923	2.61	920.15	925.37
14	924	2.53	921.02	926.08
15	924	2.45	921.89	926.80
16	925	2.37	922.76	927.51
17	926	2.37	923.47	928.22
18	927	2.37	924.19	928.93
19	927	2.37	924.82	929.57
20	928	2.37	925.53	930.28
21	929	2.37	926.24	930.99
22	929	2.37	926.96	931.70
23	930	2.37	927.67	932.42
24	931	2.37	928.38	933.13
25	931	2.45	929.01	933.92
26	932	2.45	929.80	934.71
27	933	2.53	930.52	935.58
28	934	2.61	931.31	936.53
29	935	2.69	932.10	937.48
30	936	2.85	932.89	938.59
31	937	3.01	933.76	939.77
32	938	3.32	934.71	941.36
33	940	3.80	935.82	943.41
34	942*	4.67	937.16	946.50
35	944*	6.96	938.90	952.83
36	946*	13.85	940.17	967.87

^ Truncated

* Adjusted for end of scale effect

Table 6.1.11

Raw Score to Proficiency Level Conversion: List 1-2

Raw Score	Grade 1			Grade 2		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	5.87	5.87	A1	4.44	4.44
1	A1	0.45	6.32	A1	0.48	4.92
2	A1	1.34	7.66	A1	0.66	5.59
3	A1	1.08	8.74	A1	1.08	6.67
4	A1	5.11	13.85	A1	3.42	10.09
5	A1	0.83	14.68	A1	1.14	11.23
6	A1	1.34	16.02	A1	1.08	12.31
7	A1	1.72	17.74	A1	1.68	13.99
8	A1	0.89	18.63	A1	0.90	14.89
9	A1	0.89	19.53	A1	0.96	15.86
10	A1	0.64	20.17	A1	0.84	16.70
11	A1	1.28	21.44	A1	0.90	17.60
12	A1	1.34	22.78	A1	0.54	18.14
13	A1	0.83	23.61	A1	0.84	18.98
14	A1	1.02	24.63	A1	1.14	20.12
15	A2	1.28	25.91	A2	1.26	21.38
16	A2	1.34	27.25	A2	0.72	22.10
17	A2	0.96	28.21	A2	1.20	23.30
18	A2	1.40	29.61	A2	1.26	24.56
19	A2	1.66	31.27	A2	1.20	25.77
20	A2	2.17	33.44	A2	0.90	26.67
21	A2	2.11	35.55	A2	1.50	28.17
22	A2	1.53	37.08	A2	1.56	29.73
23	A2	1.79	38.86	A2	2.10	31.83
24	A3	1.72	40.59	A3	2.10	33.93
25	A3	2.94	43.52	A3	2.28	36.22
26	A3	2.55	46.08	A3	2.04	38.26
27	A3	2.17	48.25	A3	2.40	40.66
28	A3	3.00	51.24	A3	2.82	43.48
29	A3	3.00	54.24	A3	2.46	45.95
30	A3	3.77	58.01	A3	3.30	49.25
31	P1	3.96	61.97	P1	3.60	52.85
32	P1	4.21	66.18	P1	3.90	56.76
33	P1	7.34	73.52	P1	7.75	64.50
34	P1	5.49	79.00	P1	5.17	69.67
35	P2	7.91	86.92	P2	8.95	78.62
36	P2	13.08	100.00	P2	21.38	100.00

Table 6.1.1J

Accuracy and Consistency of Classification Indices: List 1-2

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.680	0.571		0.448	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.881		0.148	
	A2	0.620		0.274	
	A3	0.594		0.208	
	P1	0.333		0.228	
	P2	0.761		0.704	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.959	0.026	0.015	0.940
	A2/A3	0.931	0.033	0.036	0.909
	A3/P1	0.917	0.016	0.067	0.885
	P1/P2	0.852	0.057	0.091	0.778

6.1.2 Reading 1-2

Figure 6.1.2A

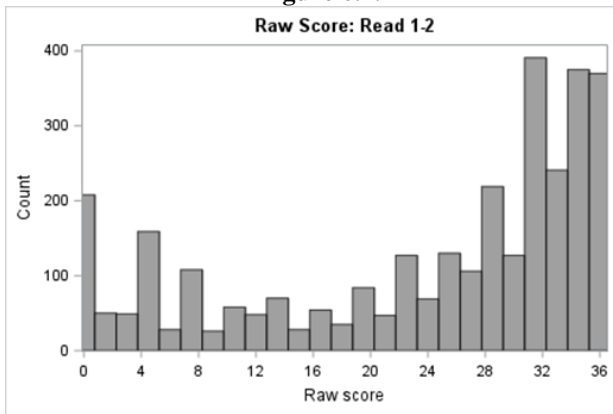


Table 6.1.2A

Raw Score Descriptive Statistics: Read 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,552	0	36	22.69	12.13
2	1,655	0	36	24.71	11.61
Total	3,207	0	36	23.74	11.90

Figure 6.1.2B

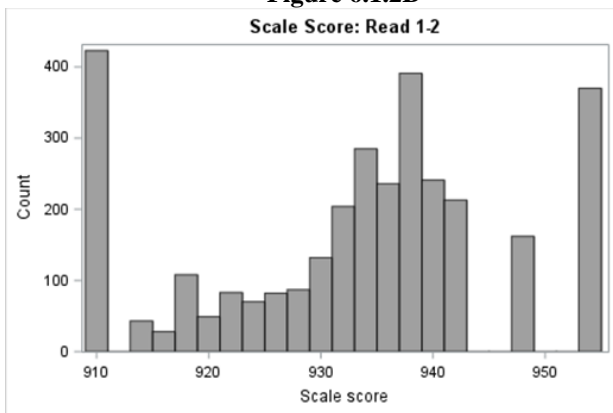


Table 6.1.2B

Scale Score Descriptive Statistics: Read 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,552	910	954	931.53	13.07
2	1,655	910	954	933.98	12.93
Total	3,207	910	954	932.79	13.05

Figure 6.1.2C

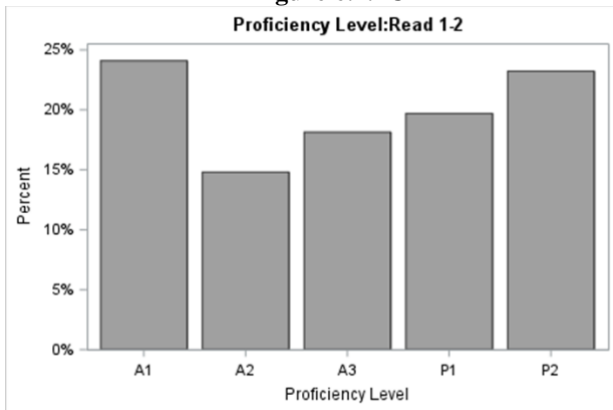


Table 6.1.2C

Proficiency Level Distribution: Read 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	415	26.74	358	21.63	773	24.10
A2	240	15.46	235	14.20	475	14.81
A3	306	19.72	276	16.68	582	18.15
P1	291	18.75	341	20.60	632	19.71
P2	300	19.33	445	26.89	745	23.23
Total	1,552	100.00	1,655	100.00	3,207	100.00

Table 6.1.2D

Equating Summary: Read 1-2

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

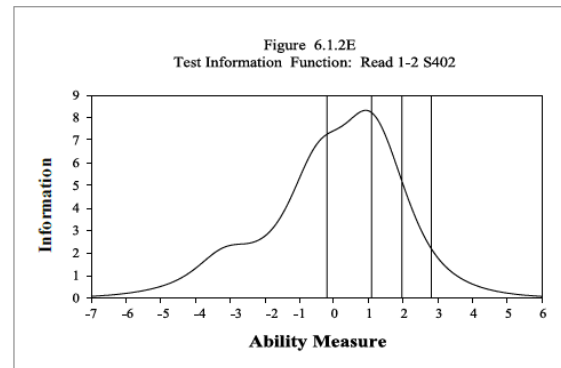
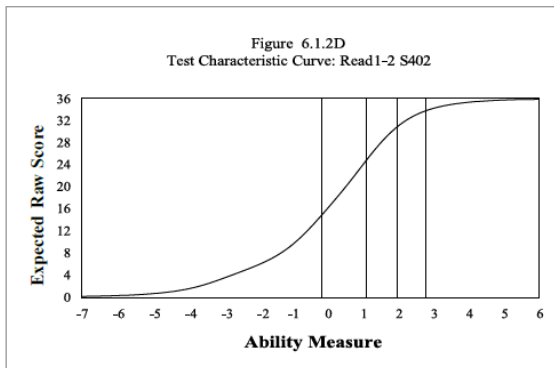


Table 6.1.2E

Reliability: Read 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,207	9	0.947	2.747

Table 6.1.2F

Item Analysis Summary: Read 1-2

Note: Item information is removed in the redacted version.

Table 6.1.2G

Complete Item Analysis: Read 1-2

Note: Item information is removed in the redacted version.

Table 6.1.2H

Raw Score to Scale Score Conversion: Read 1-2

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	11.27	910.00^	910.00^
1	910^	6.09	910.00^	910.00^
2	910^	4.40	910.00^	910.00^
3	910^	3.98	910.00^	911.57
4	910	3.92	910.00^	914.04
5	913	3.80	910.00^	916.39
6	915	3.50	911.33	918.32
7	917	3.13	913.49	919.76
8	918	2.83	915.24	920.91
9	919	2.65	916.69	921.99
10	920	2.47	917.95	922.90
11	921	2.35	919.04	923.74
12	922	2.29	920.00	924.58
13	923	2.23	920.91	925.37
14	924	2.23	921.69	926.15
15	925	2.17	922.59	926.93
16	926	2.17	923.38	927.72
17	926	2.17	924.16	928.50
18	927	2.11	924.94	929.16
19	928	2.11	925.67	929.89
20	929	2.11	926.45	930.67
21	929	2.05	927.17	931.27
22	930	2.05	927.90	931.99
23	931	2.05	928.56	932.66
24	931	2.05	929.28	933.38
25	932	2.05	929.95	934.04
26	933	2.11	930.61	934.83
27	934	2.17	931.33	935.67
28	934	2.23	932.06	936.51
29	935	2.29	932.84	937.42
30	936	2.47	933.62	938.56
31	937	2.65	934.53	939.83
32	938	2.95	935.49	941.40
33	940	3.37	936.70	943.44
34	942*	4.22	938.20	946.64
35	947*	6.03	940.55	952.60
36	954*	11.03	942.84	964.90

^ Truncated

* Adjusted for end of scale effect

Table 6.1.21

Raw Score to Proficiency Level Conversion: Read 1-2

Raw Score	Grade 1			Grade 2		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	7.41	7.41	A1	5.62	5.62
1	A1	0.58	7.99	A1	0.54	6.16
2	A1	1.42	9.41	A1	0.60	6.77
3	A1	1.35	10.76	A1	1.69	8.46
4	A1	4.64	15.40	A1	2.66	11.12
5	A1	1.68	17.07	A1	1.03	12.15
6	A1	0.97	18.04	A1	0.79	12.93
7	A1	2.19	20.23	A1	1.99	14.92
8	A1	1.16	21.39	A1	1.39	16.31
9	A1	0.71	22.10	A1	0.91	17.22
10	A1	0.84	22.94	A1	0.60	17.82
11	A1	0.77	23.71	A1	1.39	19.21
12	A1	1.61	25.32	A1	1.39	20.60
13	A1	1.42	26.74	A1	1.03	21.63
14	A2	1.03	27.77	A2	0.91	22.54
15	A2	0.71	28.48	A2	1.03	23.56
16	A2	0.97	29.45	A2	0.66	24.23
17	A2	1.10	30.54	A2	0.66	24.89
18	A2	1.10	31.64	A2	1.09	25.98
19	A2	1.74	33.38	A2	1.51	27.49
20	A2	0.97	34.34	A2	1.03	28.52
21	A2	1.68	36.02	A2	1.27	29.79
22	A2	1.68	37.69	A2	1.63	31.42
23	A2	2.58	40.27	A2	2.05	33.47
24	A2	1.93	42.20	A2	2.36	35.83
25	A3	1.93	44.14	A3	1.87	37.70
26	A3	1.87	46.01	A3	2.42	40.12
27	A3	3.87	49.87	A3	2.78	42.90
28	A3	3.74	53.61	A3	3.14	46.04
29	A3	4.06	57.67	A3	2.78	48.82
30	A3	4.25	61.92	A3	3.69	52.51
31	P1	4.06	65.98	P1	4.05	56.56
32	P1	8.12	74.10	P1	8.16	64.71
33	P1	6.57	80.67	P1	8.40	73.11
34	P2	5.41	86.08	P2	7.79	80.91
35	P2	4.12	90.21	P2	5.92	86.83
36	P2	9.79	100.00	P2	13.17	100.00

Table 6.1.2J

Accuracy and Consistency of Classification Indices: Read 1-2

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.705	0.619		0.521	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.891		0.130	
	A2	0.649		0.218	
	A3	0.562		0.284	
	P1	0.585		0.271	
	P2	0.749		0.677	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.957	0.024	0.018	0.938
	A2/A3	0.924	0.043	0.032	0.894
	A3/P1	0.903	0.043	0.054	0.870
	P1/P2	0.908	0.024	0.067	0.873

6.1.3 Speaking 1-2

Figure 6.1.3A

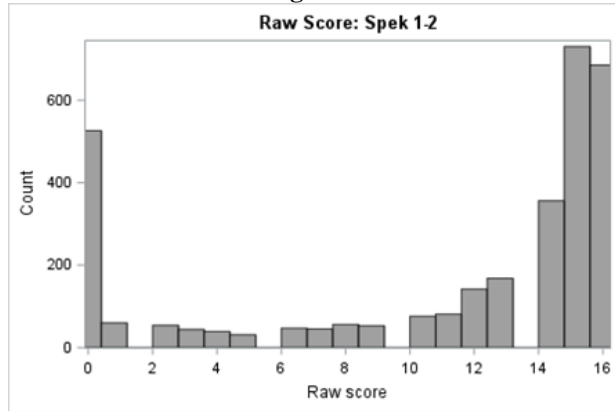


Table 6.1.3A

Raw Score Descriptive Statistics: Spek 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,556	0	16	10.35	6.13
2	1,637	0	16	11.28	5.86
Total	3,193	0	16	10.83	6.01

Figure 6.1.3B

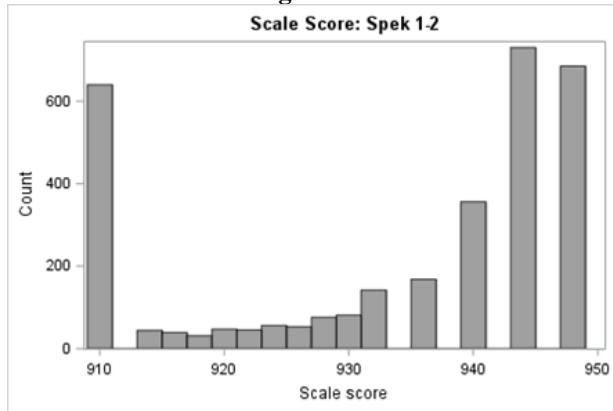


Table 6.1.3B

Scale Score Descriptive Statistics: Spek 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,556	910	948	932.28	14.56
2	1,637	910	948	934.69	14.13
Total	3,193	910	948	933.51	14.39

Figure 6.1.3C

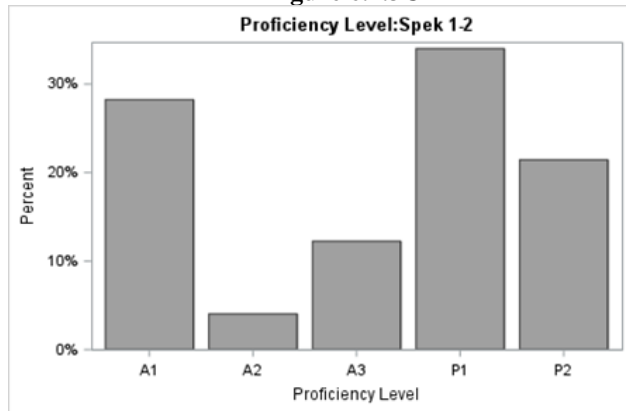


Table 6.1.3C

Proficiency Level Distribution: Spek 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	490	31.49	412	25.17	902	28.25
A2	63	4.05	66	4.03	129	4.04
A3	188	12.08	203	12.40	391	12.25
P1	539	34.64	547	33.41	1,086	34.01
P2	276	17.74	409	24.98	685	21.45
Total	1,556	100.00	1,637	100.00	3,193	100.00

Table 6.1.3D

Equating Summary: Spek 1-2

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

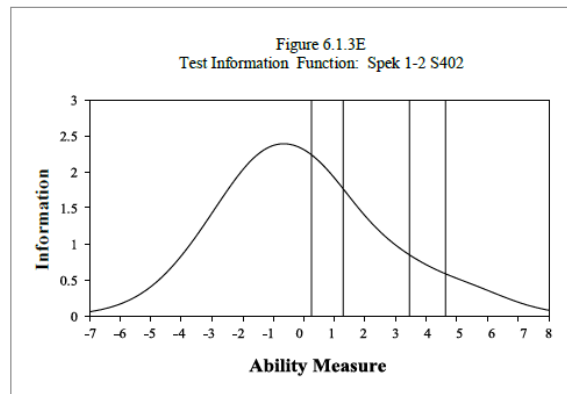
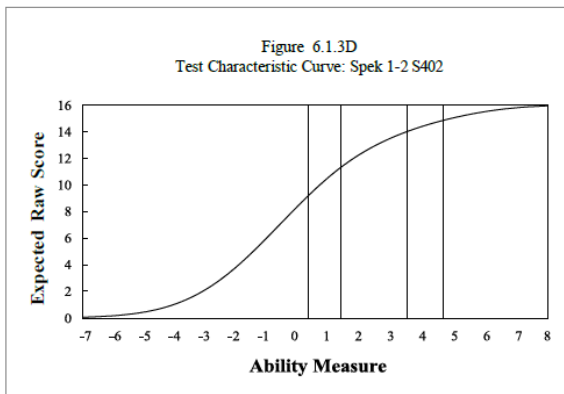


Table 6.1.3E

Reliability: Spek 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,193	8	0.962	1.175

Table 6.1.3F

Item Analysis Summary: Spek 1-2

Note: Item information is removed in the redacted version.

Table 6.1.3G

Complete Item Analysis: Spek 1-2

Note: Item information is removed in the redacted version.

Table 6.1.3H

Raw Score to Scale Score Conversion: Spek 1-2

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	8.42	910.00^	910.00^
1	910^	4.96	910.00^	911.23
2	910	3.86	910.00^	914.34
3	913	3.37	910.00^	916.73
4	916	3.10	912.61	918.81
5	918	2.97	914.78	920.72
6	920	2.88	916.82	922.58
7	922	2.88	918.68	924.44
8	923	2.88	920.54	926.30
9	925	2.97	922.40	928.34
10	927	3.06	924.35	930.47
11	930	3.28	926.39	932.95
12	932	3.59	928.70	935.88
13	936	4.08	931.49	939.65
14	940	4.83	935.17	944.83
15	944*	6.03	940.49	952.55
16	948*	8.95	945.50	963.41

^ Truncated

* Adjusted for end of scale effect

Table 6.1.31

Raw Score to Proficiency Level Conversion: Spek 1-2

Raw Score	Grade 1			Grade 2		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	17.61	17.61	A1	15.39	15.39
1	A1	2.70	20.31	A1	1.10	16.49
2	A1	1.86	22.17	A1	1.53	18.02
3	A1	1.61	23.78	A1	1.16	19.18
4	A1	1.22	25.00	A1	1.22	20.40
5	A1	1.29	26.29	A1	0.67	21.08
6	A1	2.31	28.60	A1	0.67	21.75
7	A1	1.48	30.08	A1	1.34	23.09
8	A1	1.41	31.49	A1	2.08	25.17
9	A2	1.67	33.16	A2	1.65	26.82
10	A2	2.38	35.54	A2	2.38	29.20
11	A3	2.63	38.17	A3	2.44	31.64
12	A3	4.43	42.61	A3	4.46	36.10
13	A3	5.01	47.62	A3	5.50	41.60
14	P1	12.85	60.48	P1	9.53	51.13
15	P1	21.79	82.26	P1	23.89	75.02
16	P2	17.74	100.00	P2	24.98	100.00

Table 6.1.3J

Accuracy and Consistency of Classification Indices: Spek 1-2

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.554	0.571		0.417	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.942		0.250	
	A2	0.505		0.147	
	A3	0.684		0.084	
	P1	0.416		0.405	
	P2	-		0.561	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.979	0.012	0.009	0.970
	A2/A3	0.974	0.012	0.014	0.965
	A3/P1	0.952	0.012	0.036	0.927
	P1/P2	0.646	0.354	0.000	0.681

6.1.4 Writing 1-2

Figure 6.1.4A

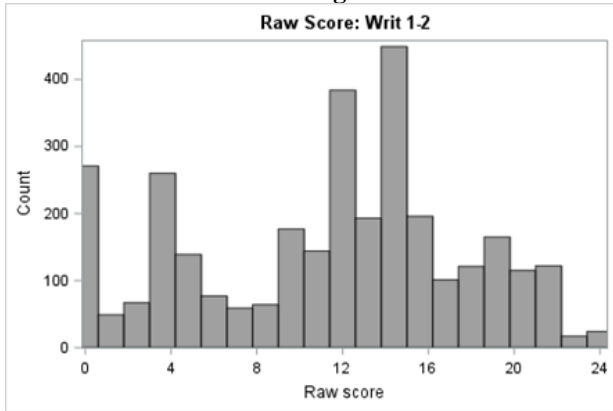


Table 6.1.4A

Raw Score Descriptive Statistics: Writ 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,561	0	24	10.50	6.18
2	1,633	0	24	11.88	6.30
Total	3,194	0	24	11.21	6.28

Figure 6.1.4B

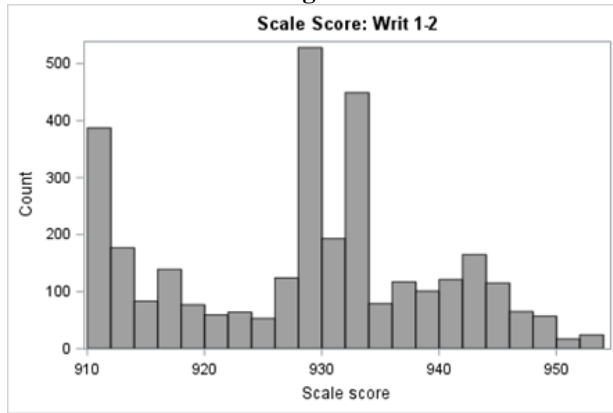


Table 6.1.4B

Scale Score Descriptive Statistics: Writ 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,561	910	953	927.05	11.04
2	1,633	910	953	929.54	11.35
Total	3,194	910	953	928.32	11.27

Figure 6.1.4C

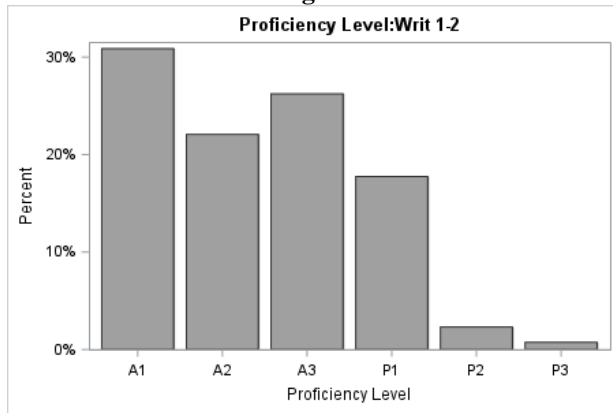


Table 6.1.4C

Proficiency Level Distribution: Writ 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	544	34.85	442	27.07	986	30.87
A2	362	23.19	343	21.00	705	22.07
A3	394	25.24	444	27.19	838	26.24
P1	231	14.80	336	20.58	567	17.75
P2	20	1.28	54	3.31	74	2.32
P3	10	0.64	14	0.86	24	0.75
Total	1,561	100.00	1,633	100.00	3,194	100.00

Table 6.1.4D

Equating Summary: Writ 1-2

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

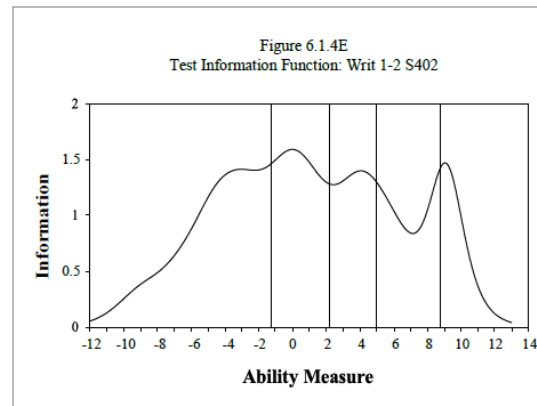
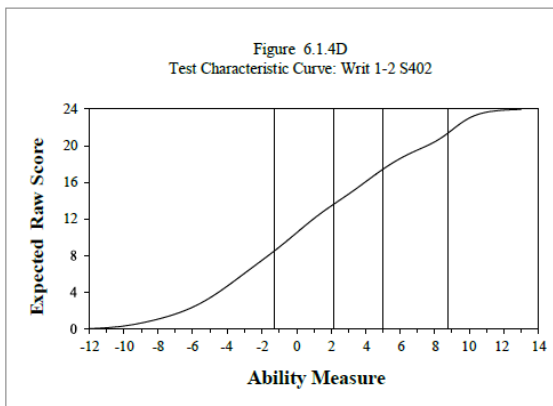


Table 6.1.4E

Reliability: Writ 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,194	10	0.93	1.650

Table 6.1.4F

Item Analysis Summary: Writ 1-2

Note: Item information is removed in the redacted version.

Table 6.1.4G

Complete Item Analysis: Writ 1-2

Note: Item information is removed in the redacted version.

