World-Class Instructional Design and Assessment

Annual Technical Report for ALTERNATE ACCESS for ELLs® English Language Proficiency Test, Series 501, 2019-2020 Administration

Annual Technical Report No. 8

Prepared by:

Psychometrics team

WIDA

November 25, 2020

Table of Contents

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

1.1. Purpose of Alternate ACCESS for ELLs 1

1.2. Format of Alternate ACCESS 1

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 2

1.3. Test Development 3

1.3.1. Item Development 3
1.3.2. Field Test 4
1.3.3. Scaling 4
1.3.4. Standard Setting 5

1.4. Reporting of Results 8

1.4.1. Scale Scores 8
1.4.2. Language Proficiency Level Scores 8

1.5. Test Administration 9

1.5.1. Test Administrator Training 9
1.5.2. Test Security 9
1.5.3. Test Accommodations 9

1.6. Scoring 9

1.6.1. Listening and Reading 9
1.6.2. Writing 10
1.6.3. Speaking 11

2 An Assessment Use Argument for Alternate ACCESS for ELLs: Focus on Assessment Records 14

2.1 The Generic Validation Framework for Alternate ACCESS 15

2.2 Focus on Assessment Records 16

2.2.1 Breakdown of Claims for the Assessment Records Produced in the Alternate ACCESS for ELLs Assessment Program 18

2.3 Evidence for Assessment Records Claims of Alternate ACCESS for ELLs 19

2.4 Summary of Assessment Records Claims, Actions, and Evidence 25

2.5 Visual Guide to Tables and Figures 26

2.5.1 Chapter 4 Visual Guide to Tables and Figures 28
2.5.2 Chapter 6 Visual Guide to Tables and Figures 29

3. Descriptions of Student Results 30

3.1 Participation 30

3.2 Scale Score Results 31
3.2.1 Mean Scale Scores Across Domain and Composite Scores 31
3.2.2 Correlations 31
3.3 Proficiency Level Results 31
3.4 Participation by Disability 32
4. Student Results 33
4.1 Students excluded from Analysis 33
4.1.1 Out-of-grade-level Test Administration 33
4.2 Participation by Grade-level Cluster 34
4.2.1 Participation by Grade-level Cluster by State 34
4.2.2 Participation by Grade-level Cluster by Gender 35
4.2.3 Participation by Grade-level Cluster by Ethnicity 35
4.3 Participation by Grade 36
4.3.1 Participation by Grade by State 36
4.3.2 Participation by Grade by Gender 37
4.3.3 Participation by Grade by Ethnicity 37
4.4 Participation by Domain 38
4.4.1 Participation by Grade-level Cluster by Domain 38
4.4.2 Participation by Grade by Domain 38
4.5 Scale Scores by Domain and Composite 39
4.5.1 Mean Scale Scores by Domain and Composite 39
4.6 Scale Scores by Grade-level Cluster 41
4.6.1 Mean Scale Scores by Gender 41
4.6.2 Mean Scale Scores by Ethnicity 43
4.7 Scale Scores by Grade 47
4.7.1 Mean Scale Scores by Gender 47
4.7.2 Mean Scale Scores by Ethnicity 53
4.8 Correlations among Scale Scores by Grade-level Cluster 65
4.8.1 Correlations among Scale Scores: Grade-level Cluster 1-2 65
4.8.2 Correlations among Scale Scores: Grade-level Cluster 3-5 65
4.8.3 Correlations among Scale Scores: Grade-level Cluster 6-8 66
4.8.4 Correlations among Scale Scores: Grade-level Cluster 9-12 66
4.9 Proficiency Levels 67
4.9.1 Proficiency Level by Grade-level Cluster 67
4.9.2 Proficiency Level by Grade 70
4.10 Participation by Disability 74
4.10.1 Participation by Disability 74
5. Analyses of Test Forms: Overview 76
5.1 Background 76
5.1.1 Measurement Models Used 76
5.1.2 Sampling 77
5.1.3 Scaling 77

5.1.4 DIF Analyses 78
5.1.5 Reliability of Composites 79
5.1.6 Accuracy and Consistency of Classification 79

5.2 Descriptions 82
5.2.1 Raw Score Information (Figure A and Table A) 82
5.2.2 Scale Score Information (Figure B and Table B) 82
5.2.3 Proficiency Level Information (Figure C and Table C) 83
5.2.4 Equating Summary Table (Table D) 83
5.2.5 Reliability (Table E) 83
5.2.6 Test Characteristic Curve (Figure D) 84
5.2.7 Test Information Function (Figure E) 85
5.2.8 Item Analysis Summary (Table F) 85
5.2.9 Complete Item Analysis Table (Table G) 85
5.2.10 Complete Raw Score to Scale Score Conversion Chart (Table H) 86
5.2.11 Raw Score to Proficiency Level Score Conversion Table (Table I) 86
5.2.12 Accuracy and Consistency of Classification Table (Table J) 87
5.2.13 Conditional Standard Error of Measurement for Composite Figure (Figure F) 87

6. Analyses of Test Forms: Results 88

6.1 Grades: 1-2 88
6.1.1 Listening 1-2 88
6.1.2 Reading 1-2 94
6.1.3 Speaking 1-2 100
6.1.4 Writing 1-2 104
6.1.5 Oral Language Composite 1-2 110
6.1.6 Literacy Composite 1-2 113
6.1.7 Comprehension Composite 1-2 116
6.1.8 Overall Composite 1-2 119

6.2 Grades: 3-5 122
6.2.1 Listening 3-5 122
6.2.2 Reading 3-5 128
6.2.3 Speaking 3-5 134
6.2.4 Writing 3-5 139
6.2.5 Oral Language Composite 3-5 145
6.2.6 Literacy Composite 3-5 148
6.2.7 Comprehension Composite 3-5 151
6.2.8 Overall Composite 3-5 154

6.3 Grades: 6-8 157
6.3.1 Listening 6-8 157
6.3.2 Reading 6-8 163
6.3.3 Speaking 6-8 169
6.3.4 Writing 6-8 174

6.3.5 Oral Language Composite 6-8 180
6.3.6 Literacy Composite 6-8 183
6.3.7 Comprehension Composite 6-8 186
6.3.8 Overall Composite 6-8 189
6.4 Grades: 9-12 192
6.4.1 Listening 9-12 192
6.4.2 Reading 9-12 198
6.4.3 Speaking 9-12 204
6.4.4 Writing 9-12 209
6.4.5 Oral Language Composite 9-12 215
6.4.6 Literacy Composite 9-12 218
6.4.7 Comprehension Composite 9-12 221
6.4.8 Overall Composite 9-12 224
References 227

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 aligned 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 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 include[1]:

- To meet federal accountability requirements for assessment practice for ELLs and students with 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

> ELLs communicate information, ideas, and concepts necessary for academic success in the

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

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.

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

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

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.

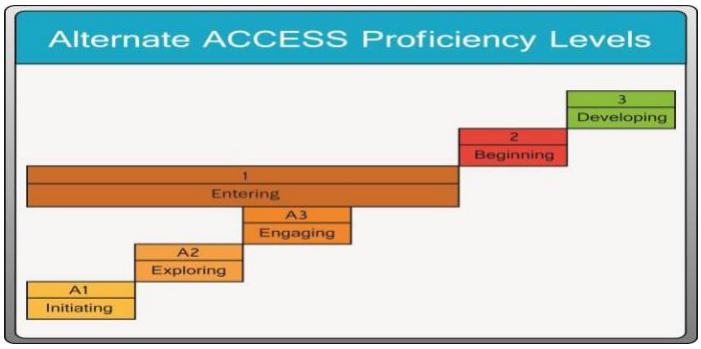


Figure 1.2.4A. Alternate ACCESS Proficiency Levels

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 to ensure that it allows students to demonstrate accomplishment of the targeted AMPI. (See the sample items at the WIDA website [https:/wida.wisc.edu/assess/alt-access] for examples.)

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 during the grant phase of test

development 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 AMPIs. 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 DRC, 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 to 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 the 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 drawings 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	Speech Features
2—Emerging	Phrases or short sentences. General language related to the task; groping for vocabulary when going beyond the highly familiar is evident. 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.
1—Entering	Single words or chunks of memorized oral language. General vocabulary from school setting and related to task. When using memorized language, is generally comprehensible; communication may be significantly impeded when going beyond the highly familiar.
A3—Engaging	Single words or chunks of mimicked oral language. Mimicked high frequency vocabulary words related to the task. When using mimicked language, is generally comprehensible; communication may be significantly impeded when going beyond mimicked language.
A2—Exploring	Single syllables or syllables of single words; speech is mimicked. Mimicked sounds and syllables of high frequency vocabulary words related to the task. Language is minimal.
A1—Initiating	Communicative vocalizations, which may be imitated (e.g., grunts). Indiscriminant sounds and syllables.