Table 6.1.4H

Raw Score to Scale Score Conversion: Writ 1-2

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	4.99	910.00^	910.00^
1	910^	3.70	910.00^	910.00^
2	910	2.90	910.00^	912.78
3	913	2.40	910.38	915.18
4	915	2.18	912.75	917.12
5	917	2.11	914.72	918.94
6	919	2.11	916.54	920.77
7	921	2.14	918.39	922.66
8	922	2.14	920.31	924.58
9	924	2.04	922.23	926.31
10	926	1.97	923.98	927.92
11	928	1.97	925.59	929.53
12	929	2.04	927.20	931.28
13	931	2.18	928.90	933.27
14	933	2.23	930.92	935.38
15	935	2.14	933.03	937.30
16	937	2.06	934.93	939.06
17	939	2.06	936.68	940.81
18	941	2.14	938.43	942.70
19	943	2.23	940.30	944.77
20	945	2.18	942.42	946.78
21	946	2.02	944.41	948.44
22	948	2.02	946.06	950.10
23	950*	2.50	947.58	952.57
24	952*	4.34	948.63	957.32

^ Truncated

* Adjusted for end of scale effect

Table 6.1.4I

Raw Score to Proficiency Level Conversion: Writ 1-2

Raw Score	Grade 1			Grade 2		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	9.67	9.67	A1	7.35	7.35
1	A1	1.79	11.47	A1	1.29	8.63
2	A1	2.43	13.90	A1	1.78	10.41
3	A1	5.57	19.47	A1	5.51	15.92
4	A1	3.52	23.00	A1	1.71	17.64
5	A1	4.42	27.42	A1	4.29	21.92
6	A1	2.95	30.37	A1	1.90	23.82
7	A1	2.18	32.54	A1	1.53	25.35
8	A1	2.31	34.85	A1	1.71	27.07
9	A2	1.41	36.26	A2	1.90	28.97
10	A2	4.16	40.42	A2	3.61	32.58
11	A2	5.25	45.68	A2	3.80	36.37
12	A2	12.36	58.04	A2	11.70	48.07
13	A3	5.70	63.74	A3	6.37	54.44
14	A3	13.52	77.26	A3	14.57	69.01
15	A3	2.31	79.56	A3	2.63	71.65
16	A3	3.72	83.28	A3	3.61	75.26
17	P1	2.69	85.97	P1	3.61	78.87
18	P1	3.46	89.43	P1	4.10	82.98
19	P1	4.61	94.04	P1	5.70	88.67
20	P1	2.56	96.60	P1	4.59	93.26
21	P1	1.47	98.08	P1	2.57	95.84
22	P2	1.02	99.10	P2	2.51	98.35
23	P2	0.26	99.36	P2	0.80	99.14
24	P3	0.64	100.00	P3	0.86	100.00

Table 6.1.4J

Accuracy and Consistency of Classification Indices: Writ 1-2

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.733	0.650		0.535	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.899		0.147	
	A2	0.708		0.250	
	A3	0.652		0.312	
	P1	0.606		0.648	
	P2	-		0.180	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.943	0.033	0.024	0.919
	A2/A3	0.917	0.036	0.046	0.886
	A3/P1	0.903	0.026	0.070	0.865
	P1/P2	0.966	0.034	0.000	0.960

6.1.5 Oral Language Composite 1-2

Figure 6.1.5A

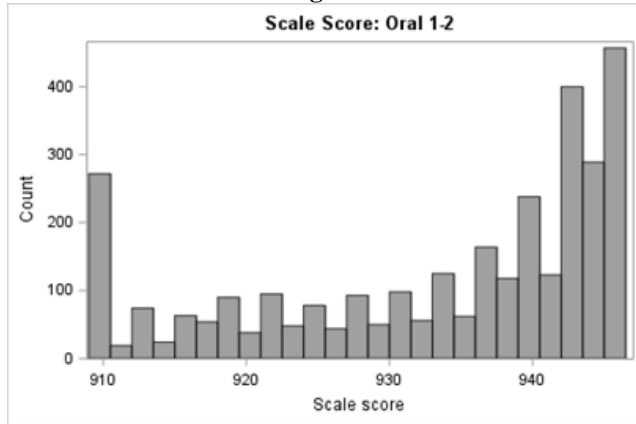


Table 6.1.5A

Scale Score Descriptive Statistics: Oral 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,546	910	946	932.08	12.02
2	1,626	910	946	934.31	11.68
Total	3,172	910	946	933.22	11.90

Figure 6.1.5B

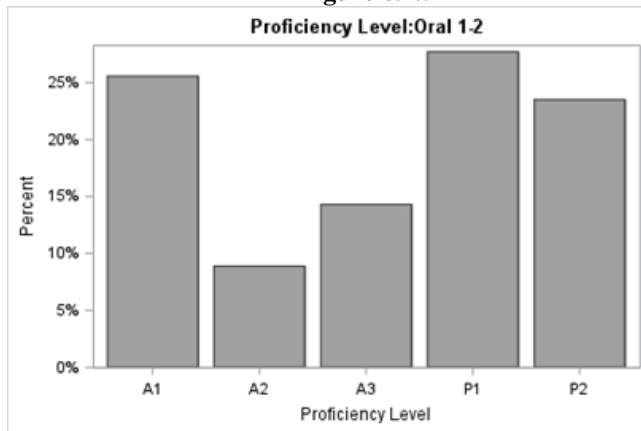


Table 6.1.5C

Proficiency Level Distribution: Oral 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	439	28.40	372	22.88	811	25.57
A2	146	9.44	136	8.36	282	8.89
A3	231	14.94	223	13.71	454	14.31
P1	445	28.78	434	26.69	879	27.71
P2	285	18.43	461	28.35	746	23.52
Total	1,546	100.00	1,626	100.00	3,172	100.00

Table 6.1.5D

n/a

Figure 6.1.5D
n/a

Figure 6.1.5E
n/a

Table 6.1.5E

Reliability: Oral 1-2

Component	Weight	Variance	Reliability
Listening	0.5	122.614	0.940
Speaking	0.5	207.068	0.962
Oral		141.584	0.973

*Variances from students who had results in all four domains

Table 6.1.5F
n/a

Table 6.1.5G
n/a

Table 6.1.5H
n/a

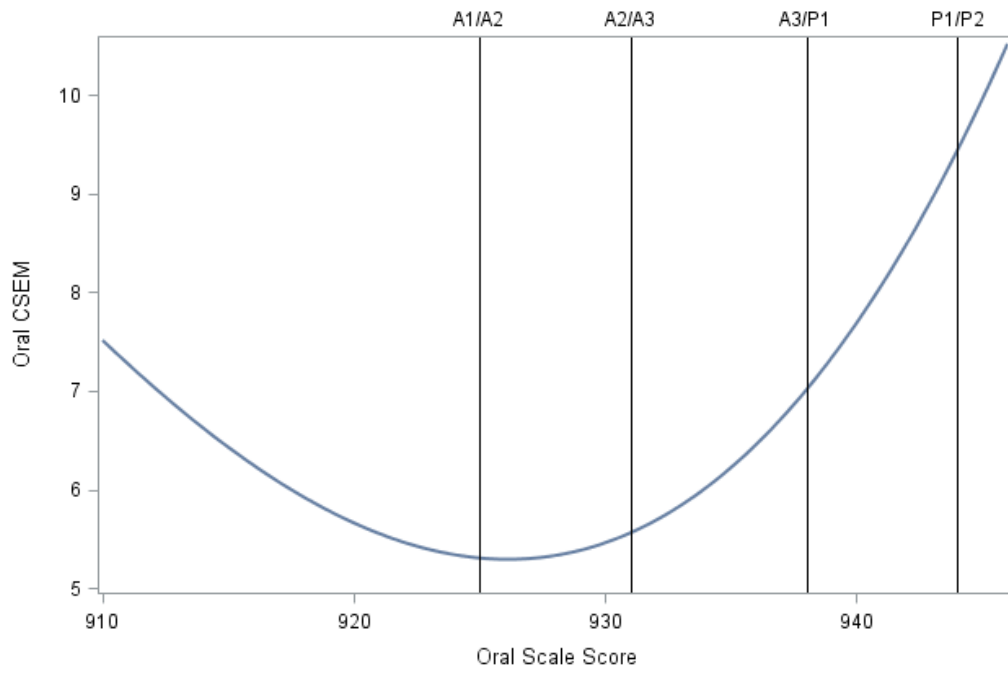
Table 6.1.5I
n/a

Table 6.1.5J

Accuracy and Consistency of Classification Indices: Oral 1-2

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.740	0.656		0.552	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.947		0.921	
	A2	0.639		0.517	
	A3	0.739		0.629	
	P1	0.632		0.507	
	P2	0.679		0.620	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.974	0.014	0.013	0.962
	A2/A3	0.967	0.017	0.016	0.953
	A3/P1	0.958	0.017	0.025	0.941
	P1/P2	0.841	0.067	0.092	0.794

Figure 6.1.5F CSEM for Oral Composite 1-2



6.1.6 Literacy Composite 1-2

Figure 6.1.6A

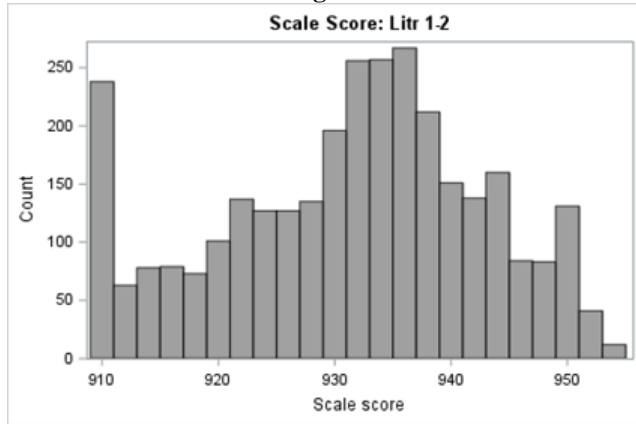


Table 6.1.6A

Scale Score Descriptive Statistics: Litr1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,529	910	954	929.59	11.12
2	1,617	910	954	932.13	11.27
Total	3,146	910	954	930.90	11.26

Figure 6.1.6B

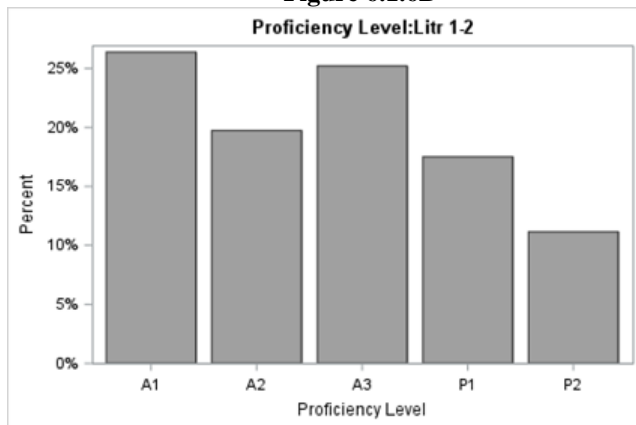


Table 6.1.6C

Proficiency Level Distribution: Litr 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	451	29.50	379	23.44	830	26.38
A2	311	20.34	310	19.17	621	19.74
A3	400	26.16	393	24.30	793	25.21
P1	239	15.63	312	19.29	551	17.51
P2	128	8.37	223	13.79	351	11.16
Total	1,529	100.00	1,617	100.00	3,146	100.00

Table 6.1.6D

n/a

Figure 6.1.6D

n/a

Figure 6.1.6E

n/a

Table 6.1.6E

Reliability: Litr 1-2

Component	Weight	Variance	Reliability
Reading	0.5	170.391	0.947
Writing	0.5	126.938	0.931
Literacy		126.877	0.965

*Variances from students who had results in all four domains

Table 6.1.6F

n/a

Table 6.1.6G

n/a

Table 6.1.6H

n/a

Table 6.1.6I

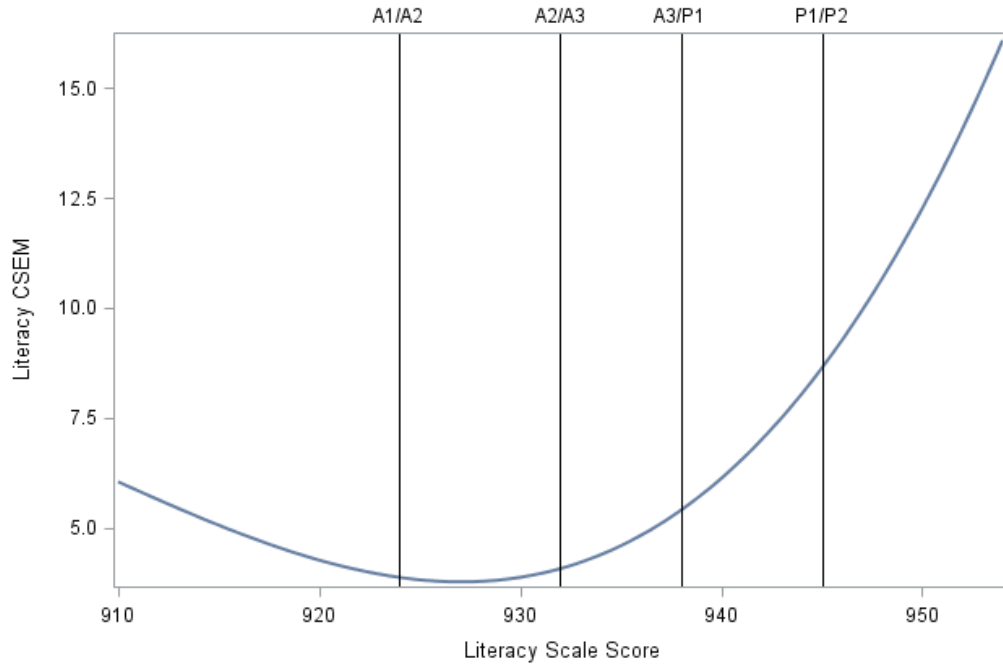
n/a

Table 6.1.6J

Accuracy and Consistency of Classification Indices: Litr 1-2

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.749	0.684		0.596	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.929		0.894	
	A2	0.736		0.635	
	A3	0.795		0.701	
	P1	0.564		0.521	
	P2	0.677		0.540	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.964	0.019	0.017	0.948
	A2/A3	0.939	0.035	0.026	0.914
	A3/P1	0.936	0.018	0.046	0.912
P1/P2	0.910	0.081	0.009	0.906	

Figure 6.1.6F CSEM for Literacy Composite 1-2



6.1.7 Comprehension Composite 1-2

Figure 6.1.7A

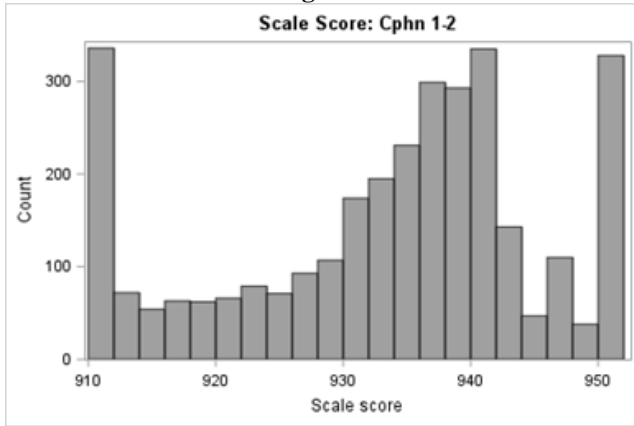


Table 6.1.7A

Scale Score Descriptive Statistics: Cphn 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,548	910	951	931.56	12.15
2	1,648	910	951	933.86	11.93
Total	3,196	910	951	932.74	12.09

Figure 6.1.7B

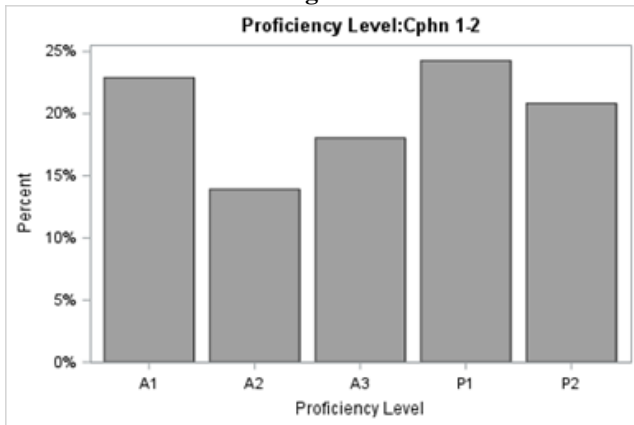


Table 6.1.7C

Proficiency Level Distribution: Cphn 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	389	25.13	343	20.81	732	22.90
A2	231	14.92	214	12.99	445	13.92
A3	303	19.57	274	16.63	577	18.05
P1	366	23.64	410	24.88	776	24.28
P2	259	16.73	407	24.70	666	20.84
Total	1,548	100.00	1,648	100.00	3,196	100.00

Table 6.1.7D

n/a

Figure 6.1.7D

n/a

Figure 6.1.7E

n/a

Table 6.1.7E

Reliability: Cphn 1-2

Component	Weight	Variance	Reliability
Listening	0.3	122.614	0.940
Reading	0.7	170.391	0.947
Comprehension		146.106	0.965

*Variances from students who had results in all four domains

Table 6.1.7F

n/a

Table 6.1.7G

n/a

Table 6.1.7H

n/a

Table 6.1.7I

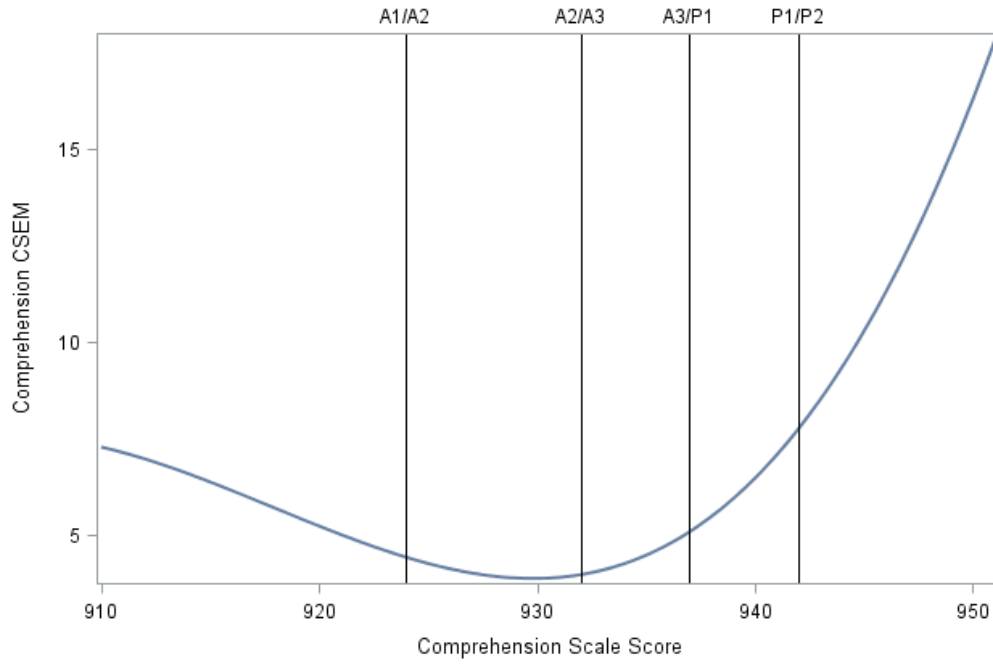
n/a

Table 6.1.7J

Accuracy and Consistency of Classification Indices: Cphn 1-2

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.750	0.666		0.580	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.922		0.883	
	A2	0.728		0.621	
	A3	0.663		0.545	
	P1	0.681		0.552	
	P2	0.730		0.674	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.971	0.016	0.012	0.958
	A2/A3	0.947	0.030	0.023	0.925
	A3/P1	0.927	0.031	0.042	0.900
	P1/P2	0.902	0.022	0.076	0.868

Figure 6.1.7F CSEM for Comprehension Composite 1-2



6.1.8 Overall Composite 1-2

Figure 6.1.8A

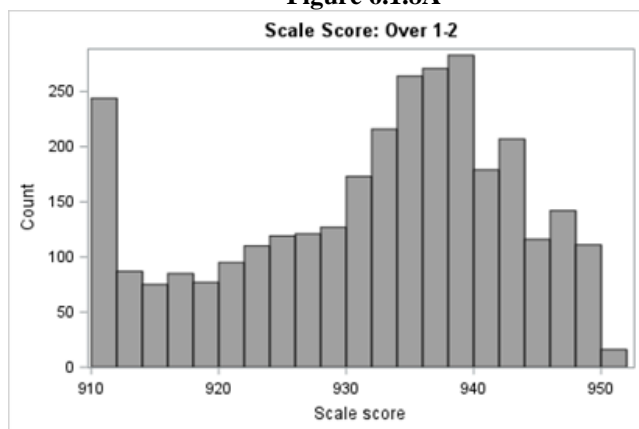


Table 6.1.8A

Scale Score Descriptive Statistics: Over 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	1,522	910	951	930.12	10.96
2	1,596	910	951	932.61	10.91
Total	3,118	910	951	931.40	11.01

Figure 6.1.8B

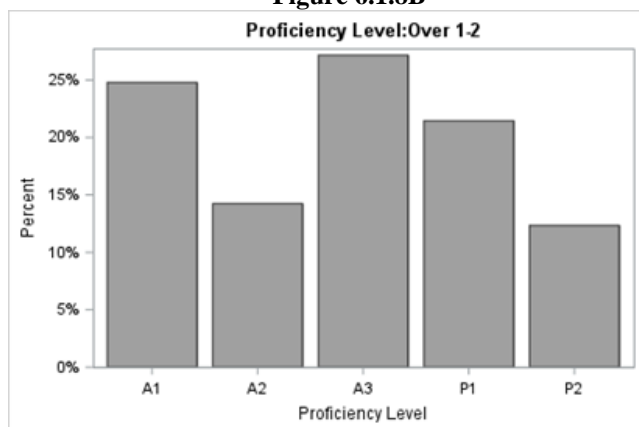


Table 6.1.8C

Proficiency Level Distribution: Over 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	423	27.79	350	21.93	773	24.79
A2	230	15.11	214	13.41	444	14.24
A3	442	29.04	405	25.38	847	27.16
P1	289	18.99	380	23.81	669	21.46
P2	138	9.07	247	15.48	385	12.35
Total	1,522	100.00	1,596	100.00	3,118	100.00

Table 6.1.8D

n/a

Figure 6.1.8D

n/a

Figure 6.1.8E

n/a

Table 6.1.8E

Reliability: Over 1-2

Component	Weight	Variance	Reliability
Listening	0.15	122.614	0.940
Reading	0.35	170.391	0.947
Speaking	0.15	207.068	0.962
Writing	0.35	126.938	0.931
Overall Composite		121.137	0.979

*Variances from students who had results in all four domains

Table 6.1.8F

n/a

Table 6.1.8G

n/a

Table 6.1.8H

n/a

Table 6.1.8I

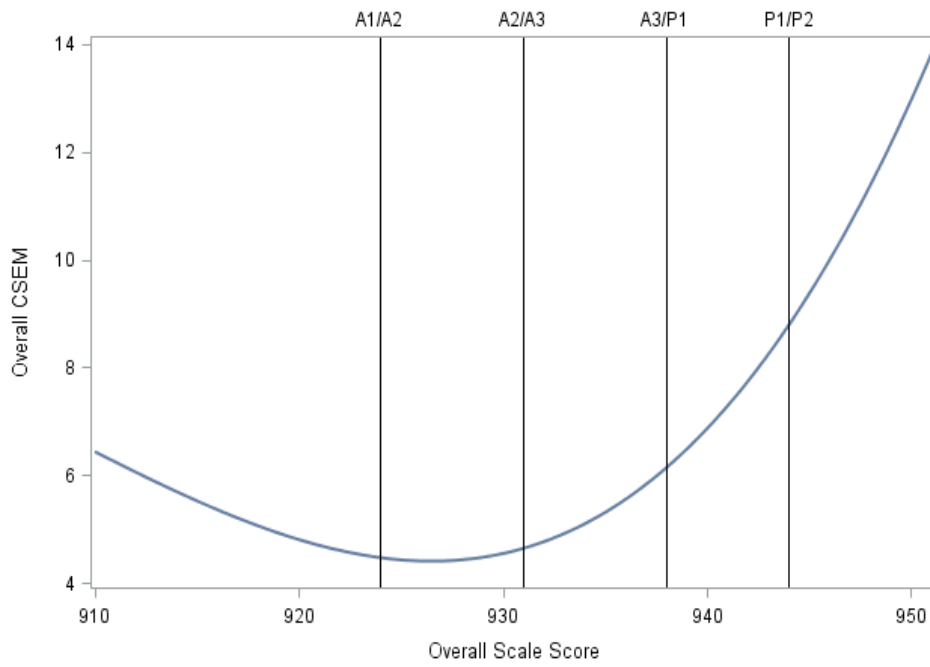
n/a

Table 6.1.8J

Accuracy and Consistency of Classification Indices: Over 1-2

Overall Indices	Accuracy	Consistency		Kappa (k)	
		0.761	0.722		0.644
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.950		0.926	
	A2	0.765		0.670	
	A3	0.875		0.813	
	P1	0.556		0.550	
	P2	-		0.539	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.976	0.013	0.011	0.966
	A2/A3	0.961	0.023	0.017	0.944
	A3/P1	0.952	0.013	0.035	0.933
	P1/P2	0.872	0.128	0.000	0.878

Figure 6.1.8F CSEM for Overall Composite 1-2



6.2 Grades: 3-5

6.2.1 Listening 3-5

Figure 6.2.1A

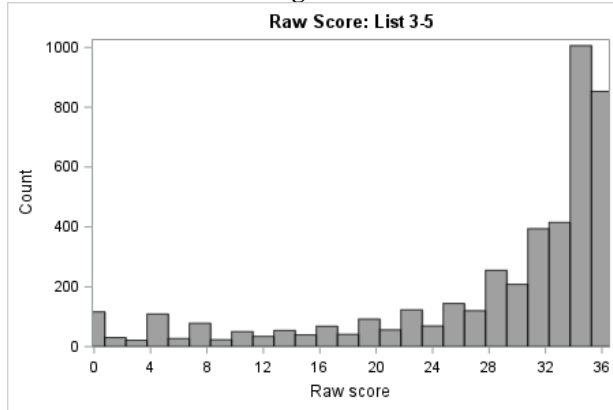


Table 6.2.1A

Raw Score Descriptive Statistics: List 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,539	0	36	26.92	10.25
4	1,510	0	36	28.56	9.80
5	1,376	0	36	29.47	9.12
Total	4,425	0	36	28.27	9.81

Figure 6.2.1B

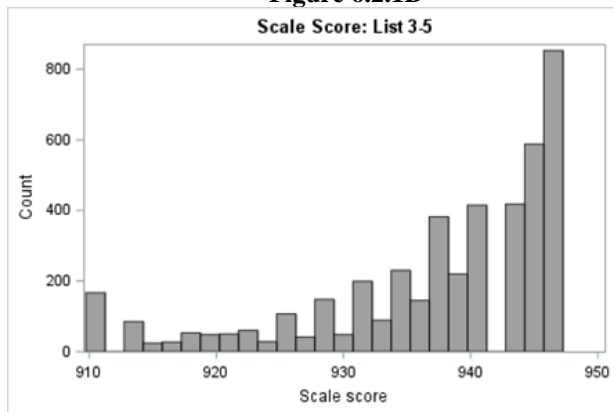


Table 6.2.1B

Scale Score Descriptive Statistics: List 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,539	910	947	935.75	10.28
4	1,510	910	947	937.59	10.02
5	1,376	910	947	938.58	9.46
Total	4,425	910	947	937.26	10.01

Figure 6.2.1C

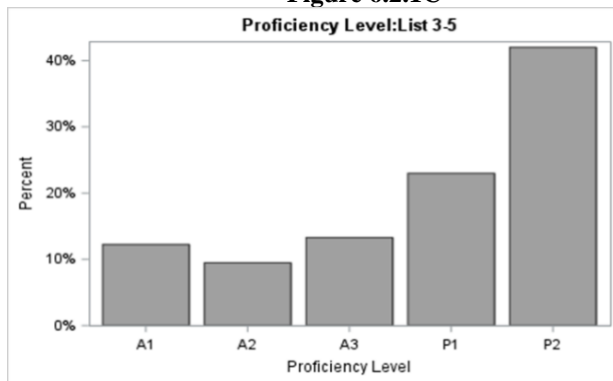


Table 6.2.1C

Proficiency Level Distribution: List 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	230	14.94	181	11.99	131	9.52	542	12.25
A2	178	11.57	124	8.21	117	8.50	419	9.47
A3	249	16.18	187	12.38	152	11.05	588	13.29
P1	352	22.87	349	23.11	316	22.97	1,017	22.98
P2	530	34.44	669	44.30	660	47.97	1,859	42.01
Total	1,539	100.00	1,510	100.00	1,376	100.00	4,425	100.00

Table 6.2.1D

Equating Summary: List 3-5

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 102. Thus, the results from the S102 of the Alternate ACCESS were used to determine raw-to-scale score conversions.

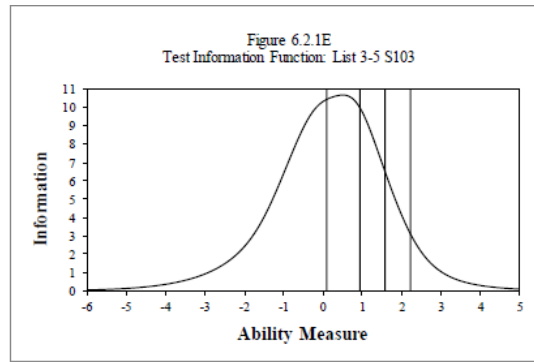
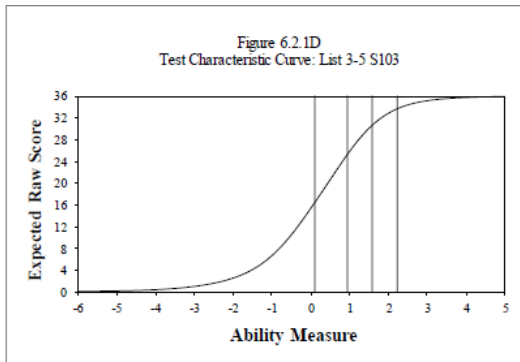


Table 6.2.1E

Reliability: List 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
4,425	9	0.933	2.542

Table 6.2.1F

Item Analysis Summary: List 3-5

Note: Item information is removed in the redacted version.

Table 6.2.1G

Complete Item Analysis: List 3-5

Note: Item information is removed in the redacted version.