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.

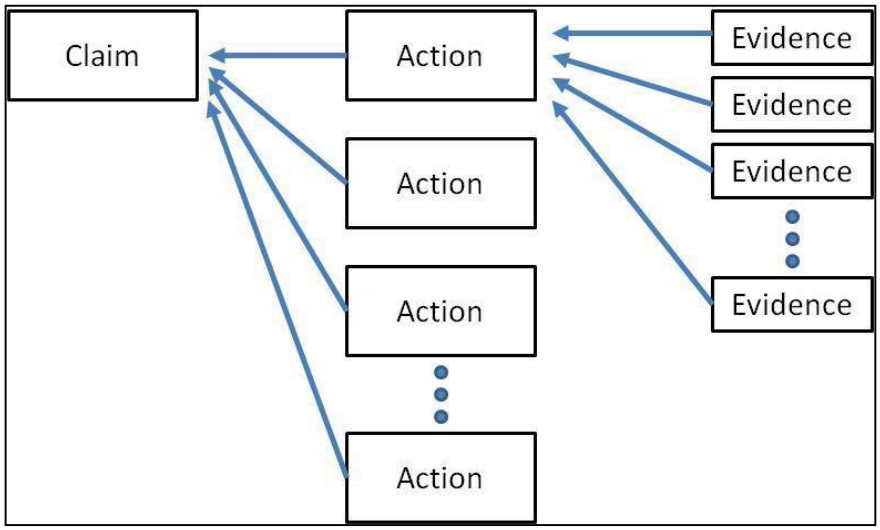


Figure 2A: General Argument Structure for Assessment Validation

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.

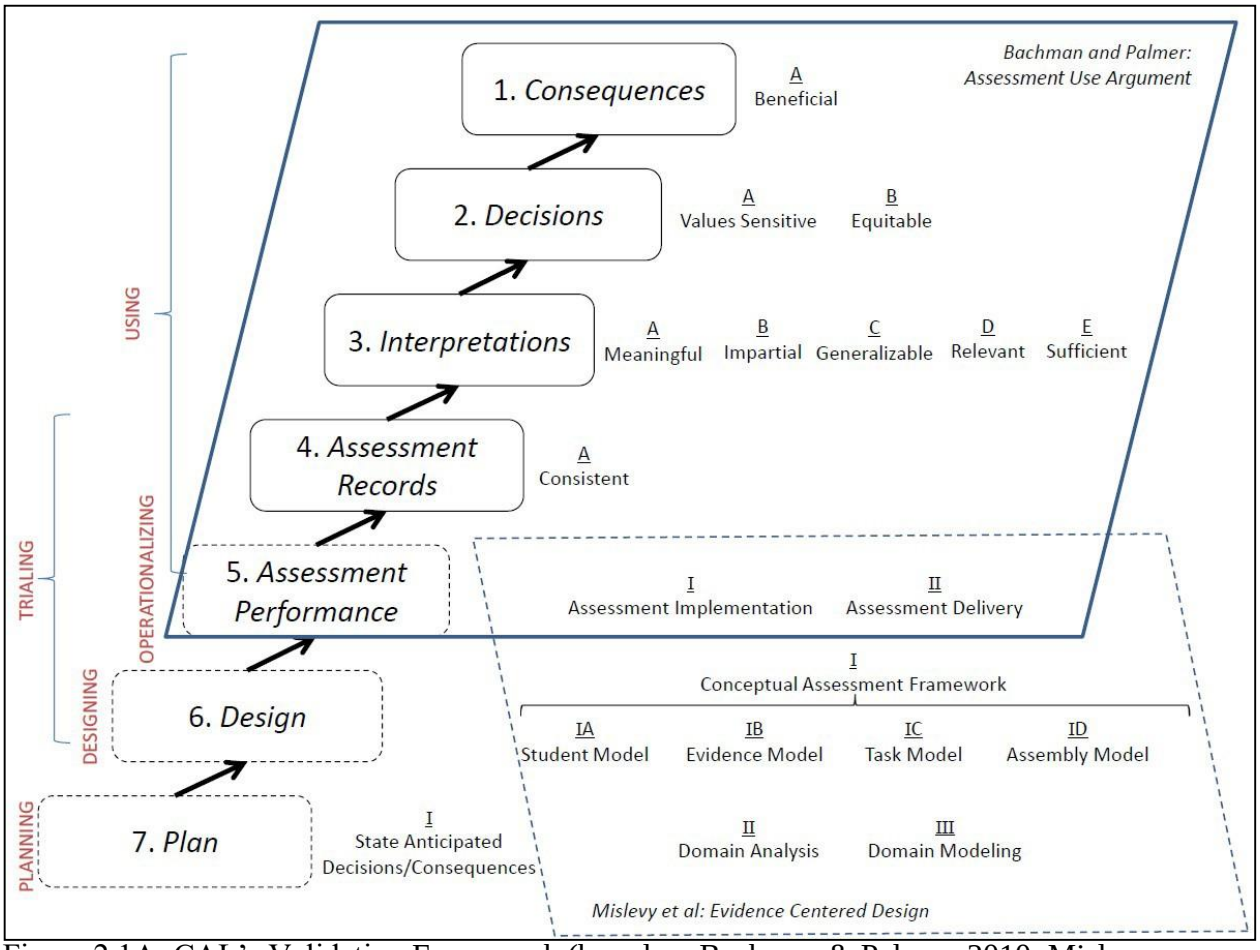


Figure 2.1A: CAL's Validation Framework (based on Bachman & Palmer, 2010; Mislevy, Almond, & Lukas, 2004)

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.

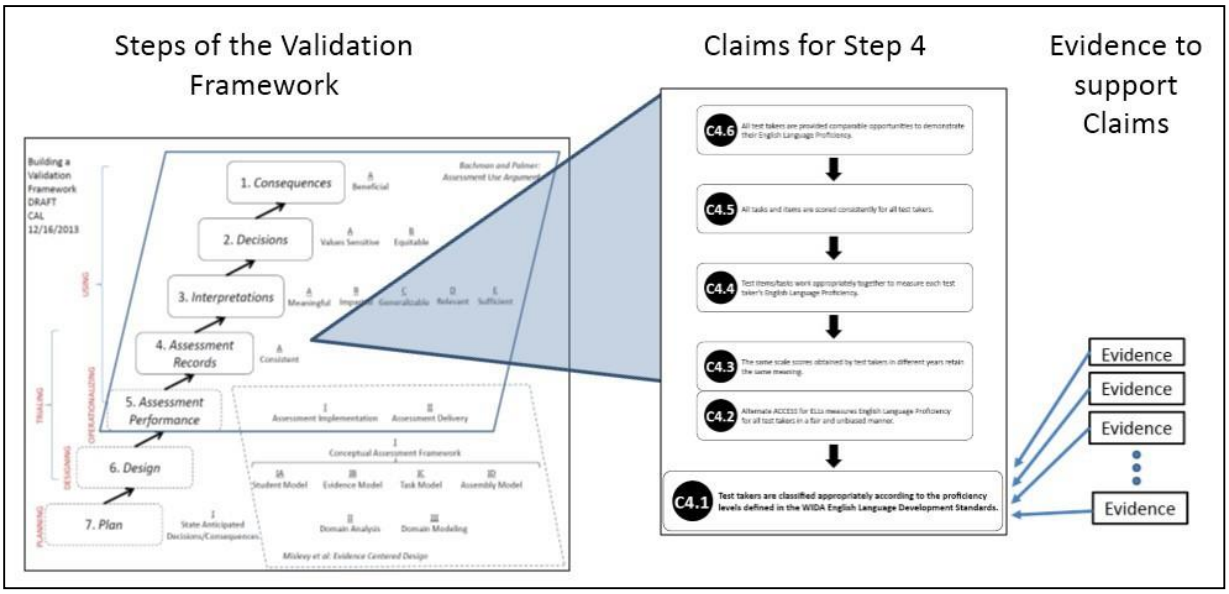


Figure 2.2A: Structure of the Argument-Based Approach Supporting Step 4 Contained in this Chapter

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.