Table 6.2.1H

Raw Score to Scale Score Conversion: List 3-5

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	14.64	910.00^	910.00^
1	910^	8.07	910.00^	910.00^
2	910^	5.70	910.00^	910.97
3	910^	4.59	910.00^	913.19
4	911	4.04	910.00^	915.01
5	913	3.64	910.00^	916.43
6	914	3.40	910.97	917.78
7	916	3.24	912.47	918.96
8	917	3.09	913.90	920.07
9	918	2.93	915.16	921.02
10	919	2.77	916.35	921.89
11	920	2.69	917.38	922.76
12	921	2.61	918.41	923.63
13	922	2.53	919.28	924.34
14	923	2.45	920.15	925.06
15	923	2.45	920.94	925.85
16	924	2.45	921.65	926.56
17	925	2.37	922.44	927.19
18	926	2.37	923.24	927.98
19	926	2.37	923.95	928.70
20	927	2.45	924.58	929.49
21	928	2.45	925.37	930.28
22	929	2.45	926.08	930.99
23	929	2.45	926.88	931.78
24	930	2.53	927.59	932.65
25	931	2.53	928.38	933.44
26	932	2.61	929.17	934.39
27	933	2.69	929.96	935.34
28	934	2.77	930.83	936.37
29	935	2.93	931.78	937.64
30	936	3.09	932.73	938.90
31	937	3.32	933.76	940.41
32	939	3.56	935.03	942.15
33	940	4.04	936.37	944.44
34	941*	4.91	937.95	947.77
35	942*	7.04	940.09	954.18
36	943*	13.85	941.44	969.13

^ Truncated

* Adjusted for end of scale effect

Table 6.2.11

Raw Score to Proficiency Level Conversion: List 3-5

Raw Score	Grade 3			Grade 4			Grade 5		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	3.31	3.31	A1	2.52	2.52	A1	1.96	1.96
1	A1	0.06	3.38	A1	0.20	2.72	A1	0.22	2.18
2	A1	0.58	3.96	A1	0.66	3.38	A1	0.29	2.47
3	A1	0.45	4.42	A1	0.46	3.84	A1	0.51	2.98
4	A1	2.47	6.89	A1	1.99	5.83	A1	1.24	4.22
5	A1	0.58	7.47	A1	0.46	6.29	A1	0.58	4.80
6	A1	0.71	8.19	A1	0.40	6.69	A1	0.73	5.52
7	A1	1.30	9.49	A1	1.39	8.08	A1	0.87	6.40
8	A1	0.71	10.20	A1	0.53	8.61	A1	0.44	6.83
9	A1	0.71	10.92	A1	0.40	9.01	A1	0.44	7.27
10	A1	0.52	11.44	A1	0.73	9.74	A1	0.36	7.63
11	A1	0.84	12.28	A1	0.53	10.26	A1	0.36	7.99
12	A1	1.10	13.39	A1	0.60	10.86	A1	0.58	8.58
13	A1	0.84	14.23	A1	0.46	11.32	A1	0.44	9.01
14	A1	0.71	14.94	A1	0.66	11.99	A1	0.51	9.52
15	A2	0.84	15.79	A2	0.93	12.91	A2	0.87	10.39
16	A2	0.84	16.63	A2	0.93	13.84	A2	0.65	11.05
17	A2	0.78	17.41	A2	0.66	14.50	A2	0.73	11.77
18	A2	1.30	18.71	A2	0.73	15.23	A2	0.73	12.50
19	A2	1.17	19.88	A2	0.86	16.09	A2	0.87	13.37
20	A2	1.49	21.38	A2	1.13	17.22	A2	0.65	14.03
21	A2	1.36	22.74	A2	1.06	18.28	A2	1.38	15.41
22	A2	1.62	24.37	A2	0.73	19.01	A2	0.87	16.28
23	A2	2.14	26.51	A2	1.19	20.20	A2	1.74	18.02
24	A3	2.34	28.85	A3	1.26	21.46	A3	1.02	19.04
25	A3	1.36	30.21	A3	1.46	22.91	A3	0.87	19.91
26	A3	2.34	32.55	A3	1.99	24.90	A3	1.67	21.58
27	A3	2.86	35.41	A3	2.58	27.48	A3	2.69	24.27
28	A3	3.57	38.99	A3	1.99	29.47	A3	1.82	26.09
29	A3	3.70	42.69	A3	3.11	32.58	A3	2.98	29.07
30	P1	4.94	47.63	P1	4.90	37.48	P1	4.22	33.28
31	P1	3.57	51.20	P1	4.04	41.52	P1	4.22	37.50
32	P1	4.74	55.95	P1	5.10	46.62	P1	5.09	42.59
33	P1	9.62	65.56	P1	9.07	55.70	P1	9.45	52.03
34	P2	9.68	75.24	P2	9.60	65.30	P2	9.01	61.05
35	P2	10.46	85.71	P2	14.64	79.93	P2	14.97	76.02
36	P2	14.29	100.00	P2	20.07	100.00	P2	23.98	100.00

Table 6.2.1J

Accuracy and Consistency of Classification Indices: List 3-5

Overall Indices	Accuracy	Consistency		Kappa (k)	
		0.483	0.506		0.338
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.899		0.158	
	A2	0.625		0.222	
	A3	0.633		0.100	
	P1	0.369		0.361	
	P2	-		0.598	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.976	0.014	0.010	0.966
	A2/A3	0.961	0.019	0.020	0.947
	A3/P1	0.937	0.011	0.052	0.910
	P1/P2	0.605	0.395	0.000	0.652

6.2.2 Reading 3-5

Figure 6.2.2A

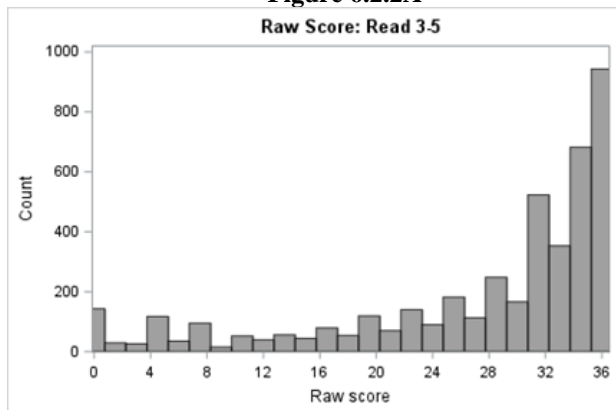


Table 6.2.2A

Raw Score Descriptive Statistics: Read 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,532	0	36	25.75	10.67
4	1,510	0	36	27.80	10.16
5	1,371	0	36	28.81	9.54
Total	4,413	0	36	27.40	10.23

Figure 6.2.2B

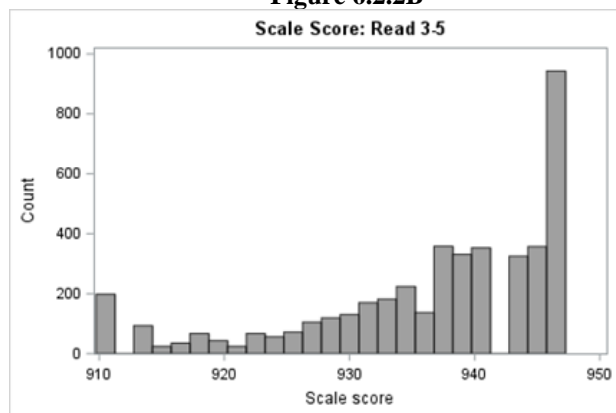


Table 6.2.2B

Scale Score Descriptive Statistics: Read 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,532	910	947	934.59	10.51
4	1,510	910	947	936.84	10.35
5	1,371	910	947	937.96	9.82
Total	4,413	910	947	936.41	10.34

Figure 6.2.2C

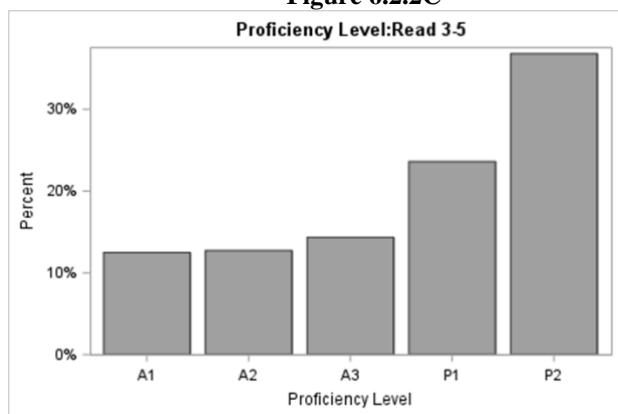


Table 6.2.2C

Proficiency Level Distribution: Read 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	231	15.08	182	12.05	138	10.07	551	12.49
A2	250	16.32	170	11.26	142	10.36	562	12.74
A3	245	15.99	216	14.30	172	12.55	633	14.34
P1	384	25.07	354	23.44	304	22.17	1,042	23.61
P2	422	27.55	588	38.94	615	44.86	1,625	36.82
Total	1,532	100.00	1,510	100.00	1,371	100.00	4,413	100.00

Table 6.2.2D

Equating Summary: Read 3-5

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 102. Thus, the results from the S102 of the Alternate ACCESS were used to determine raw-to-scale score conversions.

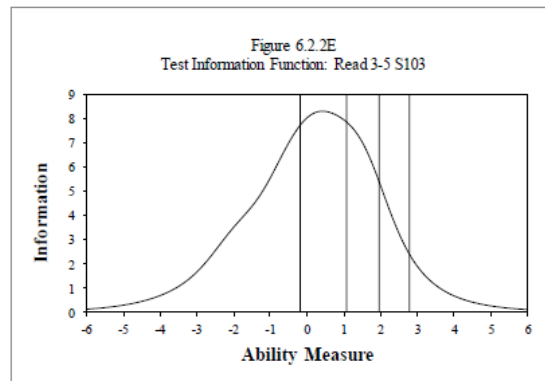
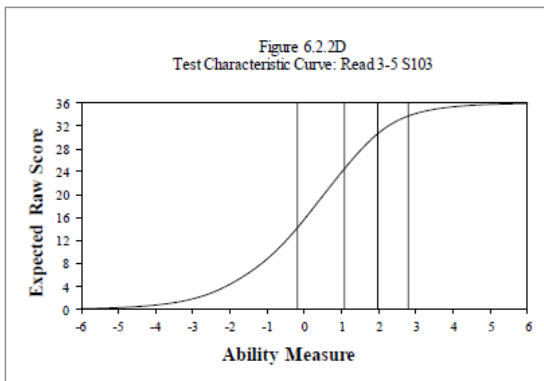


Table 6.2.2E

Reliability: Read 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
4,413	9	0.943	2.451

Table 6.2.2F

Item Analysis Summary: Read 3-5

Note: Item information is removed in the redacted version.

Table 6.2.2G

Complete Item Analysis: Read 3-5

Note: Item information is removed in the redacted version.

Table 6.2.2H

Raw Score to Scale Score Conversion: Read 3-5

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	11.45	910.00^	910.00^
1	910^	6.33	910.00^	910.00^
2	910^	4.52	910.00^	910.00^
3	910^	3.92	910.00^	911.57
4	910	3.74	910.00^	913.86
5	912	3.68	910.00^	916.09
6	915	3.50	911.02	918.01
7	916	3.19	913.19	919.58
8	918	2.95	915.00	920.91
9	919	2.71	916.57	921.99
10	920	2.53	917.89	922.96
11	921	2.41	919.04	923.86
12	922	2.35	920.06	924.76
13	923	2.23	921.03	925.49
14	924	2.17	921.93	926.27
15	925	2.17	922.71	927.05
16	926	2.11	923.50	927.72
17	926	2.11	924.22	928.44
18	927	2.05	925.00	929.10
19	928	2.05	925.73	929.83
20	928	2.05	926.39	930.49
21	929	2.05	927.11	931.21
22	930	2.05	927.78	931.87
23	931	2.05	928.50	932.60
24	931	2.11	929.16	933.38
25	932	2.11	929.89	934.10
26	933	2.17	930.55	934.89
27	934	2.17	931.33	935.67
28	934	2.29	932.06	936.63
29	935	2.35	932.90	937.60
30	936	2.47	933.74	938.68
31	937	2.65	934.65	939.95
32	939	2.95	935.67	941.58
33	940	3.37	936.88	943.62
34	941*	4.16	938.44	946.76
35	942*	6.03	940.67	952.72
36	943*	11.03	942.90	964.96

^ Truncated

* Adjusted for end of scale effect

Table 6.2.2I

Raw Score to Proficiency Level Conversion: Read 3-5

Raw Score	Grade 3			Grade 4			Grade 5		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	4.37	4.37	A1	2.91	2.91	A1	2.33	2.33
1	A1	0.20	4.57	A1	0.20	3.11	A1	0.22	2.55
2	A1	0.65	5.22	A1	0.66	3.77	A1	0	2.55
3	A1	0.59	5.81	A1	0.60	4.37	A1	0.58	3.14
4	A1	2.02	7.83	A1	2.38	6.75	A1	1.90	5.03
5	A1	0.85	8.68	A1	0.26	7.02	A1	0.51	5.54
6	A1	0.85	9.53	A1	0.73	7.75	A1	0.80	6.35
7	A1	1.70	11.23	A1	1.59	9.34	A1	1.24	7.59
8	A1	0.78	12.01	A1	0.79	10.13	A1	0.29	7.88
9	A1	0.46	12.47	A1	0.26	10.40	A1	0.29	8.17
10	A1	0.85	13.32	A1	0.33	10.73	A1	0.44	8.61
11	A1	0.98	14.30	A1	0.73	11.46	A1	0.15	8.75
12	A1	0.78	15.08	A1	0.60	12.05	A1	1.31	10.07
13	A2	0.85	15.93	A2	0.60	12.65	A2	0.22	10.28
14	A2	0.91	16.84	A2	0.46	13.11	A2	0.73	11.01
15	A2	1.24	18.08	A2	0.99	14.11	A2	0.73	11.74
16	A2	0.98	19.06	A2	0.46	14.57	A2	0.44	12.18
17	A2	1.70	20.76	A2	0.99	15.56	A2	0.73	12.91
18	A2	1.70	22.45	A2	0.99	16.56	A2	0.95	13.86
19	A2	1.76	24.22	A2	1.13	17.68	A2	1.31	15.17
20	A2	1.70	25.91	A2	1.39	19.07	A2	0.73	15.90
21	A2	1.76	27.68	A2	1.59	20.66	A2	1.39	17.29
22	A2	1.70	29.37	A2	1.13	21.79	A2	1.24	18.53
23	A2	2.02	31.40	A2	1.52	23.31	A2	1.90	20.42
24	A3	2.55	33.94	A3	1.79	25.10	A3	1.75	22.17
25	A3	2.28	36.23	A3	1.85	26.95	A3	2.19	24.36
26	A3	2.22	38.45	A3	2.45	29.40	A3	1.31	25.67
27	A3	2.22	40.67	A3	2.85	32.25	A3	2.63	28.30
28	A3	3.46	44.13	A3	1.99	34.24	A3	2.04	30.34
29	A3	3.26	47.39	A3	3.38	37.62	A3	2.63	32.97
30	P1	3.59	50.98	P1	3.91	41.52	P1	3.79	36.76
31	P1	4.37	55.35	P1	4.64	46.16	P1	4.01	40.77
32	P1	8.36	63.71	P1	7.22	53.38	P1	6.86	47.63
33	P1	8.75	72.45	P1	7.68	61.06	P1	7.51	55.14
34	P2	6.79	79.24	P2	7.22	68.28	P2	8.17	63.31
35	P2	6.92	86.16	P2	7.28	75.56	P2	10.28	73.60
36	P2	13.84	100.00	P2	24.44	100.00	P2	26.40	100.00

Table 6.2.2J

Accuracy and Consistency of Classification Indices: Read 3-5

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.527	0.528		0.389	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.896		0.125	
	A2	0.706		0.253	
	A3	0.653		0.110	
	P1	0.385		0.373	
	P2	-		0.591	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.973	0.016	0.011	0.961
	A2/A3	0.954	0.022	0.024	0.937
	A3/P1	0.939	0.012	0.049	0.912
	P1/P2	0.659	0.341	0.000	0.693

6.2.3 Speaking 3-5

Figure 6.2.3A

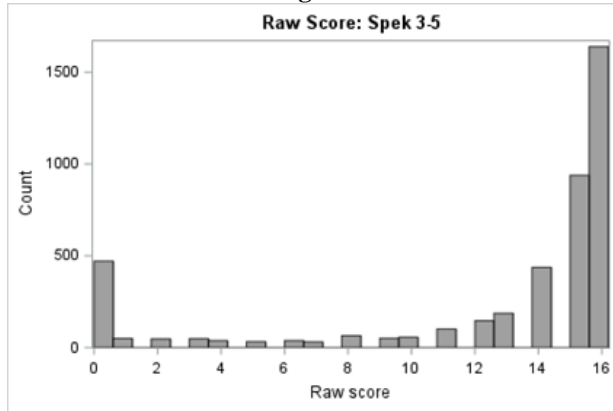


Table 6.2.3A

Raw Score Descriptive Statistics: Spek 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,519	0	16	11.83	5.62
4	1,494	0	16	12.63	5.25
5	1,364	0	16	12.82	5.16
Total	4,377	0	16	12.41	5.37

Figure 6.2.3B

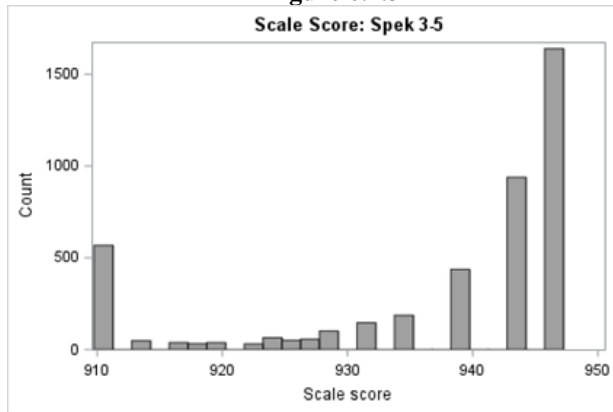


Table 6.2.3B

Scale Score Descriptive Statistics: Spek 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,519	910	947	935.45	13.25
4	1,494	910	947	937.61	12.60
5	1,364	910	947	938.12	12.36
Total	4,377	910	947	937.02	12.81

Figure 6.2.3C

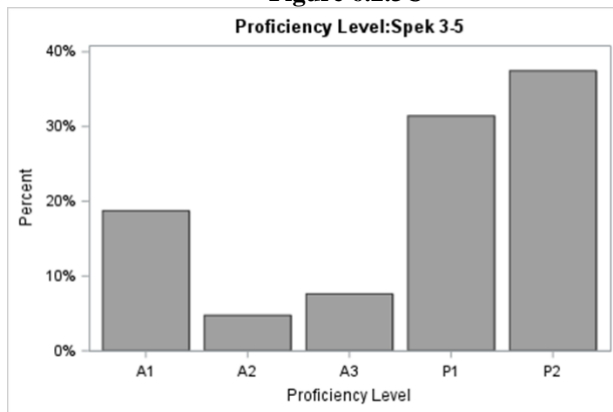


Table 6.2.3C

Proficiency Level Distribution: Spek 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	332	21.86	271	18.14	217	15.91	820	18.73
A2	87	5.73	60	4.02	61	4.47	208	4.75
A3	133	8.76	104	6.96	96	7.04	333	7.61
P1	511	33.64	452	30.25	413	30.28	1,376	31.44
P2	456	30.02	607	40.63	577	42.30	1,640	37.47
Total	1,519	100.00	1,494	100.00	1,364	100.00	4,377	100.00

Table 6.2.3D

Equating Summary: Spek 3-5

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

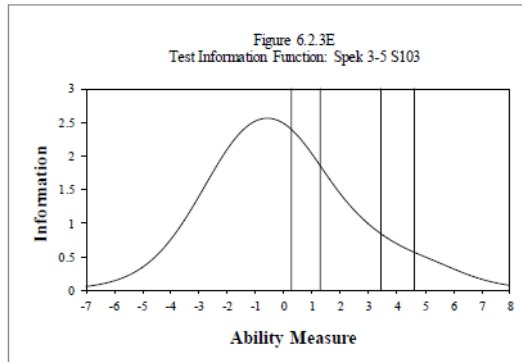
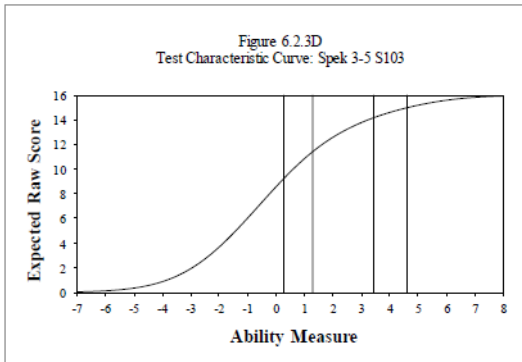


Table 6.2.3E

Reliability: Spek 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
4,377	8	0.965	1.003

Table 6.2.3F

Item Analysis Summary: Spek 3-5

Note: Item information is removed in the redacted version.

Table 6.2.3G

Complete Item Analysis: Spek 3-5

Note: Item information is removed in the redacted version.

Table 6.2.3H

Raw Score to Scale Score Conversion: Spek 3-5

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	8.33	910.00^	910.00^
1	910^	4.88	910.00^	912.34
2	911	3.72	910.00^	915.13
3	914	3.24	910.88	917.35
4	916	3.01	913.27	919.30
5	918	2.88	915.35	921.12
6	920	2.79	917.22	922.80
7	922	2.75	919.03	924.53
8	924	2.79	920.72	926.30
9	925	2.84	922.45	928.12
10	927	2.97	924.22	930.16
11	929	3.15	926.13	932.42
12	932	3.46	928.25	935.17
13	935	3.95	930.83	938.72
14	939	4.70	934.19	943.59
15	943*	5.94	939.20	951.08
16	947*	8.95	943.99	961.90

^ Truncated

* Adjusted for end of scale effect

Table 6.2.3I

Raw Score to Proficiency Level Conversion: Spek 3-5

Raw Score	Grade 3			Grade 4			Grade 5		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	12.64	12.64	A1	9.77	9.77	A1	9.75	9.75
1	A1	1.38	14.02	A1	0.94	10.71	A1	1.10	10.85
2	A1	1.38	15.40	A1	1.20	11.91	A1	0.59	11.44
3	A1	1.38	16.79	A1	0.80	12.72	A1	1.10	12.54
4	A1	0.66	17.45	A1	1.27	13.99	A1	0.66	13.20
5	A1	0.92	18.37	A1	0.60	14.59	A1	0.66	13.86
6	A1	0.92	19.29	A1	1.20	15.80	A1	0.44	14.30
7	A1	1.18	20.47	A1	0.54	16.33	A1	0.37	14.66
8	A1	1.38	21.86	A1	1.81	18.14	A1	1.25	15.91
9	A2	1.38	23.24	A2	0.74	18.88	A2	1.39	17.30
10	A2	1.51	24.75	A2	1.07	19.95	A2	1.25	18.55
11	A2	2.83	27.58	A2	2.21	22.16	A2	1.83	20.38
12	A3	3.75	31.34	A3	3.21	25.37	A3	3.01	23.39
13	A3	5.00	36.34	A3	3.75	29.12	A3	4.03	27.42
14	P1	11.32	47.66	P1	9.30	38.42	P1	9.24	36.66
15	P1	22.32	69.98	P1	20.95	59.37	P1	21.04	57.70
16	P2	30.02	100.00	P2	40.63	100.00	P2	42.30	100.00

Table 6.2.3J

Accuracy and Consistency of Classification Indices: Spek 3-5

Overall Indices	Accuracy	Consistency		Kappa (k)	
		0.563	0.584		0.435
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.946		0.236	
	A2	0.522		0.148	
	A3	0.721		0.078	
	P1	0.421		0.406	
	P2	-		0.572	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.980	0.012	0.008	0.972
	A2/A3	0.975	0.012	0.013	0.966
	A3/P1	0.959	0.011	0.030	0.939
	P1/P2	0.646	0.354	0.000	0.687

6.2.4 Writing 3-5

Figure 6.2.4A

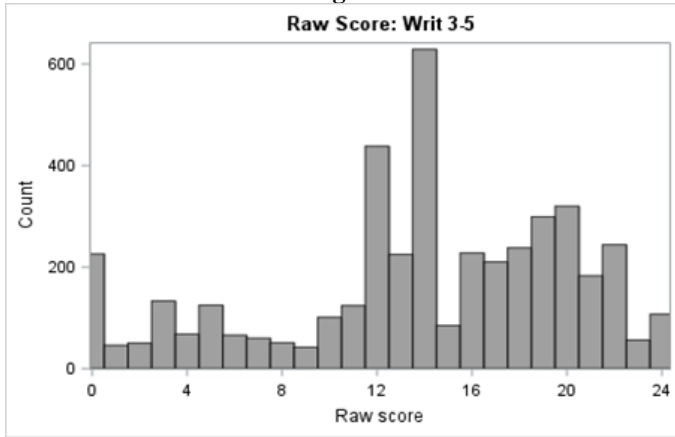


Table 6.2.4A

Raw Score Descriptive Statistics: Writ 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,513	0	24	12.96	6.16
4	1,493	0	24	14.17	6.32
5	1,348	0	24	14.52	6.30
Total	4,354	0	24	13.86	6.29

Figure 6.2.4B

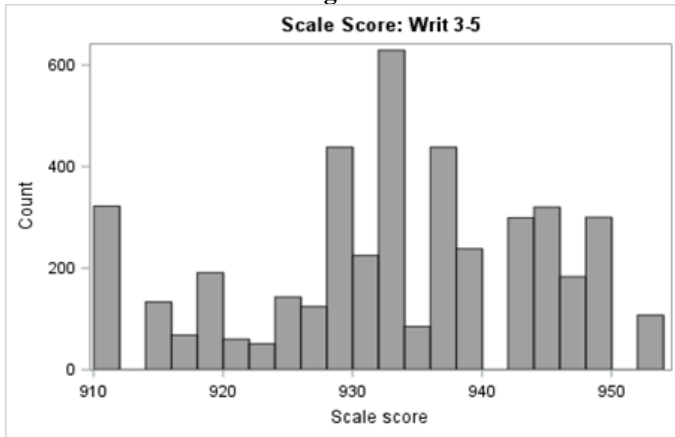


Table 6.2.4B

Scale Score Descriptive Statistics: Writ 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,513	910	953	930.84	10.76
4	1,493	910	953	933.14	11.29
5	1,348	910	953	933.75	11.28
Total	4,354	910	953	932.53	11.17

Figure 6.2.4C

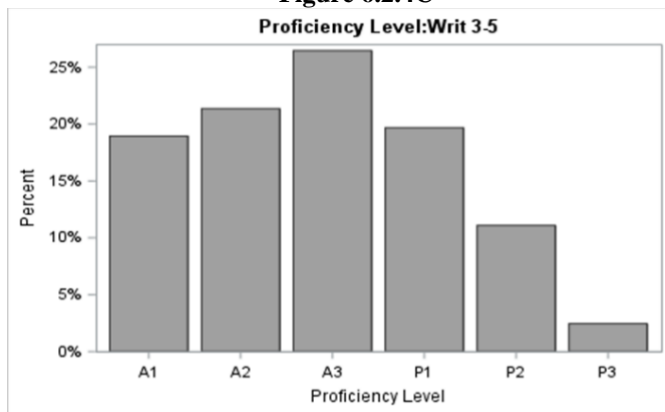


Table 6.2.4C

Proficiency Level Distribution: Writ 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	329	21.74	273	18.29	223	16.54	825	18.95
A2	376	24.85	301	20.16	253	18.77	930	21.36
A3	415	27.43	378	25.32	359	26.63	1,152	26.46
P1	254	16.79	313	20.96	290	21.51	857	19.68
P2	120	7.93	188	12.59	175	12.98	483	11.09
P3	19	1.26	40	2.68	48	3.56	107	2.46
Total	1,513	100.00	1,493	100.00	1,348	100.00	4,354	100.00

Table 6.2.4D

Equating Summary: Writ 3-5

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

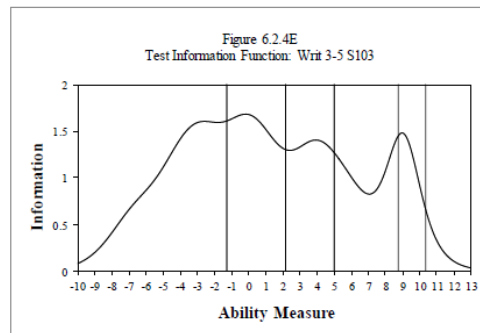
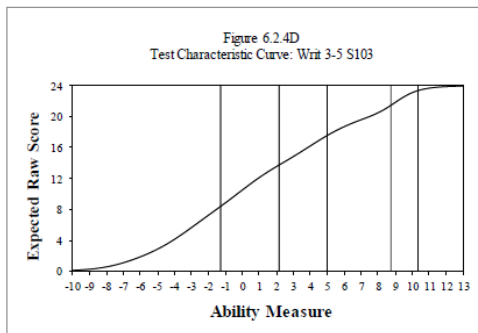


Table 6.2.4E

Reliability: Writ 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
4,354	10	0.933	1.628

Table 6.2.4F

Item Analysis Summary: Writ 3-5

Note: Item information is removed in the redacted version.

Table 6.2.4G

Complete Item Analysis: Writ 3-5

Note: Item information is removed in the redacted version.

Table 6.2.4H

Raw Score to Scale Score Conversion: Writ 3-5

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	4.92	910.00^	910.00^
1	910^	3.43	910.00^	911.00
2	911	2.66	910.00^	914.05
3	914	2.26	911.60	916.11
4	916	2.06	913.71	917.84
5	918	2.02	915.49	919.52
6	919	1.99	917.17	921.15
7	921	1.99	918.82	922.81
8	922	1.97	920.46	924.39
9	924	1.92	922.06	925.90
10	925	1.87	923.60	927.34
11	927	1.90	925.06	928.86
12	928	1.97	926.53	930.46
13	930	2.06	928.11	932.24
14	932	2.14	929.89	934.16
15	934	2.09	931.78	935.96
16	936	2.04	933.58	937.66
17	937	2.06	935.29	939.42
18	939	2.21	937.04	941.46
19	942	2.50	939.01	944.00
20	944	2.57	941.77	946.90
21	947	2.11	944.48	948.70
22	948	1.99	946.28	950.26
23	949*	2.40	947.74	952.54
24	950*	4.20	948.63	957.03

^ Truncated

* Adjusted for end of scale effect

Table 6.2.4I

Raw Score to Proficiency Level Conversion: Writ 3-5

Raw Score	Grade 3			Grade 4			Grade 5		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	5.88	5.88	A1	4.82	4.82	A1	4.82	4.82
1	A1	1.26	7.14	A1	1.07	5.89	A1	0.82	5.64
2	A1	0.93	8.06	A1	1.27	7.17	A1	1.26	6.90
3	A1	3.57	11.63	A1	2.68	9.85	A1	2.89	9.79
4	A1	1.85	13.48	A1	1.34	11.19	A1	1.48	11.28
5	A1	3.64	17.12	A1	2.81	14.00	A1	2.08	13.35
6	A1	1.59	18.70	A1	1.54	15.54	A1	1.41	14.76
7	A1	1.59	20.29	A1	1.61	17.15	A1	0.89	15.65
8	A1	1.45	21.74	A1	1.14	18.29	A1	0.89	16.54
9	A2	1.12	22.87	A2	1.00	19.29	A2	0.74	17.28
10	A2	3.11	25.97	A2	2.08	21.37	A2	1.71	18.99
11	A2	3.77	29.74	A2	2.48	23.84	A2	2.23	21.22
12	A2	11.04	40.78	A2	9.65	33.49	A2	9.42	30.64
13	A2	5.82	46.60	A2	4.96	38.45	A2	4.67	35.31
14	A3	15.27	61.86	A3	13.80	52.24	A3	14.24	49.55
15	A3	2.12	63.98	A3	1.67	53.92	A3	2.08	51.63
16	A3	4.49	68.47	A3	6.16	60.08	A3	5.04	56.68
17	A3	5.55	74.03	A3	3.68	63.76	A3	5.27	61.94
18	P1	5.22	79.25	P1	5.16	68.92	P1	6.08	68.03
19	P1	6.74	85.99	P1	7.10	76.02	P1	6.75	74.78
20	P1	4.82	90.81	P1	8.71	84.73	P1	8.68	83.46
21	P2	2.71	93.52	P2	4.76	89.48	P2	5.27	88.72
22	P2	4.30	97.82	P2	6.56	96.05	P2	6.01	94.73
23	P2	0.93	98.74	P2	1.27	97.32	P2	1.71	96.44
24	P3	1.26	100.00	P3	2.68	100.00	P3	3.56	100.00

Table 6.2.4J

Accuracy and Consistency of Classification Indices: Writ 3-5

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.624	0.561		0.445	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.870		0.158	
	A2	0.639		0.245	
	A3	0.668		0.196	
	P1	0.474		0.454	
	P2	-		0.500	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.949	0.029	0.022	0.927
	A2/A3	0.921	0.039	0.040	0.892
	A3/P1	0.914	0.021	0.065	0.879
	P1/P2	0.835	0.165	0.000	0.835

6.2.5 Oral Language Composite 3-5

Figure 6.2.5A

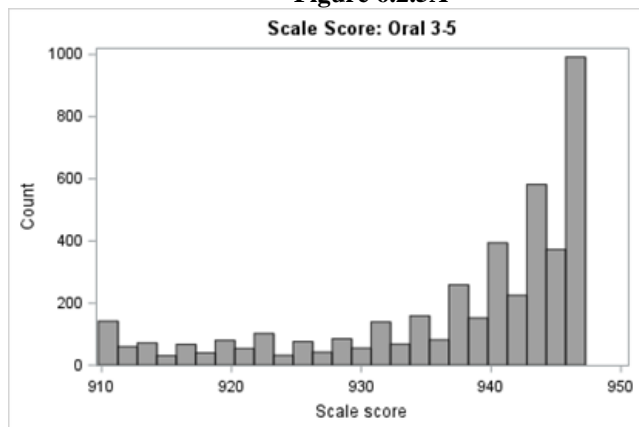


Table 6.2.5A

Scale Score Descriptive Statistics: Oral 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,516	910	947	935.81	10.82
4	1,487	910	947	937.77	10.48
5	1,360	910	947	938.56	10.10
Total	4,363	910	947	937.33	10.54

Figure 6.2.5B

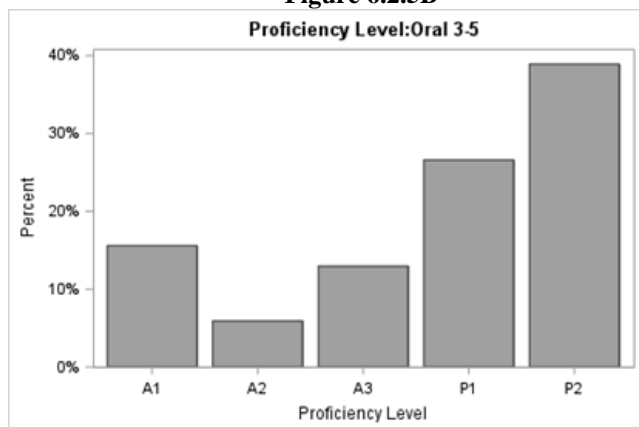


Table 6.2.5C

Proficiency Level Distribution: Oral 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	275	18.14	227	15.27	178	13.09	680	15.59
A2	117	7.72	76	5.11	65	4.78	258	5.91
A3	235	15.50	179	12.04	152	11.18	566	12.97
P1	420	27.70	387	26.03	354	26.03	1,161	26.61
P2	469	30.94	618	41.56	611	44.93	1,698	38.92
Total	1,516	100.00	1,487	100.00	1,360	100.00	4,363	100.00

Table 6.2.5D

n/a

Figure 6.2.5D
n/a

Figure 6.2.5E
n/a

Table 6.2.5E

Reliability: Oral 3-5

Component	Weight	Variance	Reliability
Listening	0.5	100.169	0.933
Speaking	0.5	163.981	0.965
Oral		111.185	0.972

*Variances from students who had results in all four domains

Table 6.2.5F
n/a

Table 6.2.5G
n/a

Table 6.2.5H
n/a

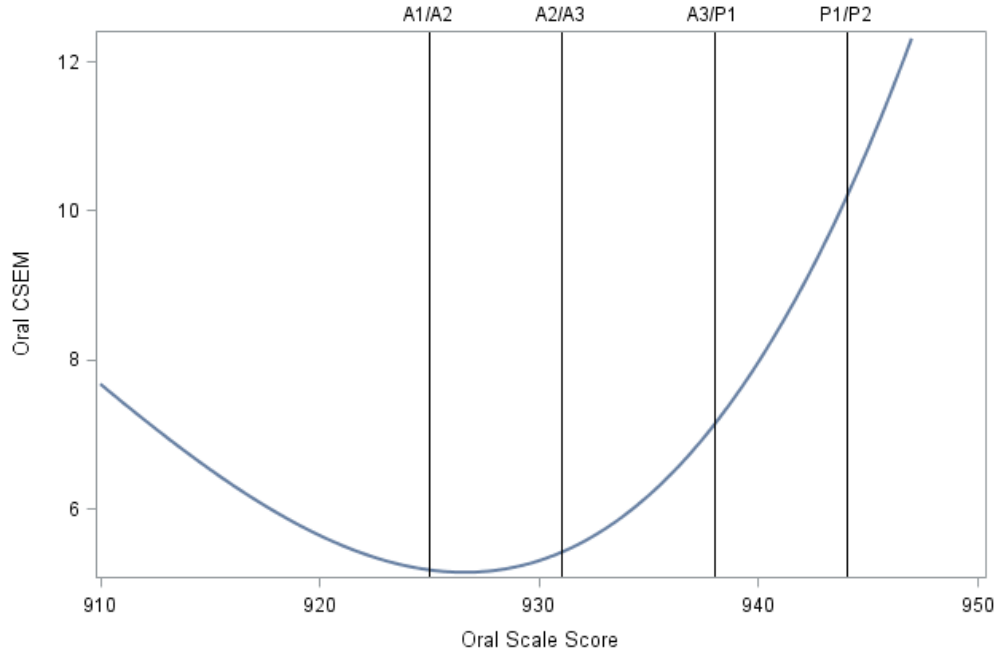
Table 6.2.5I
n/a

Table 6.2.5J

Accuracy and Consistency of Classification Indices: Oral 3-5

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.607	0.608		0.467	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.949		0.923	
	A2	0.632		0.509	
	A3	0.799		0.706	
	P1	0.487		0.500	
	P2	-		0.569	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.983	0.009	0.009	0.975
	A2/A3	0.976	0.013	0.011	0.966
	A3/P1	0.969	0.010	0.021	0.957
	P1/P2	0.680	0.320	0.000	0.708

Figure 6.2.5F CSEM for Oral Composite 3-5



6.2.6 Literacy Composite 3-5

Figure 6.2.6A

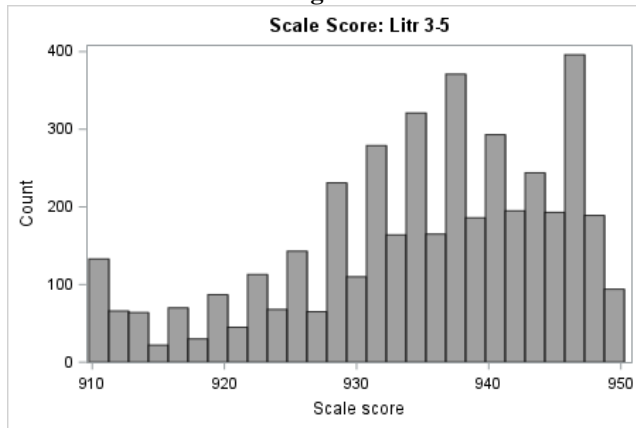


Table 6.2.6A

Scale Score Descriptive Statistics: Litr 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,507	910	950	933.05	9.96
4	1,488	910	950	935.31	10.20
5	1,342	910	950	936.23	9.87
Total	4,337	910	950	934.81	10.10

Figure 6.2.6B

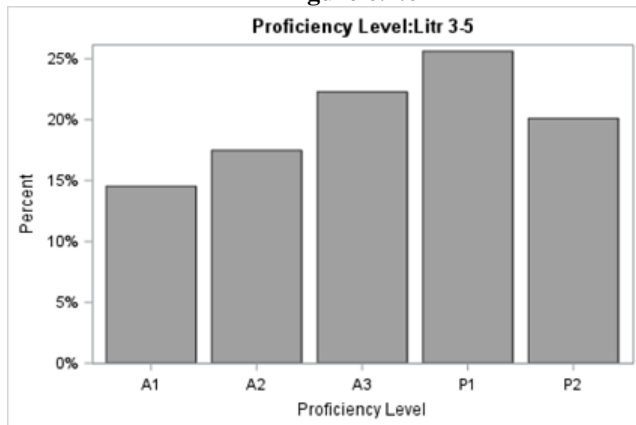


Table 6.2.6C

Proficiency Level Distribution: Litr 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	267	17.72	211	14.18	152	11.33	630	14.53
A2	306	20.31	239	16.06	213	15.87	758	17.48
A3	367	24.35	322	21.64	277	20.64	966	22.27
P1	374	24.82	368	24.73	369	27.50	1,111	25.62
P2	193	12.81	348	23.39	331	24.66	872	20.11
Total	1,507	100.00	1,488	100.00	1,342	100.00	4,337	100.00

Table 6.2.6D

n/a

Figure 6.2.6D
n/a

Figure 6.2.6E
n/a

Table 6.2.6E

Reliability: Litr 3-5

Component	Weight	Variance	Reliability
Reading	0.5	106.862	0.943
Writing	0.5	124.860	0.933
Literacy		102.029	0.964

*Variances from students who had results in all four domains

Table 6.2.6F
n/a

Table 6.2.6G
n/a

Table 6.2.6H
n/a

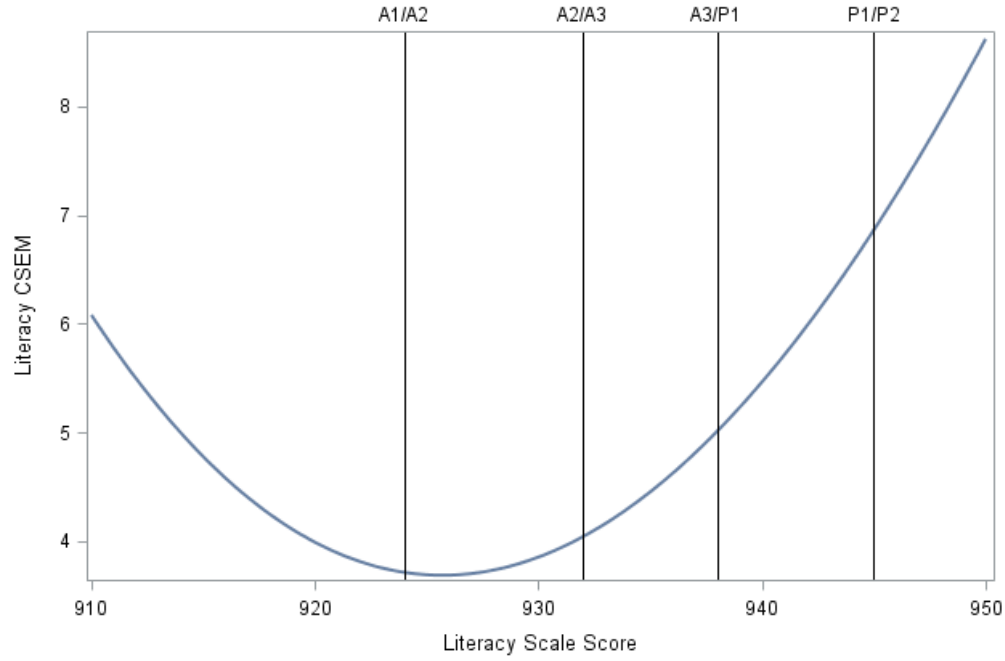
Table 6.2.6I
n/a

Table 6.2.6J

Accuracy and Consistency of Classification Indices: Litr 3-5

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.727	0.669		0.564	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.914		0.872	
	A2	0.753		0.653	
	A3	0.791		0.693	
	P1	0.625		0.602	
	P2	-		0.406	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.972	0.016	0.012	0.960
	A2/A3	0.951	0.026	0.023	0.931
	A3/P1	0.941	0.016	0.043	0.918
	P1/P2	0.863	0.137	0.000	0.857

Figure 6.2.6F CSEM for Literacy Composite 3-5



6.2.7 Comprehension Composite 3-5

Figure 6.2.7A

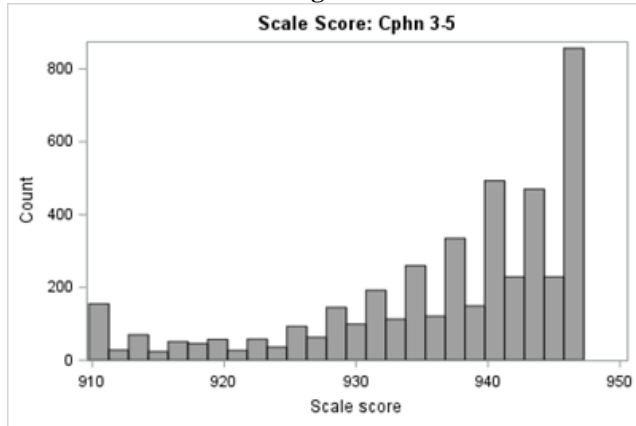


Figure 6.2.7B

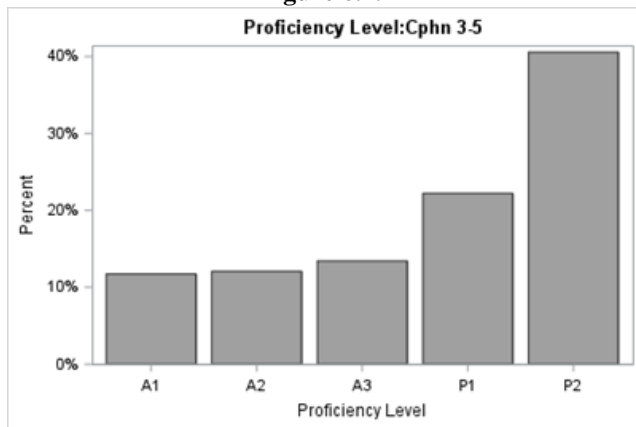


Table 6.2.7A

Scale Score Descriptive Statistics: Cphn 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,530	910	947	935.02	10.04
4	1,505	910	947	937.13	9.91
5	1,365	910	947	938.22	9.36
Total	4,400	910	947	936.73	9.88

Table 6.2.7C

Proficiency Level Distribution: Cphn 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	213	13.92	176	11.69	127	9.30	516	11.73
A2	241	15.75	162	10.76	128	9.38	531	12.07
A3	235	15.36	190	12.62	166	12.16	591	13.43
P1	362	23.66	340	22.59	275	20.15	977	22.20
P2	479	31.31	637	42.33	669	49.01	1,785	40.57
Total	1,530	100.00	1,505	100.00	1,365	100.00	4,400	100.00

Table 6.2.7D

n/a

Figure 6.2.7D

n/a

Figure 6.2.7E

n/a

Table 6.2.7E

Reliability: Cphn 3-5

Component	Weight	Variance	Reliability
Listening	0.3	100.169	0.933
Reading	0.7	106.862	0.943
Comprehension		97.534	0.963

*Variances from students who had results in all four domains

Table 6.2.7F

n/a

Table 6.2.7G

n/a

Table 6.2.7H

n/a

Table 6.2.7I

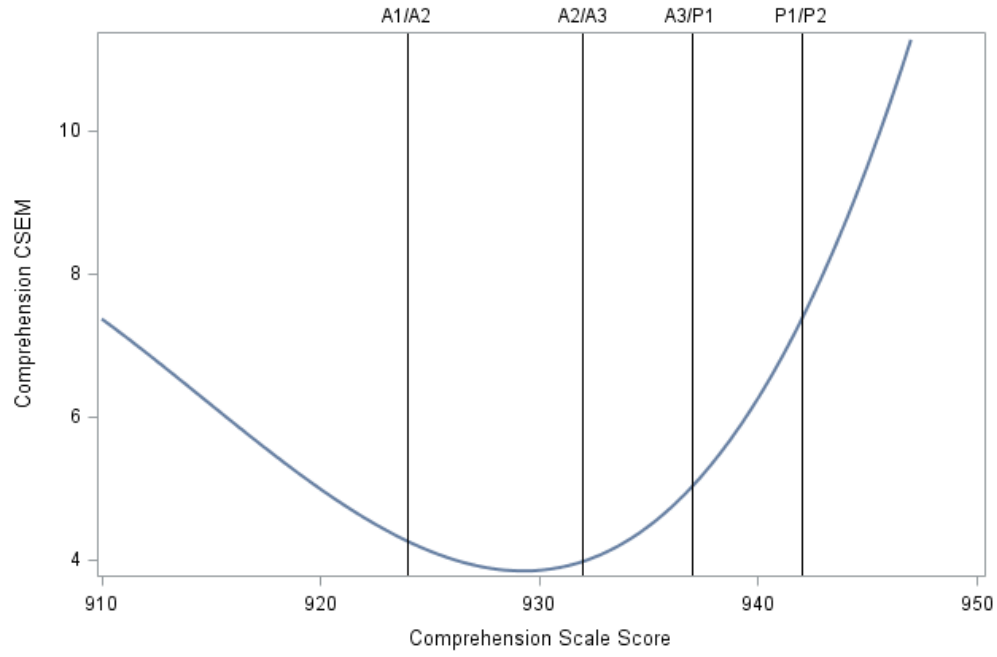
n/a

Table 6.2.7J

Accuracy and Consistency of Classification Indices: Cphn 3-5

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.562	0.560		0.423	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.918		0.880	
	A2	0.753		0.654	
	A3	0.748		0.630	
	P1	0.421		0.416	
	P2	-		0.595	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.980	0.012	0.008	0.972
	A2/A3	0.964	0.018	0.017	0.951
	A3/P1	0.957	0.009	0.034	0.940
	P1/P2	0.659	0.341	0.000	0.690

Figure 6.2.7F CSEM for Comprehension Composite 3-5



6.2.8 Overall Composite 3-5

Figure 6.2.8A

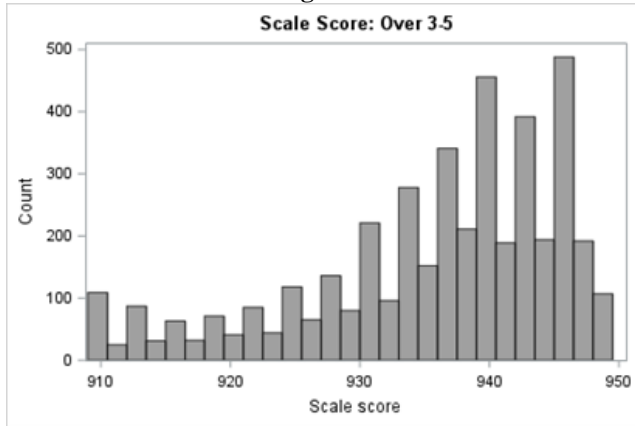


Table 6.2.8A

Scale Score Descriptive Statistics: Over 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	1,497	910	949	933.67	9.86
4	1,473	910	949	935.92	9.89
5	1,334	910	949	936.76	9.54
Total	4,304	910	949	935.40	9.86

Figure 6.2.8B

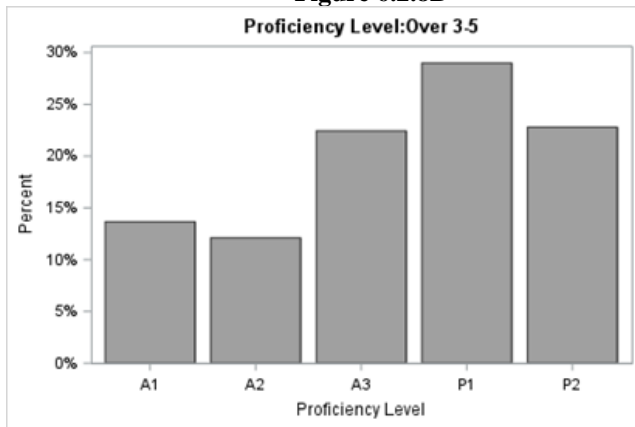


Table 6.2.8C

Proficiency Level Distribution: Over 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	249	16.63	201	13.65	138	10.34	588	13.66
A2	215	14.36	151	10.25	155	11.62	521	12.11
A3	385	25.72	323	21.93	258	19.34	966	22.44
P1	433	28.92	416	28.24	399	29.91	1,248	29.00
P2	215	14.36	382	25.93	384	28.79	981	22.79
Total	1,497	100.00	1,473	100.00	1,334	100.00	4,304	100.00

Table 6.2.8D

n/a

Figure 6.2.8D
n/a

Figure 6.2.8E
n/a

Table 6.2.8E

Reliability: Over 3-5

Component	Weight	Variance	Reliability
Listening	0.15	100.169	0.933
Reading	0.35	106.862	0.943
Speaking	0.15	163.981	0.965
Writing	0.35	124.860	0.933
Overall Composite		97.148	0.979

*Variances from students who had results in all four domains

Table 6.2.8F
n/a

Table 6.2.8G
n/a

Table 6.2.8H
n/a

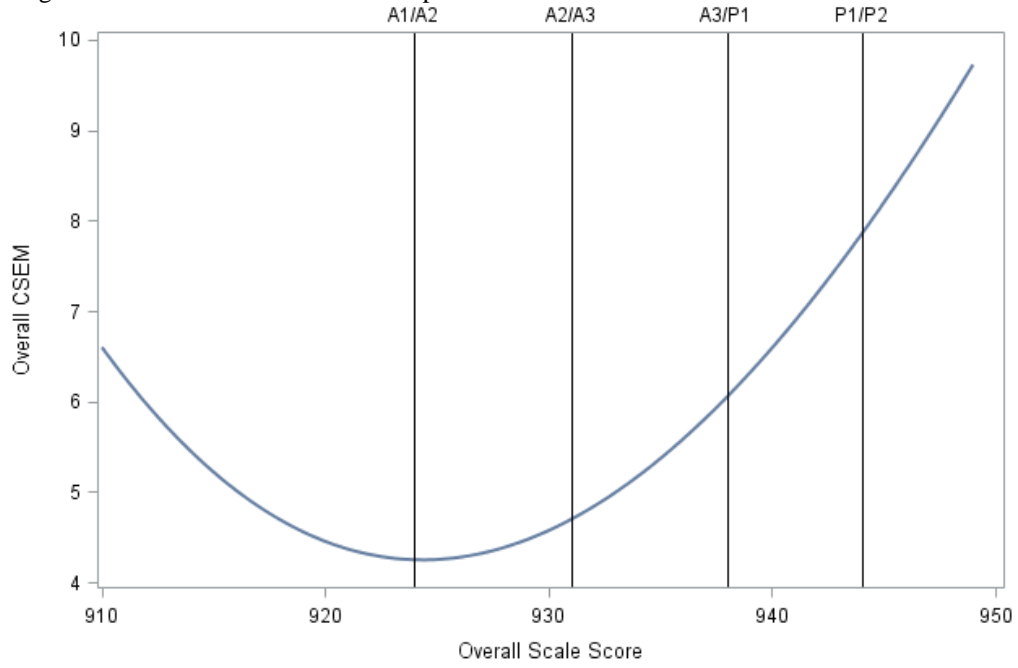
Table 6.2.8I
n/a

Table 6.2.8J

Accuracy and Consistency of Classification Indices: Over 3-5

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.708	0.671		0.567	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.941		0.913	
	A2	0.763		0.666	
	A3	0.876		0.814	
	P1	0.561		0.550	
	P2	-		0.499	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.982	0.010	0.008	0.974
	A2/A3	0.969	0.017	0.014	0.955
	A3/P1	0.958	0.010	0.031	0.943
	P1/P2	0.799	0.201	0.000	0.799

Figure 6.2.8F CSEM for Overall Composite 3-5



6.3 Grades: 6-8

6.3.1 Listening 6-8

Figure 6.3.1A

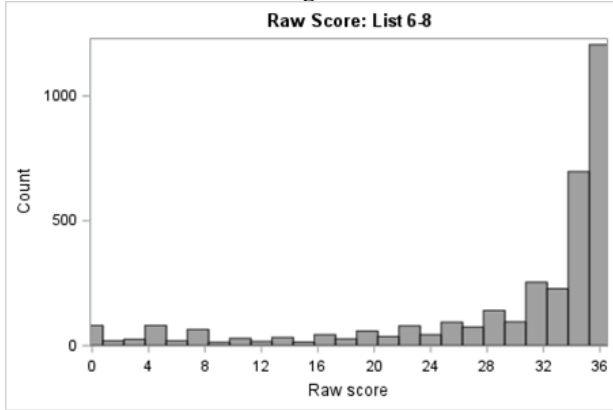


Table 6.3.1A

Raw Score Descriptive Statistics: List 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,270	0	36	29.18	9.81
7	1,127	0	36	29.81	9.58
8	1,076	0	36	29.73	9.82
Total	3,473	0	36	29.55	9.74

Figure 6.3.1B

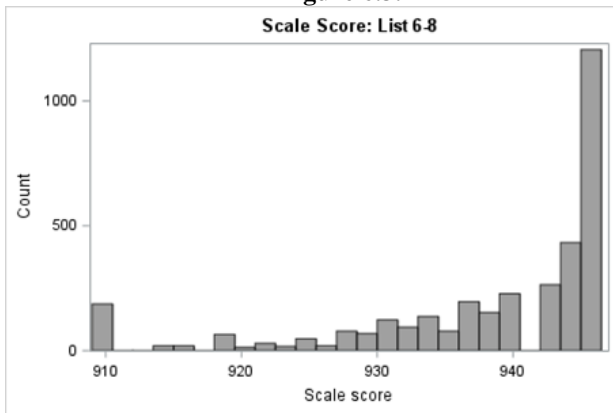


Table 6.3.1B

Scale Score Descriptive Statistics: List 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,270	910	946	937.67	10.00
7	1,127	910	946	938.45	9.82
8	1,076	910	946	938.33	10.12
Total	3,473	910	946	938.13	9.98

Figure 6.3.1C

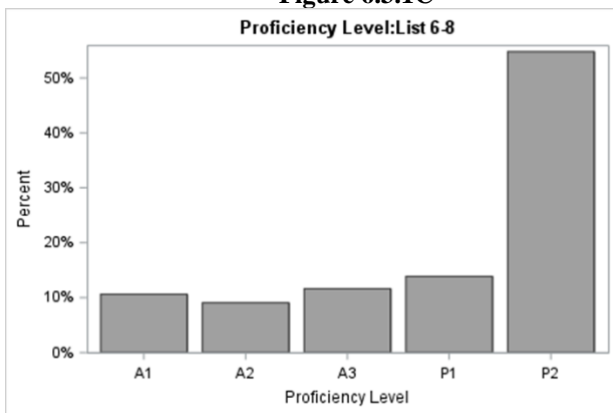


Table 6.3.1C

Proficiency Level Distribution: List 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	139	10.94	114	10.12	116	10.78	369	10.62
A2	128	10.08	103	9.14	84	7.81	315	9.07
A3	171	13.46	116	10.29	117	10.87	404	11.63
P1	174	13.70	154	13.66	154	14.31	482	13.88
P2	658	51.81	640	56.79	605	56.23	1,903	54.79
Total	1,270	100.00	1,127	100.00	1,076	100.00	3,473	100.00

Table 6.3.1D

Equating Summary: List 6-8

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 102. Thus, the results from the S102 of the Alternate ACCESS were used to determine raw-to-scale score conversion.

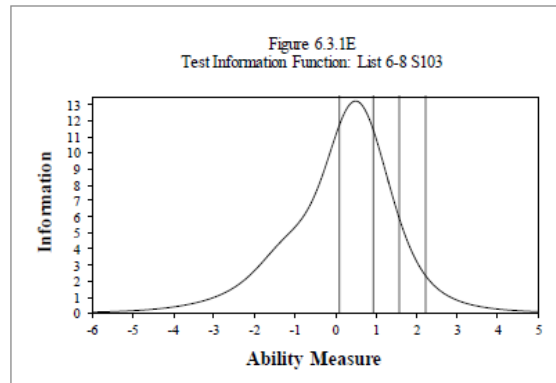
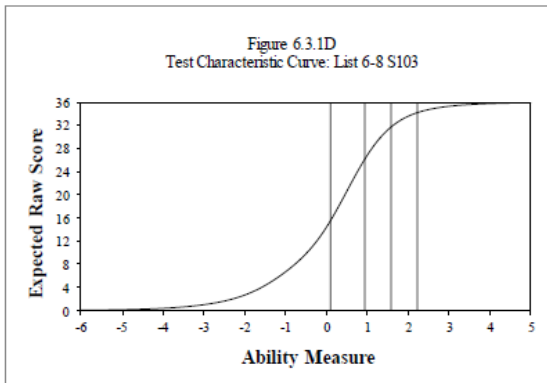


Table 6.3.1E

Reliability: List 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,473	9	0.942	2.344

Table 6.3.1F

Item Analysis Summary: List 6-8

Note: Item information is removed in the redacted version.

Table 6.3.1G

Complete Item Analysis: List 6-8

Note: Item information is removed in the redacted version.

Table 6.3.1H

Raw Score to Scale Score Conversion: List 6-8

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	14.72	910.00^	910.00^
1	910^	7.99	910.00^	910.00^
2	910^	5.54	910.00^	910.00^
3	910^	4.67	910.00^	911.13
4	910^	4.43	910.00^	913.42
5	911	4.35	910.00^	915.80
6	914	4.19	910.00^	917.93
7	916	3.96	911.84	919.75
8	918	3.64	913.98	921.26
9	919	3.32	915.80	922.44
10	920	3.09	917.38	923.55
11	922	2.85	918.73	924.42
12	923	2.69	919.83	925.21
13	923	2.53	920.78	925.85
14	924	2.45	921.65	926.56
15	925	2.37	922.44	927.19
16	926	2.29	923.24	927.83
17	926	2.22	923.95	928.38
18	927	2.22	924.58	929.01
19	927	2.22	925.21	929.65
20	928	2.14	925.85	930.12
21	929	2.14	926.48	930.75
22	929	2.22	926.96	931.39
23	930	2.22	927.59	932.02
24	930	2.22	928.22	932.65
25	931	2.29	928.78	933.36
26	932	2.37	929.41	934.16
27	932	2.45	930.04	934.95
28	933	2.53	930.67	935.74
29	934	2.69	931.39	936.77
30	935	2.85	932.18	937.88
31	936	3.09	933.05	939.22
32	937	3.40	934.08	940.88
33	939	3.96	935.18	943.10
34	941*	4.91	936.69	946.50
35	943*	7.36	938.75	953.46
36	945*	14.09	940.72	968.89

^ Truncated* Adjusted for end of scale effect

Table 6.3.11

Raw Score to Proficiency Level Conversion: List 6-8

Raw Score	Grade 6			Grade 7			Grade 8		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	2.76	2.76	A1	1.95	1.95	A1	2.23	2.23
1	A1	0.24	2.99	A1	0.18	2.13	A1	0	2.23
2	A1	0.24	3.23	A1	0.44	2.57	A1	0.56	2.79
3	A1	0.79	4.02	A1	0.62	3.19	A1	0.74	3.53
4	A1	1.57	5.59	A1	1.86	5.06	A1	1.95	5.48
5	A1	0.63	6.22	A1	0.27	5.32	A1	0.74	6.23
6	A1	0.47	6.69	A1	0.35	5.68	A1	0.84	7.06
7	A1	1.34	8.03	A1	1.60	7.28	A1	1.02	8.09
8	A1	0.24	8.27	A1	0.62	7.90	A1	0.84	8.92
9	A1	0.31	8.58	A1	0.18	8.07	A1	0.65	9.57
10	A1	0.24	8.82	A1	0.27	8.34	A1	0.37	9.94
11	A1	0.79	9.61	A1	0.53	8.87	A1	0.28	10.22
12	A1	0.39	10.00	A1	0.80	9.67	A1	0.28	10.50
13	A1	0.94	10.94	A1	0.44	10.12	A1	0.28	10.78
14	A2	0.31	11.26	A2	0.62	10.74	A2	0.19	10.97
15	A2	0.39	11.65	A2	0.44	11.18	A2	0.37	11.34
16	A2	0.55	12.20	A2	0.71	11.89	A2	0.37	11.71
17	A2	1.02	13.23	A2	0.80	12.69	A2	0.28	11.99
18	A2	0.79	14.02	A2	0.80	13.49	A2	0.74	12.73
19	A2	0.79	14.80	A2	0.35	13.84	A2	1.12	13.85
20	A2	1.02	15.83	A2	0.89	14.73	A2	0.84	14.68
21	A2	1.02	16.85	A2	1.33	16.06	A2	0.74	15.43
22	A2	0.94	17.80	A2	0.89	16.95	A2	0.93	16.36
23	A2	1.18	18.98	A2	1.15	18.10	A2	1.77	18.12
24	A2	2.05	21.02	A2	1.15	19.25	A2	0.46	18.59
25	A3	1.50	22.52	A3	1.06	20.32	A3	1.12	19.70
26	A3	1.89	24.41	A3	0.53	20.85	A3	1.95	21.65
27	A3	1.81	26.22	A3	2.31	23.16	A3	2.32	23.98
28	A3	1.81	28.03	A3	1.95	25.11	A3	1.67	25.65
29	A3	3.23	31.26	A3	1.60	26.71	A3	1.77	27.42
30	A3	3.23	34.49	A3	2.84	29.55	A3	2.04	29.46
31	P1	3.15	37.64	P1	2.84	32.39	P1	2.70	32.16
32	P1	4.02	41.65	P1	4.26	36.65	P1	5.02	37.17
33	P1	6.54	48.19	P1	6.57	43.21	P1	6.60	43.77
34	P2	7.80	55.98	P2	8.25	51.46	P2	6.69	50.46
35	P2	12.44	68.43	P2	12.87	64.33	P2	12.08	62.55
36	P2	31.57	100.00	P2	35.67	100.00	P2	37.45	100.00

Table 6.3.1J

Accuracy and Consistency of Classification Indices: List 6-8

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.748	0.639		0.464	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.890		0.129	
	A2	0.606		0.229	
	A3	0.651		0.161	
	P1	0.448		0.202	
	P2	0.807		0.781	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.978	0.014	0.008	0.967
	A2/A3	0.959	0.023	0.018	0.945
	A3/P1	0.950	0.013	0.038	0.931
	P1/P2	0.855	0.029	0.116	0.770

6.3.2 Reading 6-8

Figure 6.3.2A

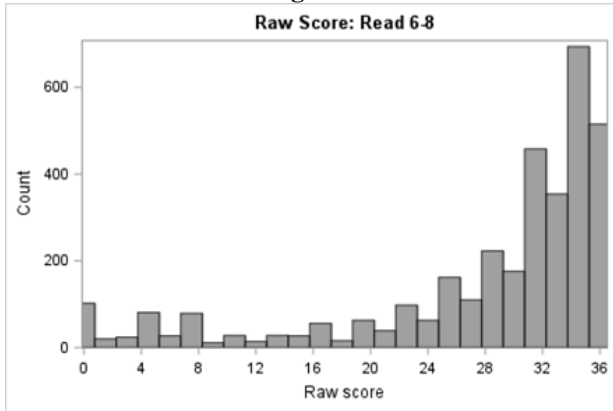


Table 6.3.2A

Raw Score Descriptive Statistics: Read 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,266	0	36	27.63	9.60
7	1,123	0	36	28.15	9.58
8	1,079	0	36	28.13	10.00
Total	3,468	0	36	27.96	9.72

Figure 6.3.2B

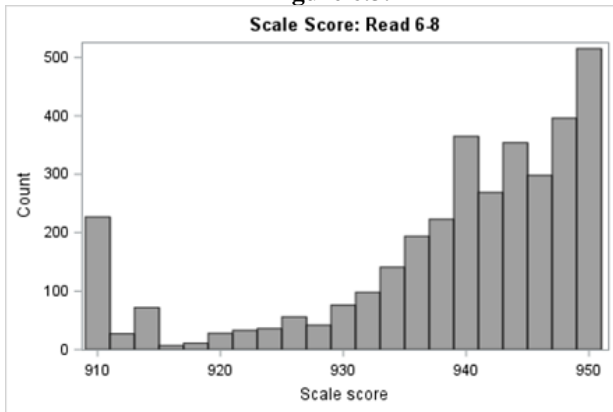


Table 6.3.2B

Scale Score Descriptive Statistics: Read 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,266	910	950	938.04	11.21
7	1,123	910	950	938.79	11.22
8	1,079	910	950	938.94	11.69
Total	3,468	910	950	938.56	11.37

Figure 6.3.2C

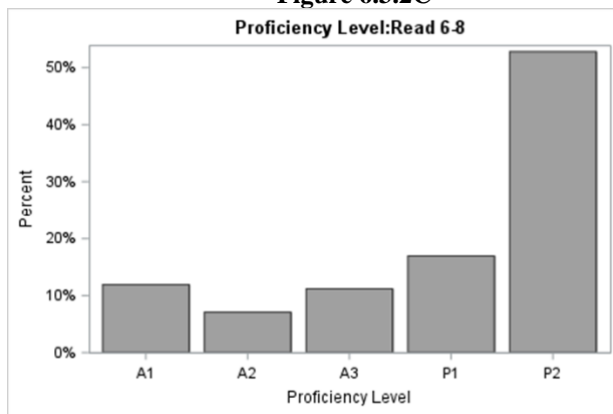


Table 6.3.2C

Proficiency Level Distribution: Read 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	151	11.93	129	11.49	134	12.42	414	11.94
A2	90	7.11	82	7.30	74	6.86	246	7.09
A3	167	13.19	117	10.42	104	9.64	388	11.19
P1	241	19.04	186	16.56	161	14.92	588	16.96
P2	617	48.74	609	54.23	606	56.16	1,832	52.83
Total	1,266	100.00	1,123	100.00	1,079	100.00	3,468	100.00

Table 6.3.2D

Equating Summary: Read 6-8

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

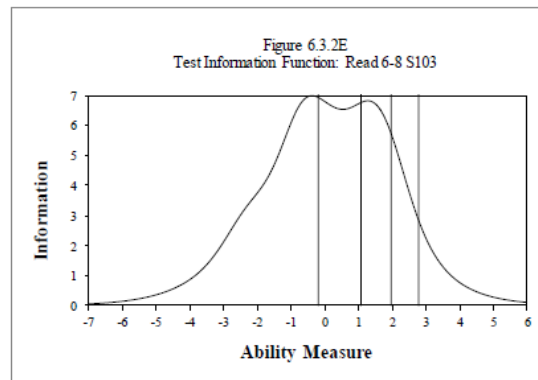
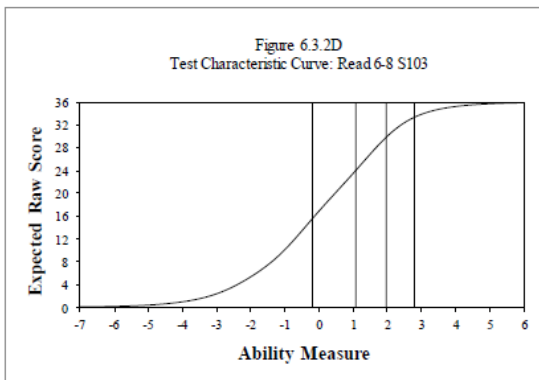


Table 6.3.2E

Reliability: Read 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,468	9	0.939	2.398

Table 6.3.2F

Item Analysis Summary: Read 6-8

Note: Item information is removed in the redacted version.

Table 6.3.2G

Complete Item Analysis: Read 6-8

Note: Item information is removed in the redacted version.

Table 6.3.2H

Raw Score to Scale Score Conversion: Read 6-8

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	11.87	910.00^	910.00^
1	910^	6.99	910.00^	910.00^
2	910^	4.82	910.00^	910.00^
3	910^	4.10	910.00^	910.00^
4	910^	3.92	910.00^	910.60
5	910^	3.86	910.00^	913.07
6	912	3.80	910.00^	915.42
7	914	3.56	910.36	917.47
8	916	3.25	912.59	919.10
9	917	3.01	914.46	920.49
10	919	2.77	916.09	921.63
11	920	2.65	917.41	922.71
12	921	2.53	918.62	923.68
13	922	2.47	919.70	924.64
14	923	2.41	920.73	925.55
15	924	2.41	921.69	926.51
16	925	2.41	922.65	927.48
17	926	2.41	923.62	928.44
18	927	2.41	924.58	929.40
19	928	2.47	925.49	930.43
20	929	2.47	926.51	931.45
21	930	2.47	927.54	932.48
22	931	2.47	928.56	933.50
23	932	2.47	929.58	934.53
24	933	2.47	930.61	935.55
25	934	2.47	931.63	936.57
26	935	2.47	932.60	937.54
27	936	2.47	933.62	938.56
28	937	2.47	934.65	939.59
29	938	2.53	935.61	940.67
30	939	2.59	936.63	941.82
31	940	2.77	937.66	943.20
32	942	3.01	938.80	944.83
33	944	3.43	940.07	946.94
34	946*	4.16	941.70	950.01
35	948*	5.97	943.93	955.86
36	950*	11.03	946.10	968.15

^ Truncated

* Adjusted for end of scale effect

Table 6.3.2I

Raw Score to Proficiency Level Conversion: Read 6-8

Raw Score	Grade 6			Grade 7			Grade 8		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	2.76	2.76	A1	2.76	2.76	A1	3.34	3.34
1	A1	0.24	3.00	A1	0.09	2.85	A1	0.09	3.43
2	A1	0.32	3.32	A1	0.36	3.21	A1	0.65	4.08
3	A1	0.55	3.87	A1	0.80	4.01	A1	0.74	4.82
4	A1	1.74	5.61	A1	1.42	5.43	A1	1.58	6.39
5	A1	0.95	6.56	A1	0.89	6.32	A1	0.37	6.77
6	A1	0.87	7.42	A1	0.53	6.86	A1	0.93	7.69
7	A1	2.29	9.72	A1	2.05	8.90	A1	1.85	9.55
8	A1	0.08	9.79	A1	0.27	9.17	A1	0.28	9.82
9	A1	0.24	10.03	A1	0.27	9.44	A1	0.46	10.29
10	A1	0.24	10.27	A1	0.09	9.53	A1	0.37	10.66
11	A1	0.55	10.82	A1	0.71	10.24	A1	0.46	11.12
12	A1	0.32	11.14	A1	0.36	10.60	A1	0.56	11.68
13	A1	0.55	11.69	A1	0.53	11.13	A1	0.56	12.23
14	A1	0.24	11.93	A1	0.36	11.49	A1	0.19	12.42
15	A2	0.63	12.56	A2	0.89	12.38	A2	0.83	13.25
16	A2	0.55	13.11	A2	0.98	13.36	A2	0.56	13.81
17	A2	1.11	14.22	A2	0.36	13.71	A2	1.30	15.11
18	A2	0.63	14.85	A2	0.45	14.16	A2	0.28	15.38
19	A2	1.03	15.88	A2	0.80	14.96	A2	0.37	15.76
20	A2	0.95	16.82	A2	1.16	16.12	A2	1.11	16.87
21	A2	1.26	18.09	A2	0.98	17.10	A2	1.11	17.98
22	A2	0.95	19.04	A2	1.69	18.79	A2	1.30	19.28
23	A3	1.74	20.77	A3	1.07	19.86	A3	1.76	21.04
24	A3	1.97	22.75	A3	1.96	21.82	A3	1.48	22.52
25	A3	2.92	25.67	A3	2.40	24.22	A3	1.30	23.82
26	A3	2.92	28.59	A3	1.96	26.18	A3	2.32	26.14
27	A3	3.63	32.23	A3	3.03	29.21	A3	2.78	28.92
28	P1	2.76	34.99	P1	2.67	31.88	P1	2.41	31.33
29	P1	4.42	39.42	P1	3.12	35.00	P1	3.80	35.13
30	P1	5.85	45.26	P1	5.16	40.16	P1	4.08	39.20
31	P1	6.00	51.26	P1	5.61	45.77	P1	4.63	43.84
32	P2	7.50	58.77	P2	8.37	54.14	P2	7.41	51.25
33	P2	10.43	69.19	P2	11.04	65.18	P2	9.08	60.33
34	P2	8.06	77.25	P2	8.37	73.55	P2	9.45	69.79
35	P2	10.66	87.91	P2	11.49	85.04	P2	12.23	82.02
36	P2	12.09	100.00	P2	14.96	100.00	P2	17.98	100.00

Table 6.3.2J

Accuracy and Consistency of Classification Indices: Read 6-8

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.786	0.720		0.561	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.927		0.152	
	A2	0.564		0.184	
	A3	0.538		0.197	
	P1	0.561		0.157	
	P2	0.847		0.825	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.980	0.011	0.009	0.971
	A2/A3	0.963	0.022	0.015	0.948
	A3/P1	0.944	0.024	0.032	0.925
	P1/P2	0.888	0.019	0.093	0.845

6.3.3 Speaking 6-8

Figure 6.3.3A

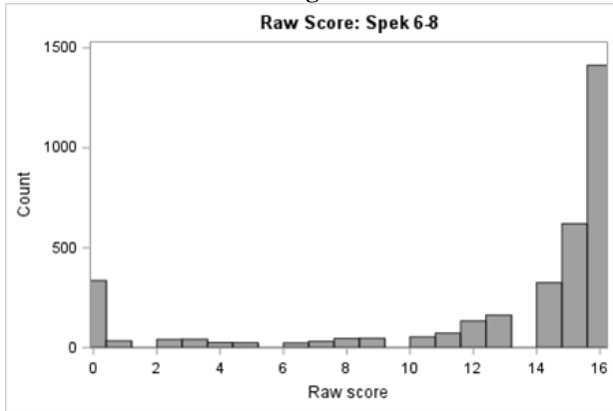


Table 6.3.3A

Raw Score Descriptive Statistics: Spek 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,255	0	16	12.50	5.23
7	1,114	0	16	12.61	5.19
8	1,060	0	16	12.62	5.33
Total	3,429	0	16	12.57	5.25

Figure 6.3.3B

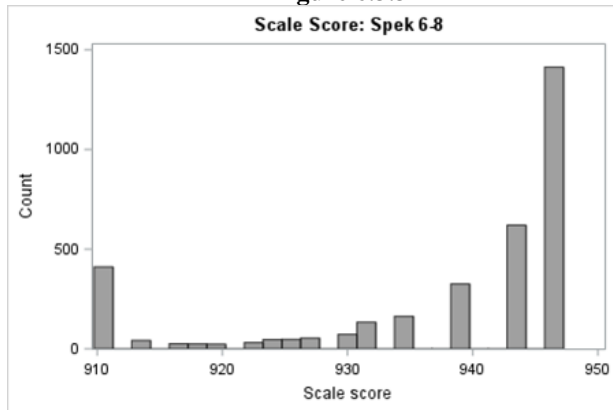


Table 6.3.3B

Scale Score Descriptive Statistics: Spek 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,255	910	947	937.16	12.55
7	1,114	910	947	937.49	12.47
8	1,060	910	947	937.66	12.82
Total	3,429	910	947	937.42	12.61

Figure 6.3.3C

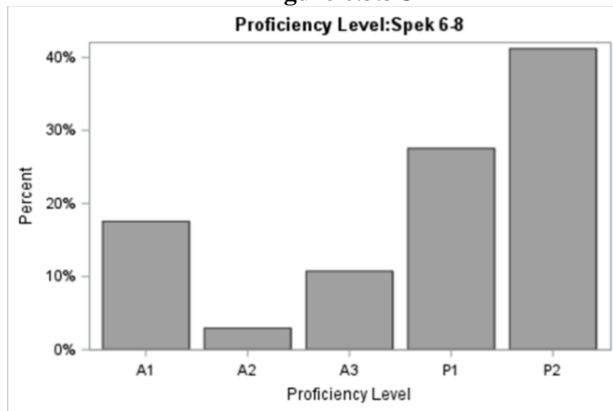


Table 6.3.3C

Proficiency Level Distribution: Spek 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	230	18.33	189	16.97	183	17.26	602	17.56
A2	28	2.23	45	4.04	28	2.64	101	2.95
A3	150	11.95	106	9.52	112	10.57	368	10.73
P1	357	28.45	324	29.08	264	24.91	945	27.56
P2	490	39.04	450	40.39	473	44.62	1,413	41.21
Total	1,255	100.00	1,114	100.00	1,060	100.00	3,429	100.00

Table 6.3.3D

Equating Summary: Spek 6-8

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

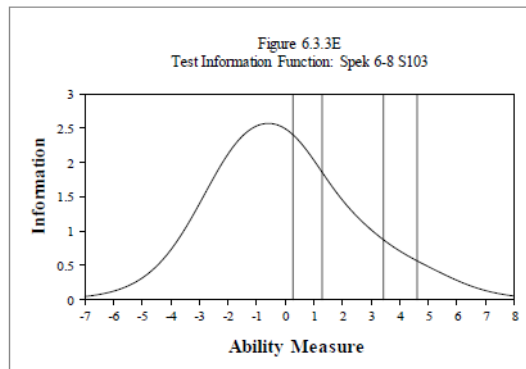
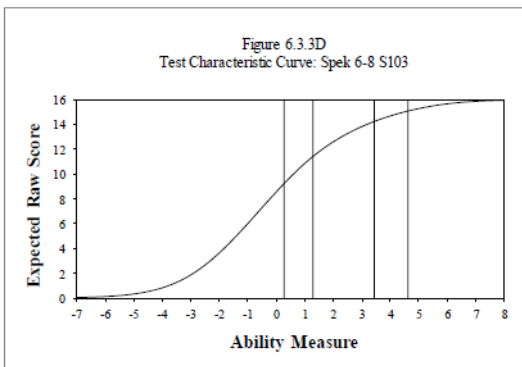


Table 6.3.3E

Reliability: Spek 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,429	8	0.964	0.990

Table 6.3.3F

Item Analysis Summary: Spek 6-8

Note: Item information is removed in the redacted version.

Table 6.3.3G

Complete Item Analysis: Spek 6-8

Note: Item information is removed in the redacted version.

Table 6.3.3H

Raw Score to Scale Score Conversion: Spek 6-8

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	8.33	910.00^	910.00^
1	910^	4.83	910.00^	911.90
2	911	3.72	910.00^	914.73
3	914	3.28	910.43	916.99
4	916	3.06	912.92	919.03
5	918	2.93	915.04	920.90
6	920	2.88	916.99	922.76
7	922	2.88	918.86	924.62
8	924	2.88	920.72	926.48
9	925	2.93	922.54	928.39
10	927	3.01	924.40	930.43
11	930	3.19	926.39	932.78
12	932	3.50	928.57	935.57
13	935	3.95	931.18	939.07
14	939	4.61	934.59	943.81
15	943*	5.94	939.43	951.31
16	947*	8.95	944.21	962.12

^ Truncated

* Adjusted for end of scale effect

Table 6.3.3I

Raw Score to Proficiency Level Conversion: Spek 6-8

Raw Score	Grade 6			Grade 7			Grade 8		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	9.80	9.80	A1	9.96	9.96	A1	9.62	9.62
1	A1	0.56	10.36	A1	0.72	10.68	A1	1.79	11.42
2	A1	1.27	11.63	A1	1.08	11.76	A1	1.23	12.64
3	A1	1.35	12.99	A1	0.81	12.57	A1	1.51	14.15
4	A1	1.12	14.10	A1	0.54	13.11	A1	0.47	14.62
5	A1	0.72	14.82	A1	0.81	13.91	A1	0.57	15.19
6	A1	0.80	15.62	A1	0.63	14.54	A1	0.57	15.75
7	A1	0.96	16.57	A1	1.08	15.62	A1	0.66	16.42
8	A1	1.75	18.33	A1	1.35	16.97	A1	0.85	17.26
9	A2	1.12	19.44	A2	2.06	19.03	A2	0.94	18.21
10	A2	1.12	20.56	A2	1.97	21.01	A2	1.70	19.91
11	A3	2.39	22.95	A3	1.89	22.89	A3	1.98	21.89
12	A3	4.38	27.33	A3	2.87	25.76	A3	4.34	26.23
13	A3	5.18	32.51	A3	4.76	30.52	A3	4.25	30.47
14	P1	9.96	42.47	P1	10.41	40.93	P1	7.92	38.40
15	P1	18.49	60.96	P1	18.67	59.61	P1	16.98	55.38
16	P2	39.04	100.00	P2	40.39	100.00	P2	44.62	100.00

Table 6.3.3J

Accuracy and Consistency of Classification Indices: Spek 6-8

Overall Indices	Accuracy	Consistency		Kappa (k)	
		0.560	0.579		0.428
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.945		0.239	
	A2	0.513		0.149	
	A3	0.708		0.080	
	P1	0.419		0.406	
	P2	-		0.568	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.980	0.012	0.009	0.971
	A2/A3	0.975	0.012	0.014	0.965
	A3/P1	0.957	0.011	0.032	0.935
	P1/P2	0.646	0.354	0.000	0.684

6.3.4 Writing 6-8

Figure 6.3.4A

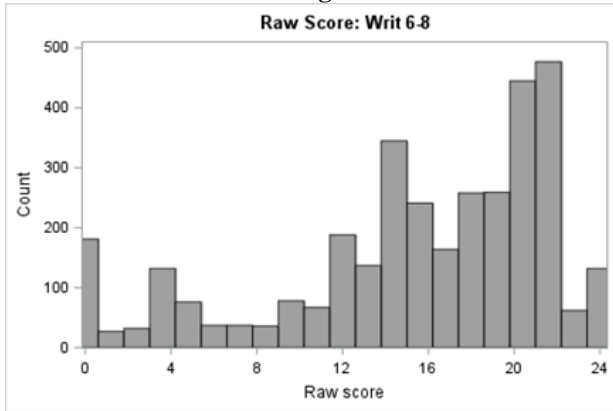


Table 6.3.4A

Raw Score Descriptive Statistics: Writ 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,243	0	24	14.94	6.51
7	1,107	0	24	15.19	6.49
8	1,061	0	24	15.60	6.44
Total	3,411	0	24	15.23	6.48

Figure 6.3.4B

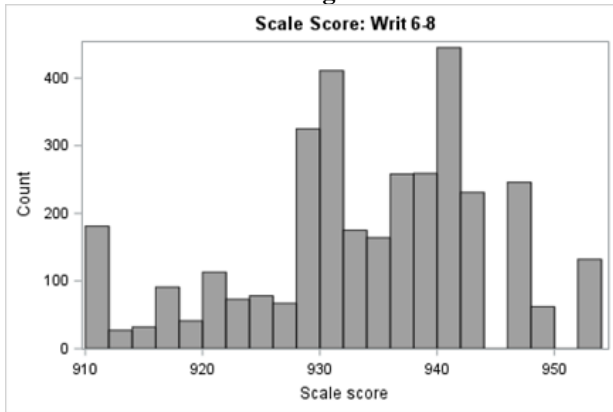


Table 6.3.4B

Scale Score Descriptive Statistics: Writ 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,243	910	953	932.98	10.30
7	1,107	910	953	933.38	10.36
8	1,061	910	953	934.13	10.40
Total	3,411	910	953	933.46	10.36

Figure 6.3.4C

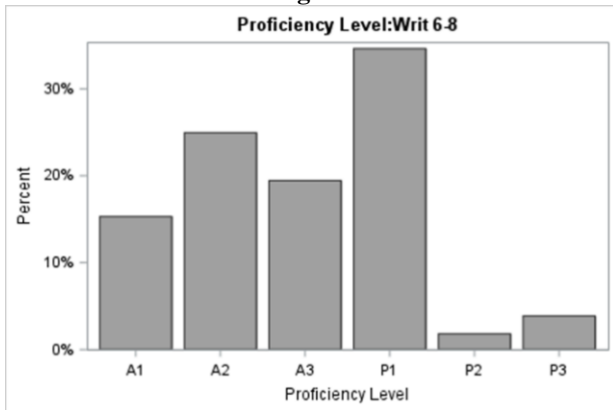


Table 6.3.4C

Proficiency Level Distribution: Writ 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	205	16.49	163	14.72	154	14.51	522	15.30
A2	317	25.50	278	25.11	256	24.13	851	24.95
A3	244	19.63	232	20.96	187	17.62	663	19.44
P1	414	33.31	371	33.51	396	37.32	1,181	34.62
P2	18	1.45	23	2.08	21	1.98	62	1.82
P3	45	3.62	40	3.61	47	4.43	132	3.87
Total	1,243	100.00	1,107	100.00	1,061	100.00	3,411	100.00

Table 6.3.4D

Equating Summary: Writ 6-8

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

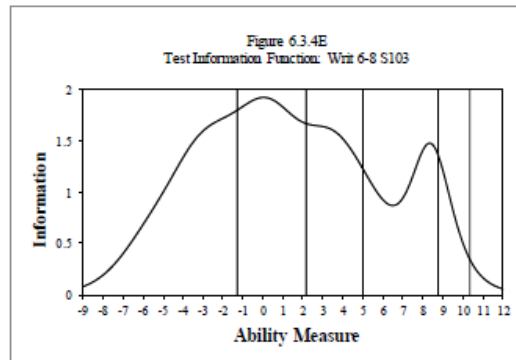
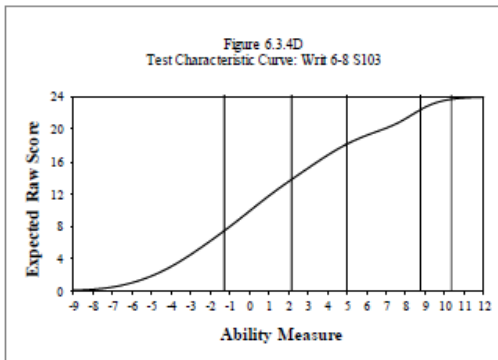


Table 6.3.4E

Reliability: Writ 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,411	10	0.941	1.576

Table 6.3.4F

Item Analysis Summary: Writ 6-8

Note: Item information is removed in the redacted version.

Table 6.3.4G

Complete Item Analysis: Writ 6-8

Note: Item information is removed in the redacted version.

Table 6.3.4H

Raw Score to Scale Score Conversion: Writ 6-8

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	4.68	910.00^	913.09
1	912	2.90	910.00^	914.98
2	915	2.28	912.54	917.10
3	917	1.99	914.70	918.68
4	918	1.82	916.38	920.02
5	920	1.73	917.79	921.25
6	921	1.68	919.06	922.42
7	922	1.66	920.26	923.58
8	923	1.63	921.42	924.68
9	924	1.63	922.52	925.78
10	925	1.63	923.65	926.91
11	926	1.63	924.75	928.02
12	928	1.68	925.86	929.22
13	929	1.73	927.01	930.46
14	930	1.75	928.23	931.74
15	931	1.78	929.50	933.06
16	933	1.82	930.82	934.47
17	934	1.92	932.17	936.01
18	936	2.09	933.66	937.83
19	938	2.42	935.41	940.26
20	941	2.76	938.00	943.52
21	943	2.33	941.17	945.82
22	946	2.16	943.35	947.67
23	949*	2.57	945.18	950.31
24	952*	4.42	946.38	955.21

^ Truncated

* Adjusted for end of scale effect

Table 6.3.4I

Raw Score to Proficiency Level Conversion: Writ 6-8

Raw Score	Grade 6			Grade 7			Grade 8		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	5.31	5.31	A1	5.51	5.51	A1	5.09	5.09
1	A1	0.80	6.11	A1	0.90	6.41	A1	0.66	5.75
2	A1	1.21	7.32	A1	0.90	7.32	A1	0.66	6.41
3	A1	2.74	10.06	A1	2.53	9.85	A1	2.73	9.14
4	A1	1.53	11.58	A1	0.90	10.75	A1	1.13	10.27
5	A1	2.65	14.24	A1	2.35	13.10	A1	1.60	11.88
6	A1	1.29	15.53	A1	0.54	13.64	A1	1.41	13.29
7	A1	0.97	16.49	A1	1.08	14.72	A1	1.23	14.51
8	A2	0.80	17.30	A2	1.63	16.35	A2	0.75	15.27
9	A2	0.72	18.02	A2	1.36	17.71	A2	0.09	15.36
10	A2	1.37	19.39	A2	1.81	19.51	A2	1.51	16.87
11	A2	2.17	21.56	A2	1.54	21.05	A2	2.17	19.04
12	A2	5.71	27.27	A2	5.33	26.38	A2	5.47	24.51
13	A2	4.51	31.78	A2	3.43	29.81	A2	4.05	28.56
14	A2	10.22	42.00	A2	10.03	39.84	A2	10.08	38.64
15	A3	2.17	44.17	A3	2.35	42.19	A3	1.23	39.87
16	A3	5.07	49.24	A3	5.51	47.70	A3	4.81	44.67
17	A3	6.03	55.27	A3	4.52	52.21	A3	3.68	48.35
18	A3	6.36	61.63	A3	8.58	60.79	A3	7.92	56.27
19	P1	7.80	69.43	P1	7.32	68.11	P1	7.63	63.90
20	P1	12.87	82.30	P1	12.47	80.58	P1	13.85	77.76
21	P1	6.36	88.66	P1	6.32	86.90	P1	7.73	85.49
22	P1	6.28	94.93	P1	7.41	94.31	P1	8.11	93.59
23	P2	1.45	96.38	P2	2.08	96.39	P2	1.98	95.57
24	P3	3.62	100.00	P3	3.61	100.00	P3	4.43	100.00

Table 6.3.4J

Accuracy and Consistency of Classification Indices: Writ 6-8

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.748	0.652		0.531	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.858		0.129	
	A2	0.776		0.300	
	A3	0.524		0.121	
	P1	0.765		0.775	
	P2	-		0.190	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.953	0.030	0.018	0.932
	A2/A3	0.932	0.022	0.046	0.907
	A3/P1	0.925	0.031	0.044	0.888
	P1/P2	0.936	0.064	0.000	0.908

6.3.5 Oral Language Composite 6-8

Figure 6.3.5A

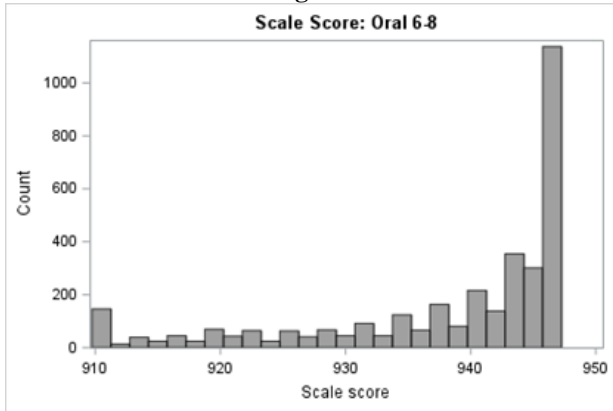


Table 6.3.5A

Scale Score Descriptive Statistics: Oral 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,254	910	947	937.88	10.68
7	1,109	910	947	938.46	10.60
8	1,059	910	947	938.46	10.87
Total	3,422	910	947	938.25	10.71

Figure 6.3.5B

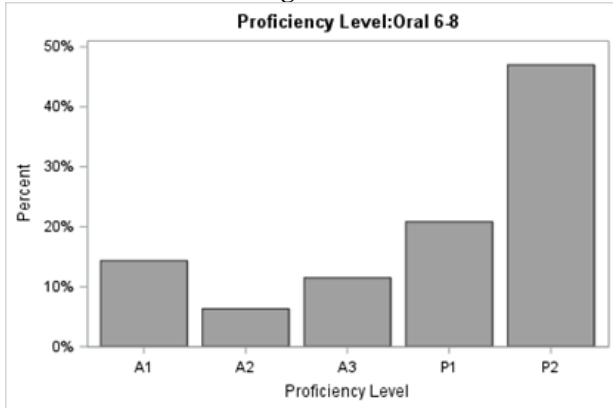


Table 6.3.5C

Proficiency Level Distribution: Oral 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	183	14.59	153	13.80	155	14.64	491	14.35
A2	82	6.54	75	6.76	59	5.57	216	6.31
A3	160	12.76	111	10.01	122	11.52	393	11.48
P1	274	21.85	249	22.45	190	17.94	713	20.84
P2	555	44.26	521	46.98	533	50.33	1,609	47.02
Total	1,254	100.00	1,109	100.00	1,059	100.00	3,422	100.00

Table 6.3.5D

n/a

Figure 6.3.5D

n/a

Figure 6.3.5E

n/a

Table 6.3.5E

Reliability: Oral 6-8

Component	Weight	Variance	Reliability
Listening	0.5	99.627	0.942
Speaking	0.5	158.907	0.964
Oral		114.742	0.975

*Variances from students who had results in all four domains

Table 6.3.5F

n/a

Table 6.3.5G

n/a

Table 6.3.5H

n/a

Table 6.3.5I

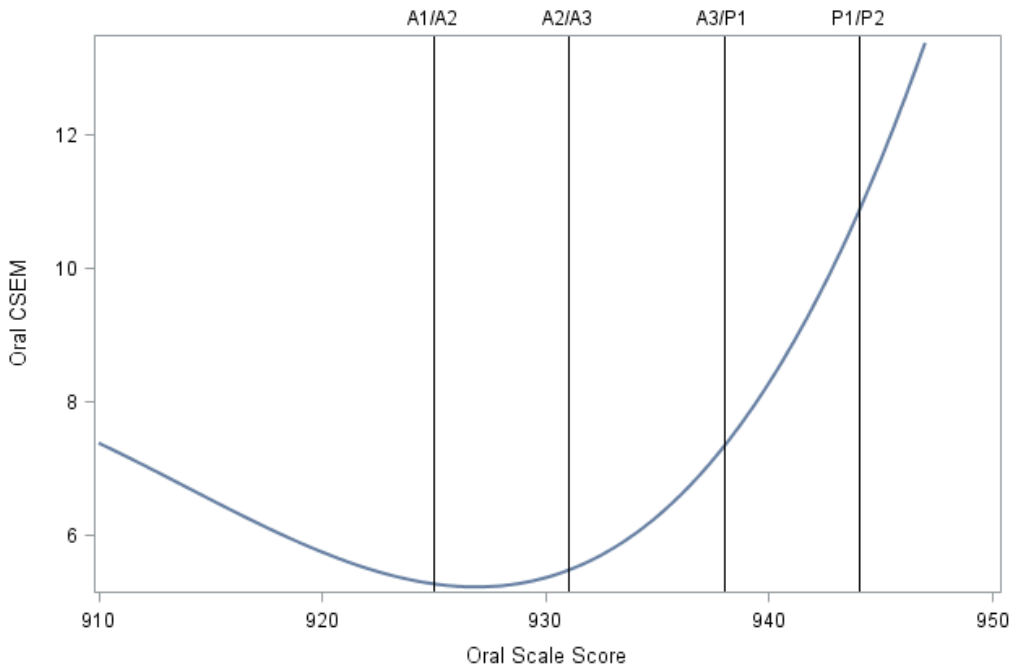
n/a

Table 6.3.5J

Accuracy and Consistency of Classification Indices: Oral 6-8

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.766	0.650		0.516	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.940		0.911	
	A2	0.648		0.526	
	A3	0.798		0.707	
	P1	0.603		0.393	
	P2	0.775		0.748	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.983	0.009	0.008	0.976
	A2/A3	0.975	0.014	0.012	0.964
	A3/P1	0.970	0.009	0.021	0.958
	P1/P2	0.838	0.048	0.114	0.749

Figure 6.3.5F CSEM for Oral Composite 6-8



6.3.6 Literacy Composite 6-8

Figure 6.3.6A

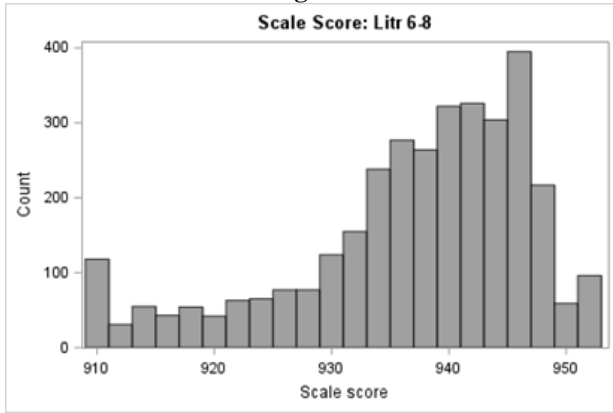


Table 6.3.6A

Scale Score Descriptive Statistics: Litr 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,241	910	952	935.84	10.00
7	1,103	910	952	936.39	10.15
8	1,058	910	952	936.89	10.34
Total	3,402	910	952	936.34	10.16

Figure 6.3.6B

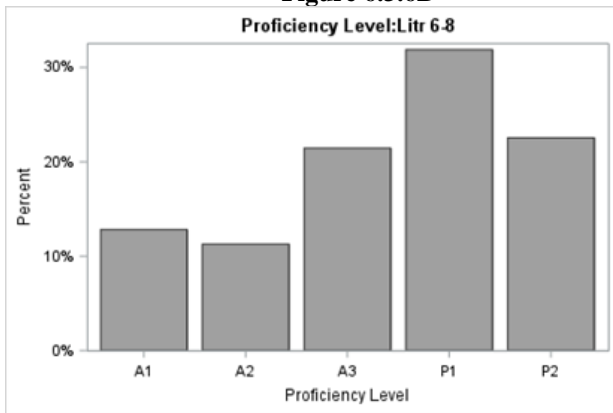


Table 6.3.6C

Proficiency Level Distribution: Litr 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	168	13.54	136	12.33	132	12.48	436	12.82
A2	147	11.85	134	12.15	103	9.74	384	11.29
A3	287	23.13	226	20.49	217	20.51	730	21.46
P1	399	32.15	362	32.82	324	30.62	1,085	31.89
P3	240	19.34	245	22.21	282	26.65	767	22.55
Total	1,241	100.00	1,103	100.00	1,058	100.00	3,402	100.00

Table 6.3.6D

n/a

Figure

6.3.6D n/a

Figure

6.3.6E n/a

Table 6.3.6E

Reliability: Litr 6-8

Component	Weight	Variance	Reliability
Reading	0.5	129.267	0.939
Writing	0.5	107.275	0.941
Literacy		103.230	0.966

*Variances from students who had results in all four domains

Table 6.3.6F

n/a

Table 6.3.6G

n/a

Table 6.3.6H

n/a

Table 6.3.6I

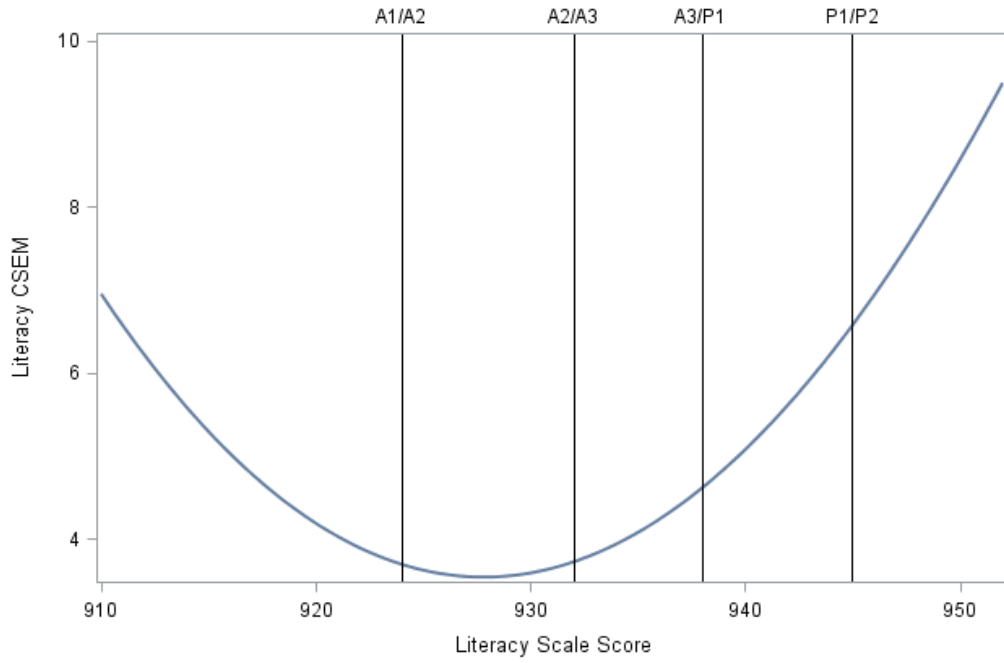
n/a

Table 6.3.6J

Accuracy and Consistency of Classification Indices: Litr 6-8

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.638	0.616		0.503	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.922		0.883	
	A2	0.740		0.636	
	A3	0.794		0.696	
	P1	0.499		0.503	
	P2	-		0.557	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.979	0.012	0.009	0.970
	A2/A3	0.959	0.023	0.018	0.943
	A3/P1	0.948	0.013	0.039	0.929
	P1/P2	0.751	0.249	0.000	0.771

Figure 6.3.6F CSEM for Literacy Composite 6-8



6.3.7 Comprehension Composite 6-8

Figure 6.3.7A

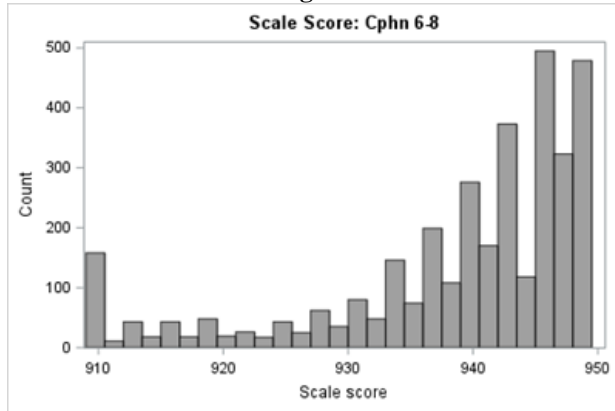


Table 6.3.7A

Scale Score Descriptive Statistics: Cphn 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,262	910	949	937.97	10.48
7	1,121	910	949	938.74	10.51
8	1,072	910	949	938.83	10.90
Total	3,455	910	949	938.49	10.63

Figure 6.3.7B

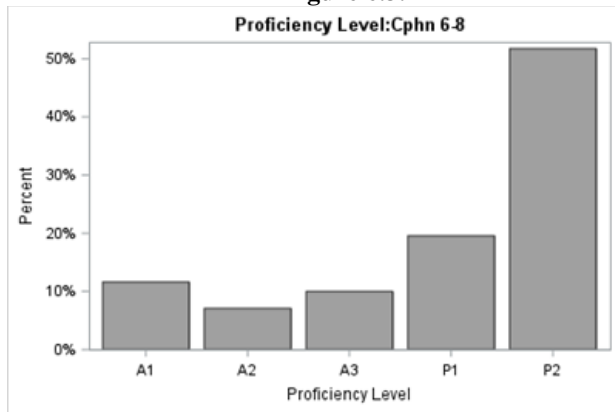


Table 6.3.7C

Proficiency Level Distribution: Cphn 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	146	11.57	124	11.06	131	12.22	401	11.61
A2	97	7.69	79	7.05	69	6.44	245	7.09
A3	140	11.09	110	9.81	95	8.86	345	9.99
P1	280	22.19	205	18.29	191	17.82	676	19.57
P2	599	47.46	603	53.79	586	54.66	1,788	51.75
Total	1,262	100.00	1,121	100.00	1,072	100.00	3,455	100.00

Table 6.3.7D

n/a

Figure 6.3.7D

n/a

Figure 6.3.7E

n/a

Table 6.3.7E

Reliability: Cphn 6-8

Component	Weight	Variance	Reliability
Listening	0.3	99.627	0.942
Reading	0.7	129.267	0.939
Comprehension		112.941	0.961

*Variances from students who had results in all four domains

Table 6.3.7F

n/a

Table 6.3.7G

n/a

Table 6.3.7H

n/a

Table 6.3.7I

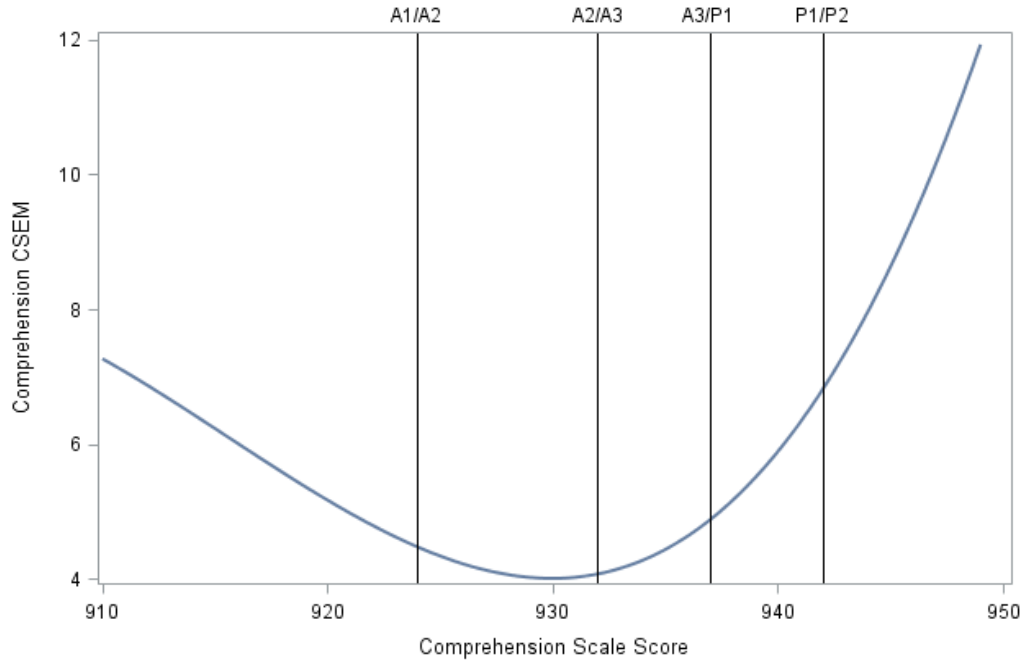
n/a

Table 6.3.7J

Accuracy and Consistency of Classification Indices: Cphn 6-8

Overall Indices	Accuracy	Consistency		Kappa (k)	
		0.823	0.763		0.633
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.948		0.920	
	A2	0.645		0.518	
	A3	0.603		0.478	
	P1	0.707		0.542	
	P2	0.870		0.849	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.985	0.007	0.008	0.979
	A2/A3	0.974	0.016	0.010	0.962
	A3/P1	0.955	0.022	0.023	0.938
	P1/P2	0.907	0.017	0.076	0.873

Figure 6.3.7F CSEM for Comprehension Composite 6-8



6.3.8 Overall Composite 6-8

Figure 6.3.8A

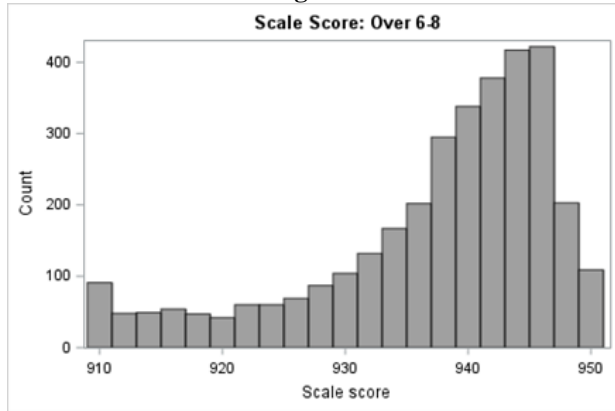


Table 6.3.8A

Scale Score Descriptive Statistics: Over 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	1,234	910	950	936.28	9.79
7	1,095	910	950	936.82	9.92
8	1,045	910	950	937.20	10.10
Total	3,374	910	950	936.74	9.93

Figure 6.3.8B

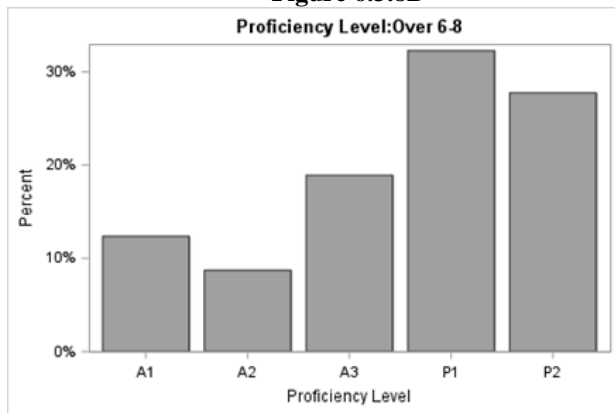


Table 6.3.8C

Proficiency Level Distribution: Over 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	158	12.80	133	12.15	126	12.06	417	12.36
A2	115	9.32	95	8.68	84	8.04	294	8.71
A3	251	20.34	204	18.63	183	17.51	638	18.91
P1	412	33.39	355	32.42	322	30.81	1,089	32.28
P2	298	24.15	308	28.13	330	31.58	936	27.74
Total	1,234	100.00	1,095	100.00	1,045	100.00	3,374	100.00

Table 6.3.8D

n/a

Figure 6.3.8D

n/a

Figure 6.3.8E

n/a

Table 6.3.8E

Reliability: Over 6-8

Component	Weight	Variance	Reliability
Listening	0.15	99.627	0.942
Reading	0.35	129.267	0.939
Speaking	0.15	158.907	0.964
Writing	0.35	107.275	0.941
Overall Composite		98.677	0.980

*Variances from students who had results in all four domains

Table 6.3.8F

n/a

Table 6.3.8G

n/a

Table 6.3.8H

n/a

Table 6.3.8I

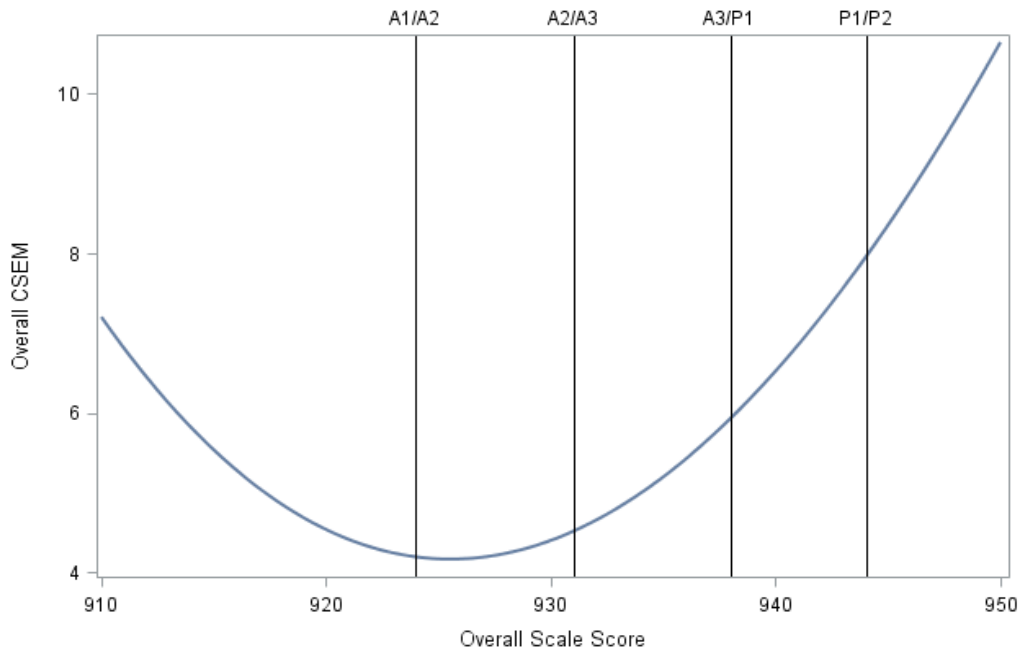
n/a

Table 6.3.8J

Accuracy and Consistency of Classification Indices: Over 6-8

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.738	0.657		0.552	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.954		0.932	
	A2	0.733		0.630	
	A3	0.887		0.829	
	P1	0.597		0.486	
	P2	0.684		0.637	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.987	0.007	0.006	0.981
	A2/A3	0.975	0.016	0.009	0.964
	A3/P1	0.963	0.010	0.027	0.950
	P1/P2	0.813	0.086	0.101	0.761

Figure 6.3.8F CSEM for Overall Composite 6-8



6.4 Grades: 9-12

6.4.1 Listening 9-12

Figure 6.4.1A

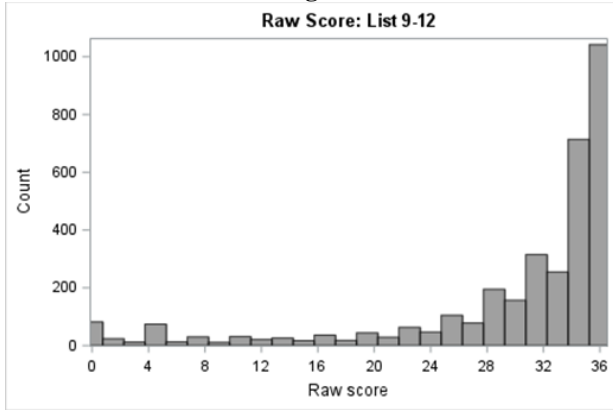


Table 6.4.1A

Raw Score Descriptive Statistics: List 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	870	0	36	29.13	9.80
10	824	0	36	29.73	9.10
11	716	0	36	30.46	8.47
12	1,029	0	36	30.00	9.14
Total	3,439	0	36	29.81	9.18

Figure 6.4.1B

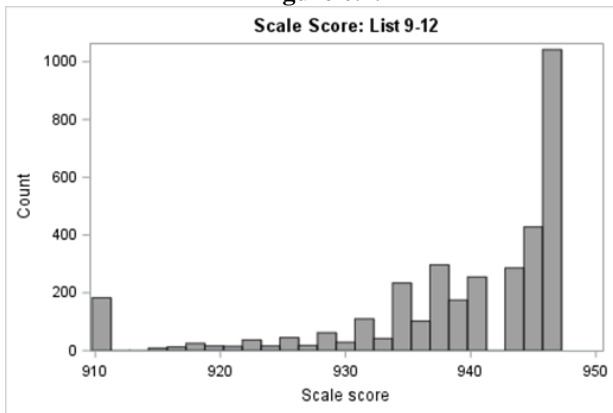


Table 6.4.1B

Scale Score Descriptive Statistics: List 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	870	910	947	938.31	10.14
10	824	910	947	938.92	9.68
11	716	910	947	939.72	9.09
12	1,029	910	947	939.28	9.70
Total	3,439	910	947	939.04	9.69

Figure 6.4.1C

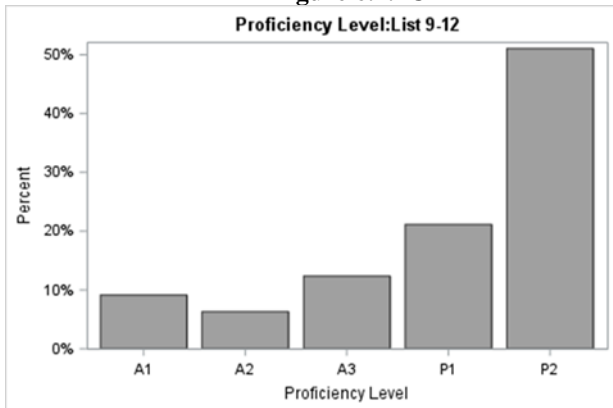


Table 6.4.1C

Proficiency Level Distribution: List 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	93	10.69	72	8.74	58	8.10	91	8.84	314	9.13
A2	65	7.47	52	6.31	39	5.45	61	5.93	217	6.31
A3	97	11.15	127	15.41	86	12.01	115	11.18	425	12.36
P1	196	22.53	157	19.05	153	21.37	221	21.48	727	21.14
P2	419	48.16	416	50.49	380	53.07	541	52.58	1,756	51.06
Total	870	100.00	824	100.00	716	100.00	1,029	100.00	3,439	100.00

Table 6.4.1D

Equating Summary: List 9-12

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 102. Thus, the results from the S102 of the Alternate ACCESS were used to determine raw-to-scale score conversion.

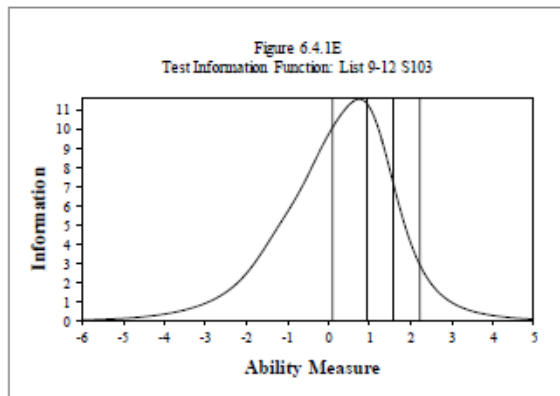
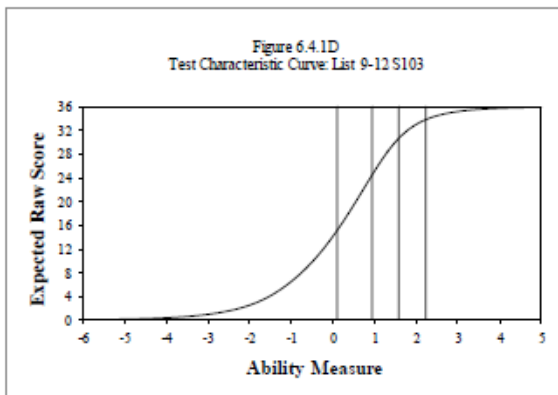


Table 6.4.1E

Reliability: List 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,439	9	0.938	2.284

Table 6.4.1F

Item Analysis Summary: List 9-12

Note: Item information is removed in the redacted version.

Table 6.4.1G

Complete Item Analysis: List 9-12

Note: Item information is removed in the redacted version.

Table 6.4.1H

Raw Score to Scale Score Conversion: List 9-12

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	14.88	910.00^	910.00^
1	910^	8.23	910.00^	910.00^
2	910^	5.70	910.00^	910.00^
3	910^	4.75	910.00^	910.81
4	910^	4.27	910.00^	912.87
5	911	4.11	910.00^	914.93
6	913	4.04	910.00^	916.98
7	915	3.88	911.05	918.80
8	917	3.64	913.03	920.31
9	918	3.40	914.85	921.65
10	920	3.17	916.43	922.76
11	921	2.93	917.86	923.71
12	922	2.77	919.04	924.58
13	923	2.69	920.07	925.45
14	924	2.61	921.02	926.24
15	924	2.53	921.89	926.96
16	925	2.45	922.76	927.67
17	926	2.45	923.55	928.46
18	927	2.37	924.34	929.09
19	927	2.37	925.06	929.80
20	928	2.37	925.85	930.60
21	929	2.37	926.56	931.31
22	930	2.37	927.27	932.02
23	930	2.37	927.98	932.73
24	931	2.37	928.70	933.44
25	932	2.45	929.33	934.24
26	933	2.45	930.12	935.03
27	933	2.53	930.83	935.90
28	934	2.61	931.54	936.77
29	935	2.69	932.34	937.72
30	936	2.85	933.13	938.82
31	937	3.01	934.08	940.09
32	938	3.32	935.03	941.67
33	940	3.88	936.13	943.89
34	942*	4.83	937.48	947.13
35	944*	7.12	939.46	953.70
36	946*	13.93	940.96	968.81

^ Truncated

* Adjusted for end of scale effect

Table 6.4.11

Raw Score to Proficiency Level Conversion: List 9-12

Raw Score	Grade 9			Grade 10			Grade 11			Grade 12		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	3.45	3.45	A1	2.55	2.55	A1	0.70	0.70	A1	2.53	2.53
1	A1	0.34	3.79	A1	0.12	2.67	A1	0.42	1.12	A1	0.10	2.62
2	A1	0.69	4.48	A1	0.12	2.79	A1	0.42	1.54	A1	0.58	3.21
3	A1	0.34	4.83	A1	0.24	3.03	A1	0.28	1.82	A1	0.49	3.69
4	A1	1.38	6.21	A1	2.43	5.46	A1	2.09	3.91	A1	1.75	5.44
5	A1	0.34	6.55	A1	0.12	5.58	A1	0.14	4.05	A1	0.39	5.83
6	A1	0.57	7.13	A1	0.36	5.95	A1	0.28	4.33	A1	0.29	6.12
7	A1	1.26	8.39	A1	0.61	6.55	A1	0.42	4.75	A1	0.58	6.71
8	A1	0.11	8.51	A1	0	6.55	A1	0.42	5.17	A1	0.10	6.80
9	A1	0.34	8.85	A1	0.49	7.04	A1	0.42	5.59	A1	0.10	6.90
10	A1	0.34	9.20	A1	0.24	7.28	A1	0.98	6.56	A1	0.29	7.19
11	A1	0.57	9.77	A1	0.61	7.89	A1	0.28	6.84	A1	0.39	7.58
12	A1	0.46	10.23	A1	0.61	8.50	A1	0.84	7.68	A1	0.58	8.16
13	A1	0.46	10.69	A1	0.24	8.74	A1	0.42	8.10	A1	0.68	8.84
14	A2	0.46	11.15	A2	0.36	9.10	A2	0.42	8.52	A2	0	8.84
15	A2	0.34	11.49	A2	0.73	9.83	A2	0.42	8.94	A2	0.49	9.33
16	A2	0.92	12.41	A2	0.24	10.07	A2	0.42	9.36	A2	0.49	9.82
17	A2	0.11	12.53	A2	0.61	10.68	A2	0.42	9.78	A2	0.87	10.69
18	A2	0.34	12.87	A2	0.36	11.04	A2	0.56	10.34	A2	0.78	11.47
19	A2	0.80	13.68	A2	0.85	11.89	A2	0.56	10.89	A2	0.29	11.76
20	A2	1.38	15.06	A2	0.73	12.62	A2	0.42	11.31	A2	0.19	11.95
21	A2	1.03	16.09	A2	0.73	13.35	A2	0.56	11.87	A2	0.97	12.93
22	A2	0.80	16.90	A2	0.97	14.32	A2	0.56	12.43	A2	0.87	13.80
23	A2	1.26	18.16	A2	0.73	15.05	A2	1.12	13.55	A2	0.97	14.77
24	A3	0.57	18.74	A3	1.94	16.99	A3	1.40	14.94	A3	1.55	16.33
25	A3	1.49	20.23	A3	0.85	17.84	A3	0.98	15.92	A3	1.46	17.78
26	A3	2.18	22.41	A3	2.31	20.15	A3	1.26	17.18	A3	1.55	19.34
27	A3	1.84	24.25	A3	2.91	23.06	A3	1.54	18.72	A3	2.62	21.96
28	A3	2.53	26.78	A3	3.64	26.70	A3	2.51	21.23	A3	2.24	24.20
29	A3	2.53	29.31	A3	3.76	30.46	A3	4.33	25.56	A3	1.75	25.95
30	P1	4.94	34.25	P1	4.25	34.71	P1	3.63	29.19	P1	5.15	31.10
31	P1	3.91	38.16	P1	3.28	37.99	P1	4.61	33.80	P1	4.47	35.57
32	P1	5.63	43.79	P1	5.46	43.45	P1	5.45	39.25	P1	4.08	39.65
33	P1	8.05	51.84	P1	6.07	49.51	P1	7.68	46.93	P1	7.77	47.42
34	P2	8.85	60.69	P2	8.13	57.65	P2	7.96	54.89	P2	8.26	55.69
35	P2	13.56	74.25	P2	12.01	69.66	P2	12.29	67.18	P2	11.95	67.64
36	P2	25.75	100.00	P2	30.34	100.00	P2	32.82	100.00	P2	32.36	100.00

Table 6.4.1J

Accuracy and Consistency of Classification Indices: List 9-12

Overall Indices	Accuracy	Consistency		Kappa (k)	
		0.756	0.652		0.470
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.913		0.150	
	A2	0.456		0.146	
	A3	0.730		0.220	
	P1	0.399		0.166	
	P2	0.838		0.810	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.981	0.010	0.009	0.971
	A2/A3	0.958	0.029	0.013	0.942
	A3/P1	0.936	0.014	0.050	0.916
	P1/P2	0.872	0.030	0.098	0.791

6.4.2 Reading 9-12

Figure 6.4.2A

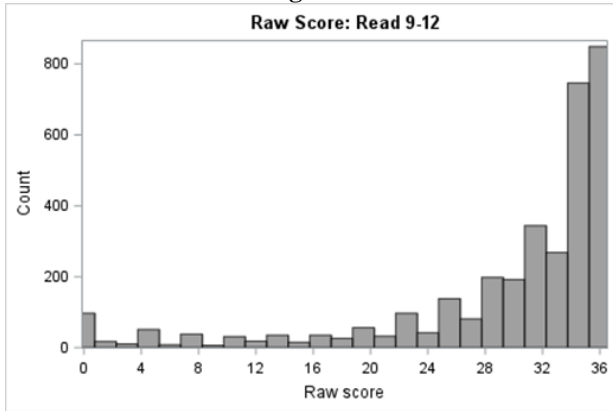


Table 6.4.2A

Raw Score Descriptive Statistics: Read 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	869	0	36	28.86	9.62
10	824	0	36	29.69	8.77
11	712	0	36	30.05	8.28
12	1,025	0	36	29.43	9.12
Total	3,430	0	36	29.47	9.01

Figure 6.4.2B

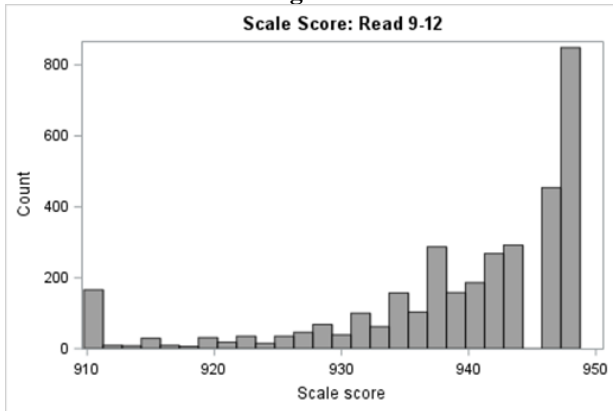


Table 6.4.2B

Scale Score Descriptive Statistics: Read 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	869	910	948	938.52	10.51
10	824	910	948	939.41	9.73
11	712	910	948	939.85	9.36
12	1,025	910	948	939.12	10.17
Total	3,430	910	948	939.19	10.00

Figure 6.4.2C

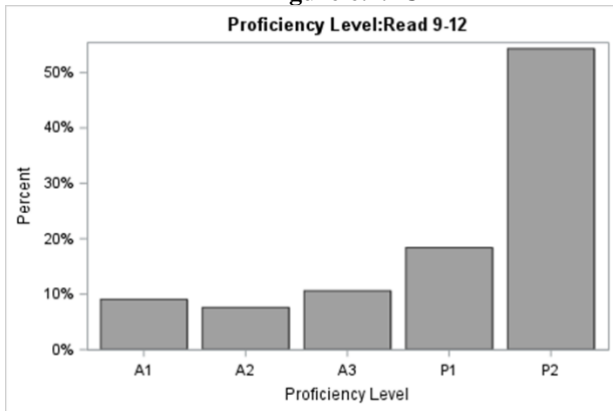


Table 6.4.2C

Proficiency Level Distribution: Read 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	88	10.13	65	7.89	61	8.57	97	9.46	311	9.07
A2	76	8.75	64	7.77	51	7.16	70	6.83	261	7.61
A3	92	10.59	96	11.65	68	9.55	108	10.54	364	10.61
P1	159	18.30	144	17.48	127	17.84	201	19.61	631	18.40
P2	454	52.24	455	55.22	405	56.88	549	53.56	1,863	54.31
Total	869	100.00	824	100.00	712	100.00	1,025	100.00	3,430	100.00

Table 6.4.2D

Equating Summary: Read 9-12

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

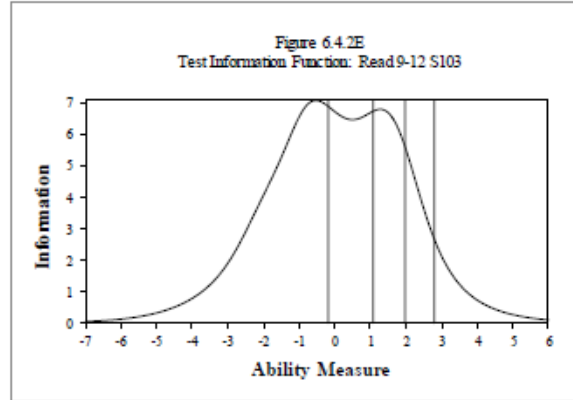
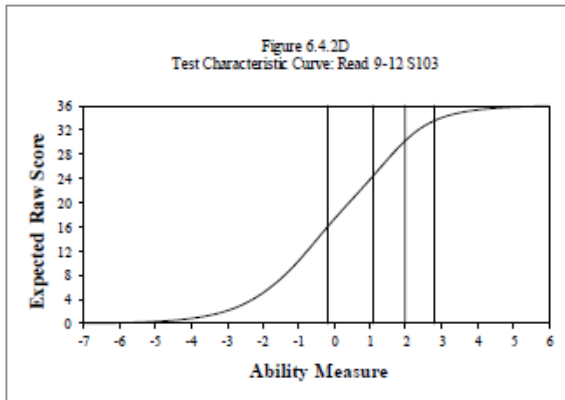


Table 6.4.2E

Reliability: Read 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,430	9	0.936	2.271

Table 6.4.2F

Item Analysis Summary: Read 9-12

Note: Item information is removed in the redacted version.

Table 6.4.2G

Complete Item Analysis: Read 9-12

Note: Item information is removed in the redacted version.

Table 6.4.2H

Raw Score to Scale Score Conversion: Read 9-12

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	11.51	910.00^	910.00^
1	910^	6.63	910.00^	910.00^
2	910^	4.70	910.00^	910.00^
3	910^	3.86	910.00^	911.45
4	910	3.50	910.00^	913.25
5	912	3.31	910.00^	915.00
6	913	3.25	910.24	916.75
7	915	3.13	912.05	918.32
8	917	2.95	913.74	919.64
9	918	2.77	915.30	920.85
10	919	2.59	916.63	921.81
11	920	2.47	917.83	922.78
12	921	2.35	918.92	923.62
13	922	2.29	919.88	924.46
14	923	2.29	920.73	925.31
15	924	2.29	921.63	926.21
16	925	2.29	922.47	927.05
17	926	2.29	923.32	927.90
18	927	2.35	924.16	928.86
19	927	2.35	925.06	929.77
20	928	2.35	926.03	930.73
21	929	2.41	926.87	931.69
22	930	2.41	927.84	932.66
23	931	2.41	928.80	933.62
24	932	2.35	929.83	934.53
25	933	2.35	930.73	935.43
26	934	2.35	931.63	936.33
27	935	2.35	932.54	937.24
28	936	2.35	933.44	938.14
29	937	2.41	934.34	939.17
30	938	2.47	935.25	940.19
31	939	2.65	936.15	941.46
32	940	2.83	937.24	942.90
33	942	3.25	938.32	944.83
34	944*	3.98	939.71	947.66
35	946*	5.72	941.64	953.09
36	948*	10.85	943.38	965.08

^ Truncated

* Adjusted for end of scale effect

Table 6.4.2I

Raw Score to Proficiency Level Conversion: Read 9-12

Raw Score	Grade 9			Grade 10			Grade 11			Grade 12		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Student	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	4.03	4.03	A1	3.03	3.03	A1	1.40	1.40	A1	2.63	2.63
1	A1	0.46	4.49	A1	0.24	3.28	A1	0.14	1.54	A1	0.10	2.73
2	A1	0.58	5.06	A1	0.12	3.40	A1	0.14	1.69	A1	0.20	2.93
3	A1	0.23	5.29	A1	0.12	3.52	A1	0.28	1.97	A1	0.49	3.41
4	A1	1.04	6.33	A1	0.97	4.49	A1	0.70	2.67	A1	1.95	5.37
5	A1	0.23	6.56	A1	0.49	4.98	A1	0.14	2.81	A1	0.20	5.56
6	A1	0.35	6.90	A1	0.12	5.10	A1	0.14	2.95	A1	0.29	5.85
7	A1	0.69	7.59	A1	0.73	5.83	A1	1.54	4.49	A1	0.59	6.44
8	A1	0.23	7.83	A1	0	5.83	A1	0.28	4.78	A1	0.49	6.93
9	A1	0.12	7.94	A1	0.12	5.95	A1	0	4.78	A1	0.39	7.32
10	A1	0.23	8.17	A1	0.12	6.07	A1	0.14	4.92	A1	0.49	7.80
11	A1	0.69	8.86	A1	0.73	6.80	A1	0.70	5.62	A1	0.49	8.29
12	A1	0.35	9.21	A1	0.24	7.04	A1	1.12	6.74	A1	0.49	8.78
13	A1	0.23	9.44	A1	0.49	7.52	A1	0.98	7.72	A1	0.29	9.07
14	A1	0.69	10.13	A1	0.36	7.89	A1	0.84	8.57	A1	0.39	9.46
15	A2	0.35	10.47	A2	0.73	8.62	A2	0.28	8.85	A2	0.39	9.85
16	A2	0.81	11.28	A2	0.73	9.34	A2	0.14	8.99	A2	0.29	10.15
17	A2	0.69	11.97	A2	0	9.34	A2	0.70	9.69	A2	0.68	10.83
18	A2	0.58	12.54	A2	0.85	10.19	A2	0.70	10.39	A2	0.88	11.71
19	A2	0.92	13.46	A2	0.24	10.44	A2	0.28	10.67	A2	0.78	12.49
20	A2	1.27	14.73	A2	1.21	11.65	A2	1.12	11.80	A2	0.68	13.17
21	A2	1.15	15.88	A2	0.85	12.50	A2	1.26	13.06	A2	0.59	13.76
22	A2	1.38	17.26	A2	1.33	13.83	A2	0.98	14.04	A2	0.88	14.63
23	A2	1.61	18.87	A2	1.82	15.66	A2	1.69	15.73	A2	1.66	16.29
24	A3	0.81	19.68	A3	1.58	17.23	A3	0.98	16.71	A3	1.46	17.76
25	A3	2.42	22.09	A3	2.06	19.30	A3	1.26	17.98	A3	1.46	19.22
26	A3	1.96	24.05	A3	2.91	22.21	A3	2.25	20.22	A3	1.85	21.07
27	A3	2.30	26.35	A3	2.43	24.64	A3	1.69	21.91	A3	2.83	23.90
28	A3	3.11	29.46	A3	2.67	27.31	A3	3.37	25.28	A3	2.93	26.83
29	P1	3.22	32.68	P1	2.31	29.61	P1	2.53	27.81	P1	2.93	29.76
30	P1	5.75	38.43	P1	4.61	34.22	P1	6.74	34.55	P1	5.46	35.22
31	P1	4.26	42.69	P1	4.98	39.20	P1	3.93	38.48	P1	5.07	40.29
32	P1	5.06	47.76	P1	5.58	44.78	P1	4.63	43.12	P1	6.15	46.44
33	P2	7.13	54.89	P2	7.77	52.55	P2	8.43	51.54	P2	8.00	54.44
34	P2	8.63	63.52	P2	9.83	62.38	P2	7.58	59.13	P2	8.00	62.44
35	P2	14.84	78.37	P2	13.35	75.73	P2	13.48	72.61	P2	11.61	74.05
36	P2	21.63	100.00	P2	24.27	100.00	P2	27.39	100.00	P2	25.95	100.00

Table 6.4.2J

Accuracy and Consistency of Classification Indices: Read 9-12

Overall Indices	Accuracy	Consistency		Kappa (k)	
		0.737	0.650		0.487
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.876		0.127	
	A2	0.605		0.207	
	A3	0.576		0.199	
	P1	0.530		0.196	
	P2	0.814		0.782	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.976	0.013	0.010	0.965
	A2/A3	0.951	0.027	0.021	0.932
	A3/P1	0.933	0.025	0.042	0.909
	P1/P2	0.867	0.030	0.103	0.809

6.4.3 Speaking 9-12

Figure 6.4.3A

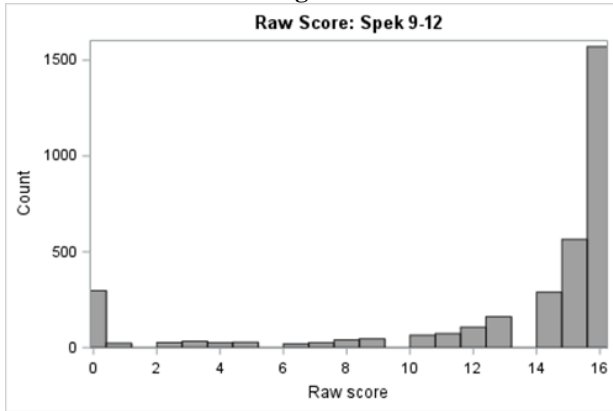


Table 6.4.3A

Raw Score Descriptive Statistics: Spek 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	857	0	16	12.85	5.09
10	815	0	16	12.99	4.89
11	708	0	16	13.09	4.91
12	1,014	0	16	12.88	5.09
Total	3,394	0	16	12.94	5.01

Figure 6.4.3B

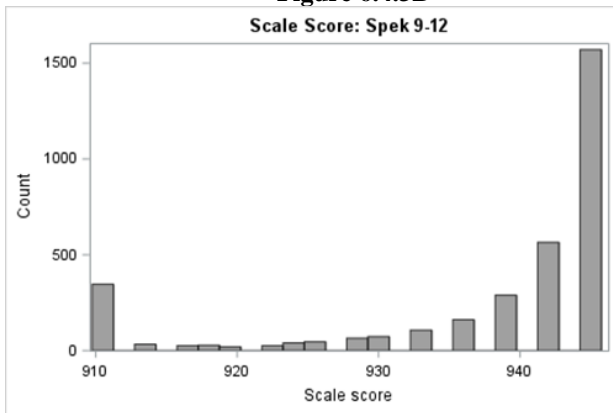


Table 6.4.3B

Scale Score Descriptive Statistics: Spek 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	857	910	945	937.19	11.58
10	815	910	945	937.45	11.20
11	708	910	945	937.76	11.24
12	1,014	910	945	937.28	11.57
Total	3,394	910	945	937.40	11.41

Figure 6.4.3C

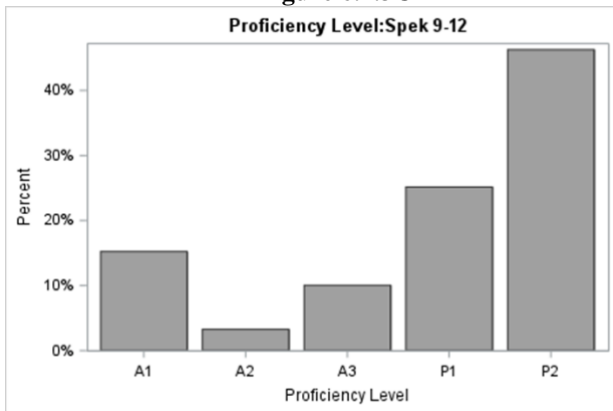


Table 6.4.3C

Proficiency Level Distribution: Spek 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	139	16.22	121	14.85	100	14.12	157	15.48	517	15.23
A2	25	2.92	28	3.44	26	3.67	32	3.16	111	3.27
A3	76	8.87	89	10.92	73	10.31	103	10.16	341	10.05
P1	244	28.47	207	25.40	155	21.89	248	24.46	854	25.16
P2	373	43.52	370	45.40	354	50.00	474	46.75	1,571	46.29
Total	857	100.00	815	100.00	708	100.00	1,014	100.00	3,394	100.00

Table 6.4.3D

Equating Summary: Spek 9-12

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

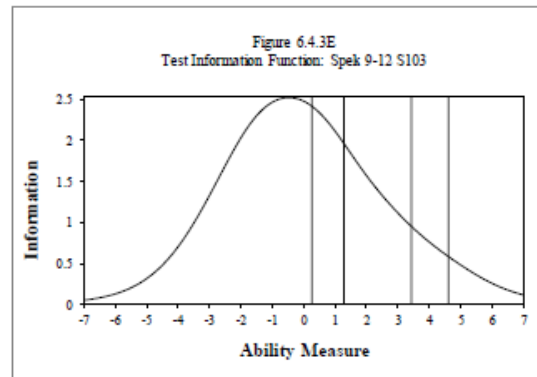
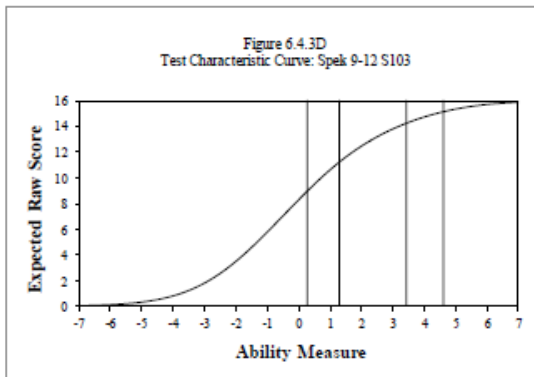


Table 6.4.3E

Reliability: Spek 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,394	8	0.963	0.951

Table 6.4.3F

Item Analysis Summary: Spek 9-12

Note: Item information is removed in the redacted version.

Table 6.4.3G

Complete Item Analysis: Spek 9-12

Note: Item information is removed in the redacted version.

Table 6.4.3H

Raw Score to Scale Score Conversion: Spek 9-12

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	8.56	910.00^	910.00^
1	910^	5.19	910.00^	911.14
2	910	3.95	910.00^	914.42
3	913	3.41	910.08	916.91
4	916	3.15	912.78	919.08
5	918	3.01	915.04	921.07
6	920	2.97	917.13	923.07
7	922	2.93	919.12	924.97
8	924	2.93	921.07	926.92
9	926	2.97	922.98	928.92
10	928	3.01	924.97	931.00
11	930	3.19	926.97	933.35
12	933	3.41	929.19	936.01
13	936	3.81	931.71	939.34
14	939	4.43	934.90	943.77
15	942*	5.67	939.25	950.60
16	945*	8.82	943.46	961.10

^ Truncated

* Adjusted for end of scale effect

Table 6.4.3I

Raw Score to Proficiency Level Conversion: Spek 9-12

Raw Score	Grade 9			Grade 10			Grade 11			Grade 12		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	8.87	8.87	A1	8.22	8.22	A1	8.19	8.19	A1	9.47	9.47
1	A1	1.17	10.04	A1	0.25	8.47	A1	0.71	8.90	A1	0.59	10.06
2	A1	0.70	10.74	A1	0.74	9.20	A1	0.71	9.60	A1	0.89	10.95
3	A1	1.28	12.02	A1	1.35	10.55	A1	0.85	10.45	A1	0.49	11.44
4	A1	0.35	12.37	A1	0.74	11.29	A1	0.85	11.30	A1	0.99	12.43
5	A1	1.17	13.54	A1	0.86	12.15	A1	0.85	12.15	A1	0.49	12.92
6	A1	0.47	14.00	A1	0.74	12.88	A1	0.28	12.43	A1	0.79	13.71
7	A1	0.82	14.82	A1	0.86	13.74	A1	0.42	12.85	A1	0.79	14.50
8	A1	1.40	16.22	A1	1.10	14.85	A1	1.27	14.12	A1	0.99	15.48
9	A2	1.63	17.85	A2	0.98	15.83	A2	1.55	15.68	A2	1.28	16.77
10	A2	1.28	19.14	A2	2.45	18.28	A2	2.12	17.80	A2	1.87	18.64
11	A3	1.52	20.65	A3	2.09	20.37	A3	2.26	20.06	A3	2.66	21.30
12	A3	2.22	22.87	A3	3.68	24.05	A3	2.97	23.02	A3	3.65	24.95
13	A3	5.13	28.00	A3	5.15	29.20	A3	5.08	28.11	A3	3.85	28.80
14	P1	9.57	37.57	P1	8.59	37.79	P1	7.77	35.88	P1	8.09	36.88
15	P1	18.90	56.48	P1	16.81	54.60	P1	14.12	50.00	P1	16.37	53.25
16	P2	43.52	100.00	P2	45.40	100.00	P2	50.00	100.00	P2	46.75	100.00

Table 6.4.3J

Accuracy and Consistency of Classification Indices: Spek 9-12

Overall Indices	Accuracy	Consistency		Kappa (k)	
		0.558	0.576		0.424
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.943		0.245	
	A2	0.502		0.150	
	A3	0.699		0.081	
	P1	0.419		0.406	
	P2	-		0.566	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.979	0.012	0.009	0.970
	A2/A3	0.974	0.012	0.014	0.964
	A3/P1	0.956	0.011	0.033	0.933
	P1/P2	0.646	0.354	0.000	0.683

6.4.4 Writing 9-12

Figure 6.4.4A

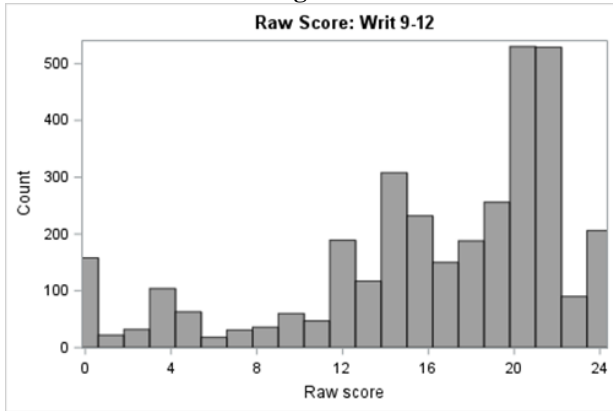


Table 6.4.4A

Raw Score Descriptive Statistics: Writ 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	849	0	24	15.90	6.46
10	809	0	24	16.30	6.34
11	703	0	24	16.08	6.15
12	1,005	0	24	15.94	6.53
Total	3,366	0	24	16.04	6.39

Figure 6.4.4B

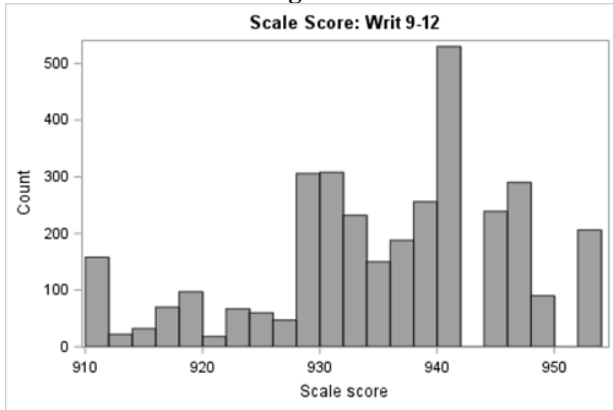


Table 6.4.4B

Scale Score Descriptive Statistics: Writ 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	849	910	953	934.67	10.71
10	809	910	953	935.48	10.62
11	703	910	953	934.99	10.24
12	1,005	910	953	934.86	10.79
Total	3,366	910	953	934.99	10.62

Figure 6.4.4C

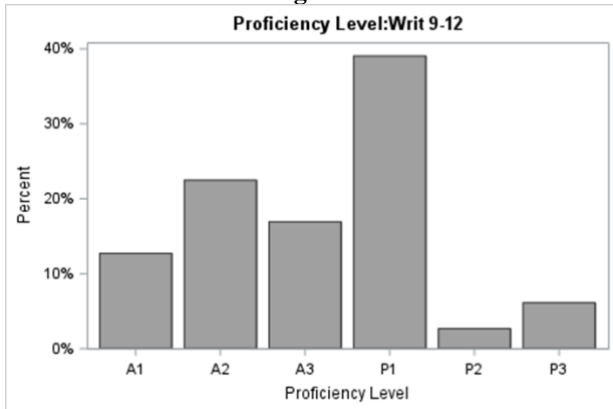


Table 6.4.4C

Proficiency Level Distribution: Writ 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	111	13.07	100	12.36	81	11.52	136	13.53	428	12.72
A2	188	22.14	170	21.01	169	24.04	230	22.89	757	22.49
A3	153	18.02	137	16.93	129	18.35	151	15.02	570	16.93
P1	323	38.04	323	39.93	262	37.27	407	40.50	1,315	39.07
P2	22	2.59	25	3.09	18	2.56	25	2.49	90	2.67
P3	52	6.12	54	6.67	44	6.26	56	5.57	206	6.12
Total	849	100.00	809	100.00	703	100.00	1,005	100.00	3,366	100.00

Table 6.4.4D

Equating Summary: Writ 9-12

No equating summary is presented because the Alternate ACCESS Series 502 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).

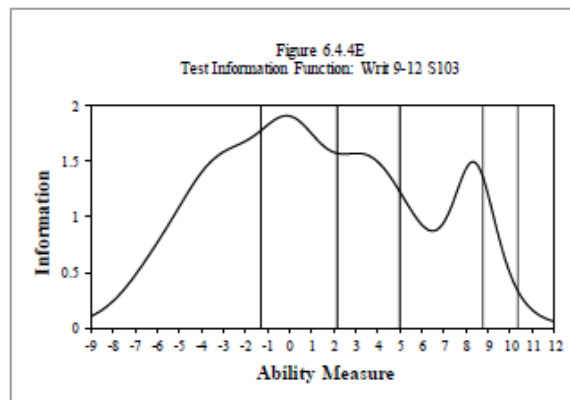
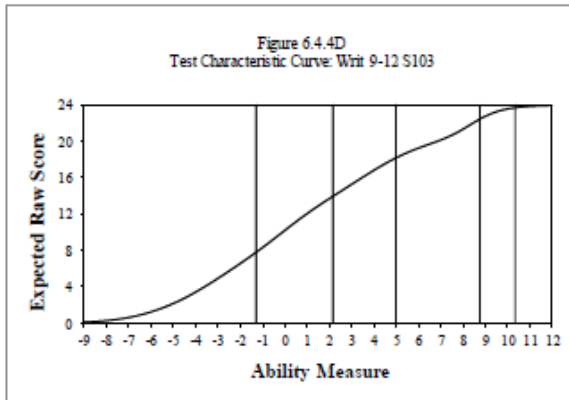


Table 6.4.4E

Reliability: Writ 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
3,366	10	0.939	1.578

Table 6.4.4F

Item Analysis Summary: Writ 9-12

Note: Item information is removed in the redacted version.

Table 6.4.4G

Complete Item Analysis: Writ 9-12

Note: Item information is removed in the redacted version.

Table 6.4.4H

Raw Score to Scale Score Conversion: Writ 9-12

Raw Score	Scale Score	SE Scaled	Low Bound	High Bound
0	910^	4.68	910.00^	912.75
1	912	2.90	910.00^	914.67
2	915	2.30	912.20	916.81
3	916	1.99	914.43	918.42
4	918	1.82	916.11	919.76
5	919	1.75	917.50	921.01
6	921	1.70	918.80	922.21
7	922	1.70	920.00	923.41
8	923	1.68	921.22	924.58
9	924	1.68	922.40	925.76
10	925	1.66	923.58	926.89
11	926	1.68	924.70	928.06
12	928	1.70	925.88	929.29
13	929	1.78	927.08	930.63
14	930	1.82	928.38	932.02
15	932	1.82	929.74	933.39
16	933	1.82	931.14	934.78
17	934	1.87	932.50	936.25
18	936	2.04	933.92	938.00
19	938	2.42	935.58	940.42
20	941	2.88	938.14	943.90
21	944	2.38	941.58	946.33
22	946	2.18	943.86	948.22
23	948*	2.59	945.68	950.86
24	950*	4.44	946.93	955.81

^ Truncated

* Adjusted for end of scale effect

Table 6.4.4I

Raw Score to Proficiency Level Conversion: Writ 9-12

Raw Score	Grade 9			Grade 10			Grade 11			Grade 12		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	5.54	5.54	A1	4.70	4.70	A1	3.13	3.13	A1	5.07	5.07
1	A1	0.82	6.36	A1	0.37	5.07	A1	0.85	3.98	A1	0.60	5.67
2	A1	0.82	7.18	A1	0.49	5.56	A1	1.14	5.12	A1	1.29	6.97
3	A1	2.59	9.78	A1	1.48	7.05	A1	2.42	7.54	A1	1.89	8.86
4	A1	0.59	10.37	A1	0.74	7.79	A1	1.14	8.68	A1	1.49	10.35
5	A1	1.53	11.90	A1	2.35	10.14	A1	1.85	10.53	A1	1.79	12.14
6	A1	0.24	12.13	A1	0.87	11.00	A1	0.43	10.95	A1	0.60	12.74
7	A1	0.94	13.07	A1	1.36	12.36	A1	0.57	11.52	A1	0.80	13.53
8	A2	0.94	14.02	A2	0.99	13.35	A2	1.85	13.37	A2	0.70	14.23
9	A2	0.47	14.49	A2	0.37	13.72	A2	0.28	13.66	A2	1.00	15.22
10	A2	0.71	15.19	A2	1.85	15.57	A2	1.00	14.65	A2	1.29	16.52
11	A2	1.06	16.25	A2	1.73	17.31	A2	1.56	16.22	A2	1.29	17.81
12	A2	5.54	21.79	A2	6.30	23.61	A2	5.83	22.05	A2	4.98	22.79
13	A2	4.12	25.91	A2	2.47	26.08	A2	3.41	25.46	A2	3.78	26.57
14	A2	9.31	35.22	A2	7.29	33.37	A2	10.10	35.56	A2	9.85	36.42
15	A3	2.71	37.93	A3	1.73	35.11	A3	2.42	37.98	A3	2.19	38.61
16	A3	5.30	43.23	A3	3.96	39.06	A3	4.41	42.39	A3	4.78	43.38
17	A3	4.95	48.17	A3	4.45	43.51	A3	5.55	47.94	A3	3.28	46.67
18	A3	5.06	53.24	A3	6.80	50.31	A3	5.97	53.91	A3	4.78	51.44
19	P1	9.31	62.54	P1	7.29	57.60	P1	6.54	60.46	P1	7.16	58.61
20	P1	14.61	77.15	P1	15.20	72.81	P1	17.21	77.67	P1	16.12	74.73
21	P1	6.36	83.51	P1	8.41	81.21	P1	6.54	84.21	P1	7.06	81.79
22	P1	7.77	91.28	P1	9.02	90.23	P1	6.97	91.18	P1	10.15	91.94
23	P2	2.59	93.88	P2	3.09	93.33	P2	2.56	93.74	P2	2.49	94.43
24	P3	6.12	100.00	P3	6.67	100.00	P3	6.26	100.00	P3	5.57	100.00

Table 6.4.4J

Accuracy and Consistency of Classification Indices: Writ 9-12

Overall Indices	Accuracy	Consistency		Kappa (k)	
		0.732	0.639		0.514
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.835		0.134	
	A2	0.763		0.302	
	A3	0.535		0.125	
	P1	0.752		0.753	
	P2	-		0.221	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.949	0.031	0.020	0.926
	A2/A3	0.926	0.025	0.049	0.899
	A3/P1	0.923	0.033	0.044	0.887
	P1/P2	0.932	0.068	0.000	0.909

6.4.5 Oral Language Composite 9-12

Figure 6.4.5A

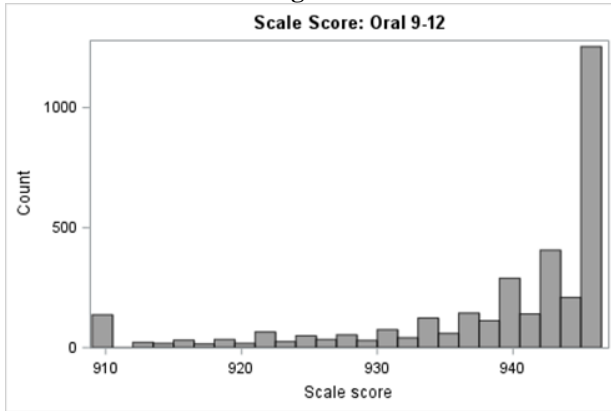


Table 6.4.5A

Scale Score Descriptive Statistics: Oral 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	853	910	946	938.11	9.89
10	815	910	946	938.43	9.65
11	704	910	946	939.00	9.35
12	1,012	910	946	938.54	9.86
Total	3,384	910	946	938.50	9.71

Figure 6.4.5B

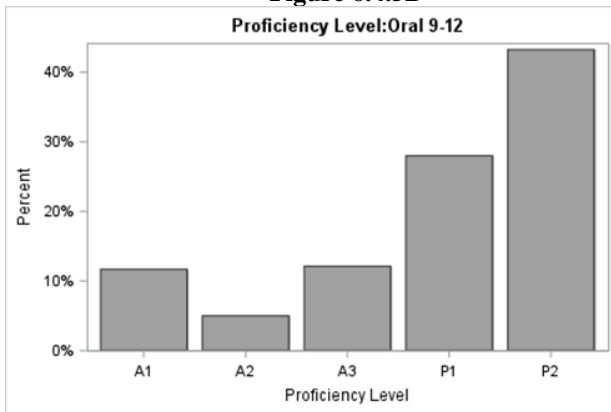


Table 6.4.5C

Proficiency Level Distribution: Oral 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	106	12.43	97	11.90	74	10.51	118	11.66	395	11.67
A2	43	5.04	38	4.66	35	4.97	53	5.24	169	4.99
A3	104	12.19	107	13.13	82	11.65	117	11.56	410	12.12
P1	259	30.36	219	26.87	193	27.41	276	27.27	947	27.98
P2	341	39.98	354	43.44	320	45.45	448	44.27	1,463	43.23
Total	853	100.00	815	100.00	704	100.00	1,012	100.00	3,384	100.00

Table 6.4.5D

n/a

Figure 6.4.5D

n/a

Figure 6.4.5E

n/a

Table 6.4.5E

Reliability: Oral 9-12

Component	Weight	Variance	Reliability
Listening	0.5	93.963	0.938
Speaking	0.5	130.299	0.964
Oral		94.377	0.972

*Variances from students who had results in all four domains

Table 6.4.5F

n/a

Table 6.4.5G

n/a

Table 6.4.5H

n/a

Table 6.4.5I

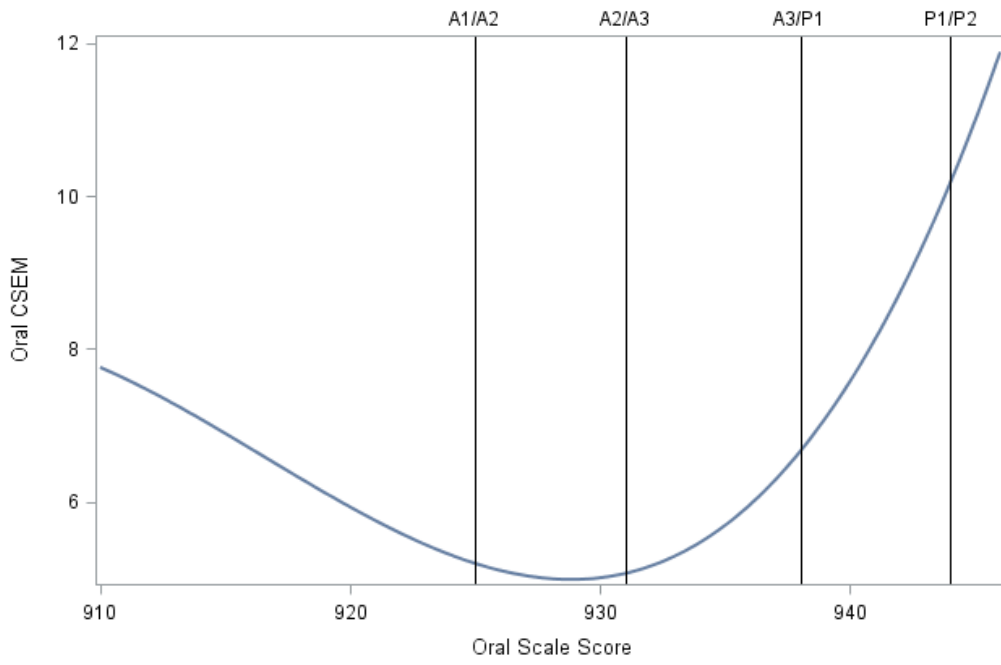
n/a

Table 6.4.5J

Accuracy and Consistency of Classification Indices: Oral 9-12

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.753	0.644		0.515	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.950		0.923	
	A2	0.589		0.463	
	A3	0.821		0.734	
	P1	0.626		0.441	
	P2	0.751		0.715	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.984	0.008	0.009	0.976
	A2/A3	0.975	0.015	0.010	0.965
	A3/P1	0.965	0.011	0.023	0.952
	P1/P2	0.829	0.057	0.114	0.749

Figure 6.4.5F CSEM for Oral Composite 9-12



6.4.6 Literacy Composite 9-12

Figure 6.4.6A

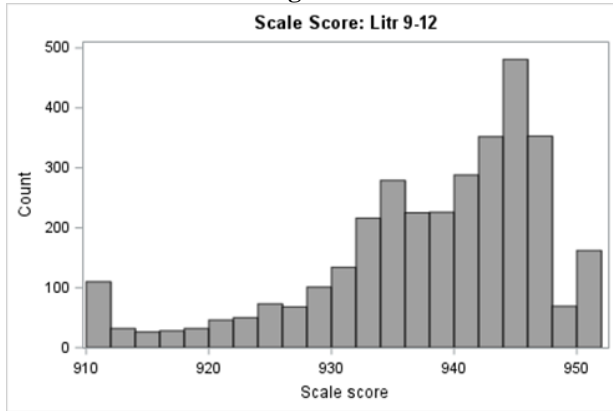


Table 6.4.6A

Scale Score Descriptive Statistics: Litr 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	844	910	951	936.96	9.83
10	807	910	951	937.75	9.47
11	698	910	951	937.66	9.15
12	1,002	910	951	937.28	9.83
Total	3,351	910	951	937.39	9.60

Figure 6.4.6B

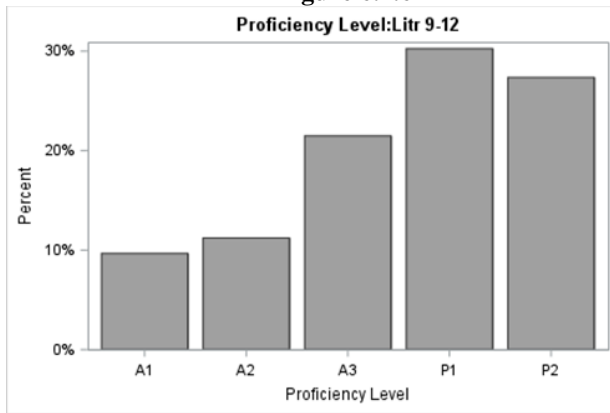


Table 6.4.6C

Proficiency Level Distribution: Litr 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	88	10.43	70	8.67	65	9.31	101	10.08	324	9.67
A2	93	11.02	96	11.90	74	10.60	113	11.28	376	11.22
A3	193	22.87	168	20.82	153	21.92	206	20.56	720	21.49
P1	266	31.52	250	30.98	204	29.23	294	29.34	1,014	30.26
P3	204	24.17	223	27.63	202	28.94	288	28.74	917	27.36
Total	844	100.00	807	100.00	698	100.00	1,002	100.00	3,351	100.00

Table 6.4.6D

n/a

Figure 6.4.6D

n/a

Figure 6.4.6E

n/a

Table 6.4.6E

Reliability: Litr 9-12

Component	Weight	Variance	Reliability
Reading	0.5	99.947	0.936
Writing	0.5	112.682	0.939
Literacy		92.244	0.964

*Variances from students who had results in all four domains

Table 6.4.6F

n/a

Table 6.4.6G

n/a

Table 6.4.6H

n/a

Table 6.4.6I

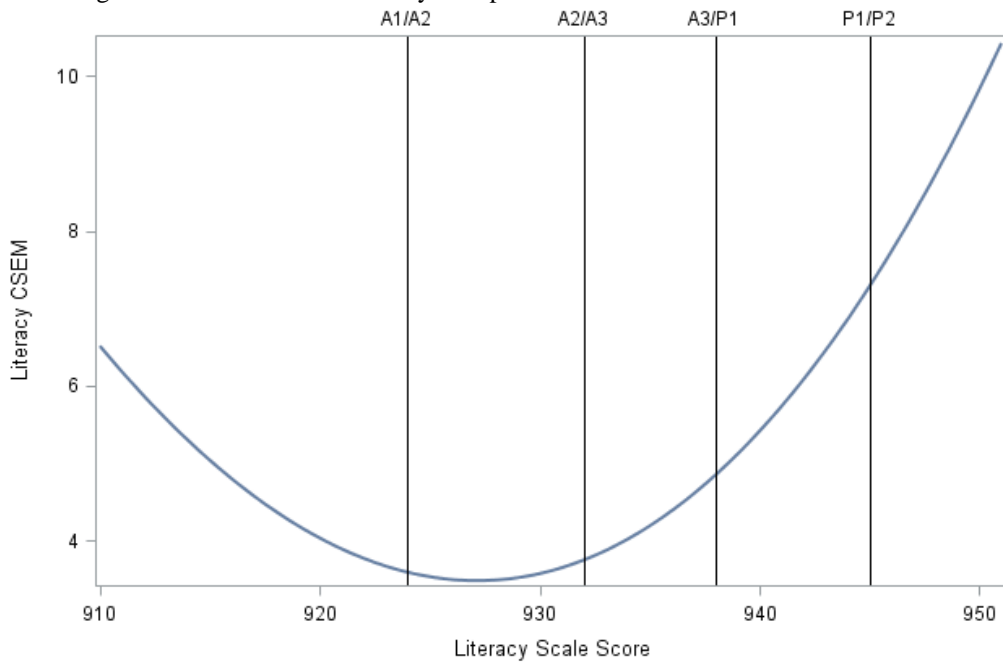
n/a

Table 6.4.6J

Accuracy and Consistency of Classification Indices: Litr 9-12

Overall Indices	Accuracy	Consistency		Kappa (k)	
		0.633	0.609		0.496
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.895		0.842	
	A2	0.747		0.644	
	A3	0.790		0.694	
	P1	0.492		0.495	
	P2	-		0.589	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.976	0.014	0.010	0.965
	A2/A3	0.952	0.026	0.022	0.933
	A3/P1	0.943	0.015	0.042	0.921
	P1/P2	0.762	0.238	0.000	0.786

Figure 6.4.6F CSEM for Literacy Composite 9-12



6.4.7 Comprehension Composite 9-12

Figure 6.4.7A

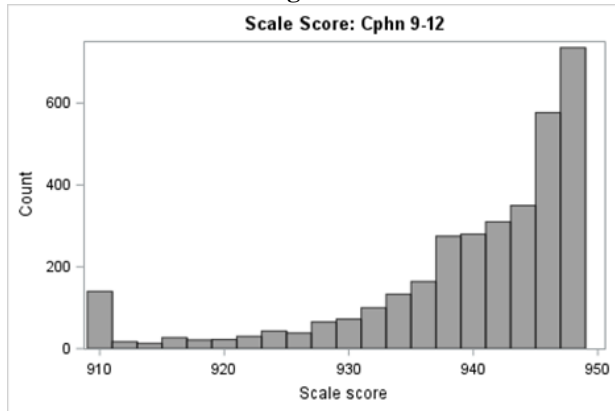


Table 6.4.7A

Scale Score Descriptive Statistics: Cphn 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	863	910	948	938.55	10.11
10	819	910	948	939.37	9.45
11	709	910	948	939.96	8.99
12	1,022	910	948	939.29	9.78
Total	3,413	910	948	939.26	9.63

Figure 6.4.7B

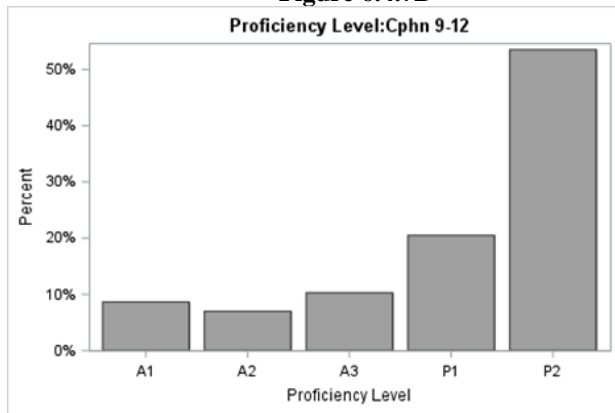


Table 6.4.7C

Proficiency Level Distribution: Cphn 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	83	9.62	64	7.81	54	7.62	94	9.20	295	8.64
A2	68	7.88	62	7.57	46	6.49	63	6.16	239	7.00
A3	93	10.78	89	10.87	68	9.59	101	9.88	351	10.28
P1	181	20.97	155	18.93	141	19.89	223	21.82	700	20.51
P2	438	50.75	449	54.82	400	56.42	541	52.94	1,828	53.56
Total	863	100.00	819	100.00	709	100.00	1,022	100.00	3,413	100.00

Table 6.4.7D

n/a

Figure 6.4.7D

n/a

Figure 6.4.7E
n/a

Table 6.4.7E

Reliability: Cphn 9-12

Component	Weight	Variance	Reliability
Listening	0.3	93.963	0.938
Reading	0.7	99.947	0.936
Comprehension		92.827	0.961

*Variances from students who had results in all four domains

Table 6.4.7F
n/a

Table 6.4.7G
n/a

Table 6.4.7H
n/a

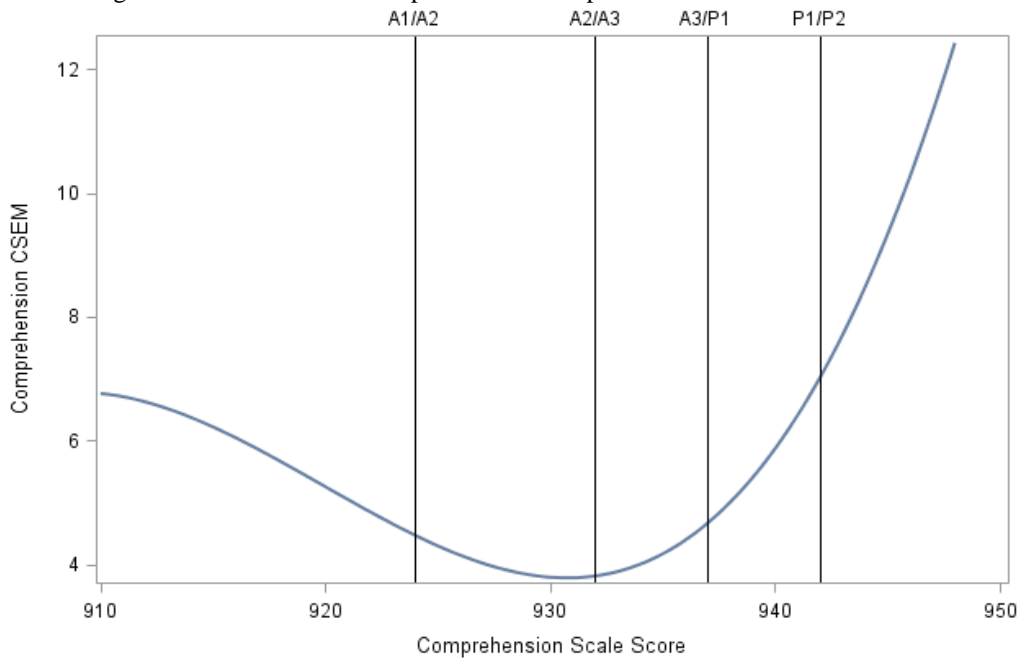
Table 6.4.7I
n/a

Table 6.4.7J

Accuracy and Consistency of Classification Indices: Cphn 9-12

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.778	0.699		0.560	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.918		0.874	
	A2	0.650		0.527	
	A3	0.661		0.538	
	P1	0.667		0.492	
	P2	0.826		0.798	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.983	0.009	0.008	0.976
	A2/A3	0.965	0.021	0.013	0.951
	A3/P1	0.945	0.023	0.032	0.925
	P1/P2	0.883	0.023	0.095	0.836

Figure 6.4.7F CSEM for Comprehension Composite 9-12



6.4.8 Overall Composite 9-12

Figure 6.4.8A

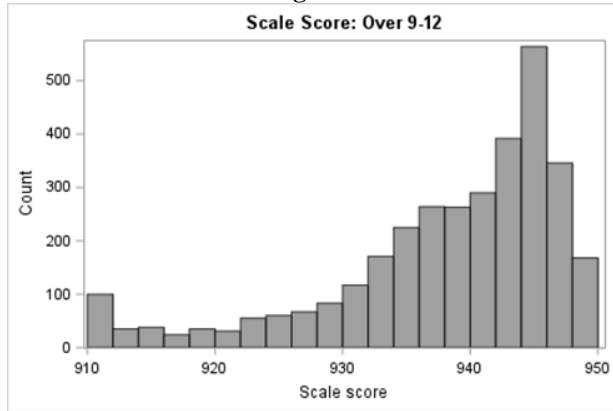


Table 6.4.8A

Scale Score Descriptive Statistics: Over 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	839	910	949	937.12	9.55
10	802	910	949	937.79	9.18
11	691	910	949	937.95	8.83
12	996	910	949	937.51	9.54
Total	3,328	910	949	937.57	9.31

Figure 6.4.8B

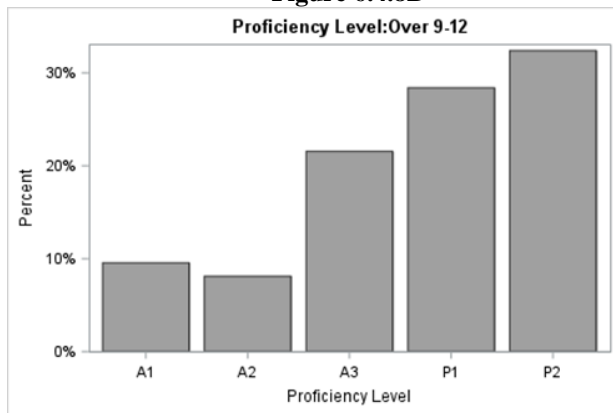


Table 6.4.8C

Proficiency Level Distribution: Over 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	90	10.73	70	8.73	61	8.83	97	9.74	318	9.56
A2	69	8.22	71	8.85	49	7.09	81	8.13	270	8.11
A3	181	21.57	171	21.32	151	21.85	214	21.49	717	21.54
P1	255	30.39	225	28.05	196	28.36	269	27.01	945	28.40
P2	244	29.08	265	33.04	234	33.86	335	33.63	1,078	32.39
Total	839	100.00	802	100.00	691	100.00	996	100.00	3,328	100.00

Table 6.4.8D

n/a

Figure 6.4.8D

n/a

Figure 6.4.8E

n/a

Table 6.4.8E

Reliability: Over 9-12

Component	Weight	Variance	Reliability
Listening	0.15	93.963	0.938
Reading	0.35	99.947	0.936
Speaking	0.15	130.299	0.964
Writing	0.35	112.682	0.939
Overall Composite		86.747	0.979

*Variances from students who had results in all four domains

Table 6.4.8F

n/a

Table 6.4.8G

n/a

Table 6.4.8H

n/a

Table 6.4.8I

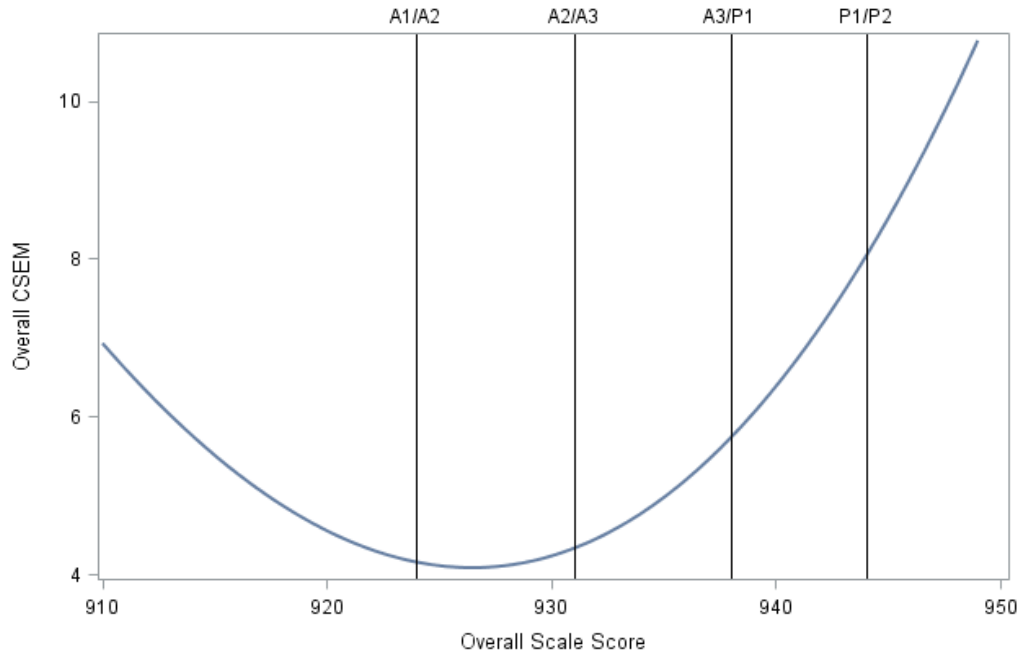
n/a

Table 6.4.8J

Accuracy and Consistency of Classification Indices: Over 9-12

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.594	0.653		0.550	
Conditional on Level	Level	Accuracy		Consistency	
	A1	0.938		0.907	
	A2	0.725		0.619	
	A3	0.894		0.840	
	P1	0.424		0.487	
	P2	1.780		0.645	
Indices at Cut Points	Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	A1/A2	0.985	0.008	0.007	0.979
	A2/A3	0.969	0.019	0.011	0.956
	A3/P1	0.957	0.011	0.033	0.940
	P1/P2	0.683	0.305	0.012	0.778

Figure 6.4.8F CSEM for Overall Composite 9-12



References

- Allen, N.L., Carlson, J.E., & Zalanak, C.A. (1999). *The NAEP 1996 technical report*. Washington, DC: National Center for Education Statistics.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, D.C.: American Psychological Association.
- Andrich D. (1978) A rating scale formulation for ordered response categories. *Psychometrika*, 43,561-573.
- Bachman, L. F & A. B. Palmer (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. UK: Oxford University Press.
- Brennan, R.L. (2004). BB-CLASS: a computer program that uses the beta-binomial model for classification consistency and accuracy. [Computer Software]. Iowa City, IA: CASMA.
- Center for Applied Linguistics. (2012a). *Alternate ACCESS for ELLs™ Standard Setting Study: Technical Brief*. Available at www.wida.wisc.edu.
- Center for Applied Linguistics (2013). *Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report*.
- Chapelle, C. A., Enright, M. & Jamieson, J. (Eds.) (2008). *Building a validity argument for the Test of English as a Foreign Language*. London: Routledge.
- Chapelle, C. A., Enright, M. K., & Jamieson, J. (2010). Does an Argument-Based Approach to Validity Make a Difference?. *Educational Measurement: Issues and Practice*, 29(1), 3-13.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.
- Cronbach, L.J., Schonemann, P., & McKie, D. (1965). Alpha coefficients for stratified-parallel tests. *Educational and Psychological Measurement*, 25, 291-312.
- Every Student Succeeds (ESSA) Act of 2015, Pub. L. No. 114-95, § 114 Stat. 1177 (2015-2016). Individuals With Disabilities Education Act, 20 U.S.C. § 1400 (2004).
- Kamata, A., Turhan, A., & Darandari, E. (2003, April). *Estimating reliability for multidimensional composite scale scores*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

- Kane, M. (2002). Validating high-stakes testing programs. *Educational measurement: Issues and practice*, 21(1), 31-41.
- Kane, M., & Case, S.M. (2004). The reliability and validity of weighted composite scores. *Applied Measurement in Education*, 17, 221-240.
- Kane, M. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73.
- Kenyon, D.M. (2006). *Development and Field Test of ACCESS for ELLs®*. (WIDA Consortium Technical Report No. 1).
- Lee, W., Hanson, B.A., & Brennan, R.L. (2002). Estimating consistency and accuracy indices for multiple classifications. *Applied Psychological Measurement*, 26, 412-432.
- Linacre, J.M. (2002, Autumn). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16, 878. Retrieved April 10, 2006, from www.rasch.org/rmt/rmt162f.htm
- Linacre, J.M. (2006). Winsteps (Version 3.60) [computer software]. Chicago, IL: Winsteps.com.
- Livingston, S.A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179-197.
- Llosa, L. (2008). Building and Supporting a Validity Argument for a Standards-Based Classroom Assessment of English Proficiency Based on Teacher Judgments. *Educational Measurement: Issues and Practice*, 27(3), 32-42.
- Meyer, J.P. (2014). jMetrik. [Computer Software]. <http://www.jmetrik.com>.
- Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2004). *A Brief Introduction to Evidence-Centered Design*. CSE Report 632. US Department of Education.
- Rudner, L. (2001, Spring). Informed test component weighting. *Educational Measurement: Issues and Practice*, 20:1, 16-19.
- WIDA Consortium. (2007). *English Language Proficiency Standards and Resource Guide, 2007 Edition, PreKindergarten through Grade 12*. Madison, Wisconsin: Board of Regents of the University of Wisconsin System.
- WIDA Consortium (2012a). *WIDA Alternate ACCESS for ELLs® Test Administration Manual*. Retrieved from <https://www.wida.wisc.edu>.
- WIDA Consortium (2013). *WIDA ACCESS for ELLs™ Test Administration Manual*. Retrieved from <https://www.wida.wisc.edu>.

- Wright, B.D., & Stone, M.H. (1979). *Best test design: Rasch measurement*. Chicago, IL: MESA Press.
- Young, M.J., & Yoon, B. (1998, April). *Estimating the consistency and accuracy of classifications in a standards-referenced assessment* (CSE Tech. Rep. 475). Los Angeles, CA: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education and Information Studies.
- Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. Holland & H. Wainer (Eds.), *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum.
- Zwick, R., Donoghue, J.R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, 30, 233-25.