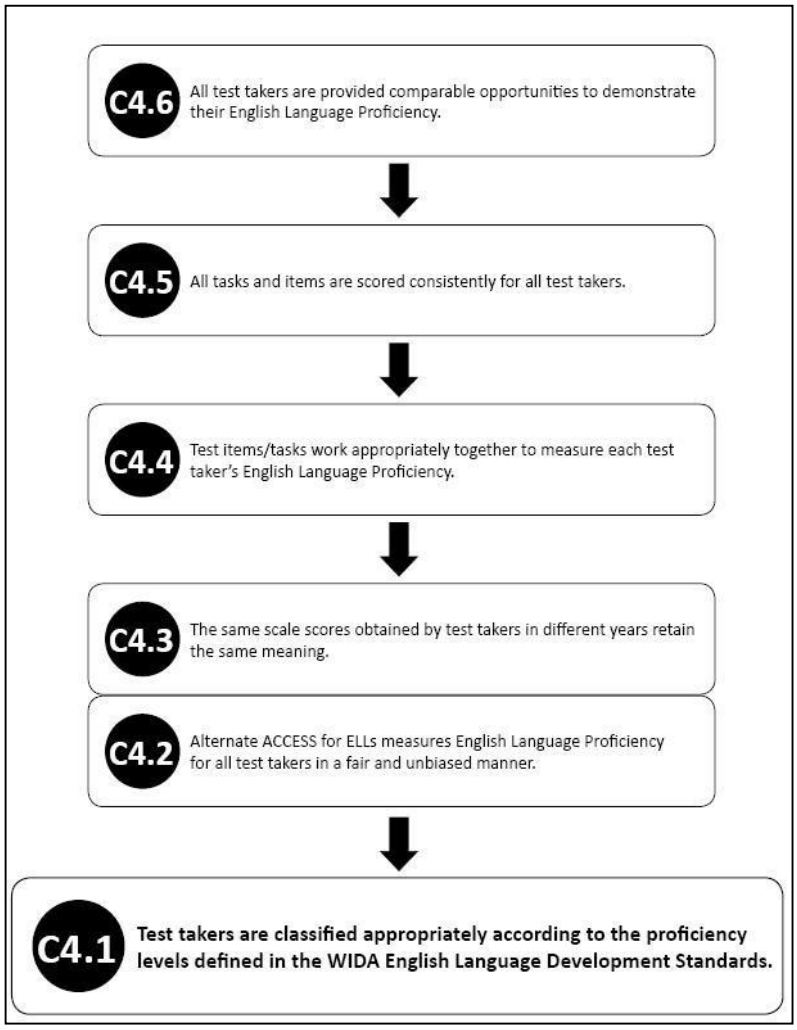


Figure 2.2.1A: Progression of Claims for Step 4: Assessment Records

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, S501).

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 is 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 (*Participation by Disability)*
	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 (*Reliability*)
	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 (*Reliability*)
	c. Analyses of Rasch model fit statistics are conducted to show that individual tasks perform appropriately.	c. Table 6G (*Complete Item Analysis*)
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 (*Equating Summary*)

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 over time.	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.	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. 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.	*a.* Table 6F *(Item Analysis Summary);* Table 6G *(Complete Item Analysis)* b. Chapter 5.1.4 (*DIF Items*)
1. Test takers are classified appropriately according to the Alternate proficiency levels defined in the WIDA English Language Development (ELD) Standards.	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.	*a.* Figure 6A *(Raw Scores)* & Table 6A *(Raw Score Descriptive Statistics)*; Figure 6B *(Scale Scores)* & Table 6B *(Scale Score Descriptive Statistics);* Figure 6C *(Proficiency Level)* & Table 6C *(Proficiency Level Distribution);* Table 6I *(Raw Score to Proficiency Level Score Conversion Chart)*; Figure 6D *(Test Characteristic Curve)*
	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	*b.* Figure 6B (*Scale Scores*) & Table 6B (*Scale Score Descriptive Statistics*); Figure 6C (*Proficiency Level*) & Table 6C (*Proficiency Level Distribution*); Figure 6D (*Test Characteristic Curve*
	c. For each test form, analyses are run to confirm that English Language Proficiency is measured with high precision at the pertinent cut points.	c. Figure 6E (*Test Information Function*); Table 6H (*Raw Score to Scale Score Conversion Chart*
	d. Classification and accuracy analyses are conducted by grade-level to confirm that proficiency level classifications are reliable for all domain and composite scores.	*d.* Table 6J (*Accuracy and Consistency of Classification Indices)*

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 2.5.1 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

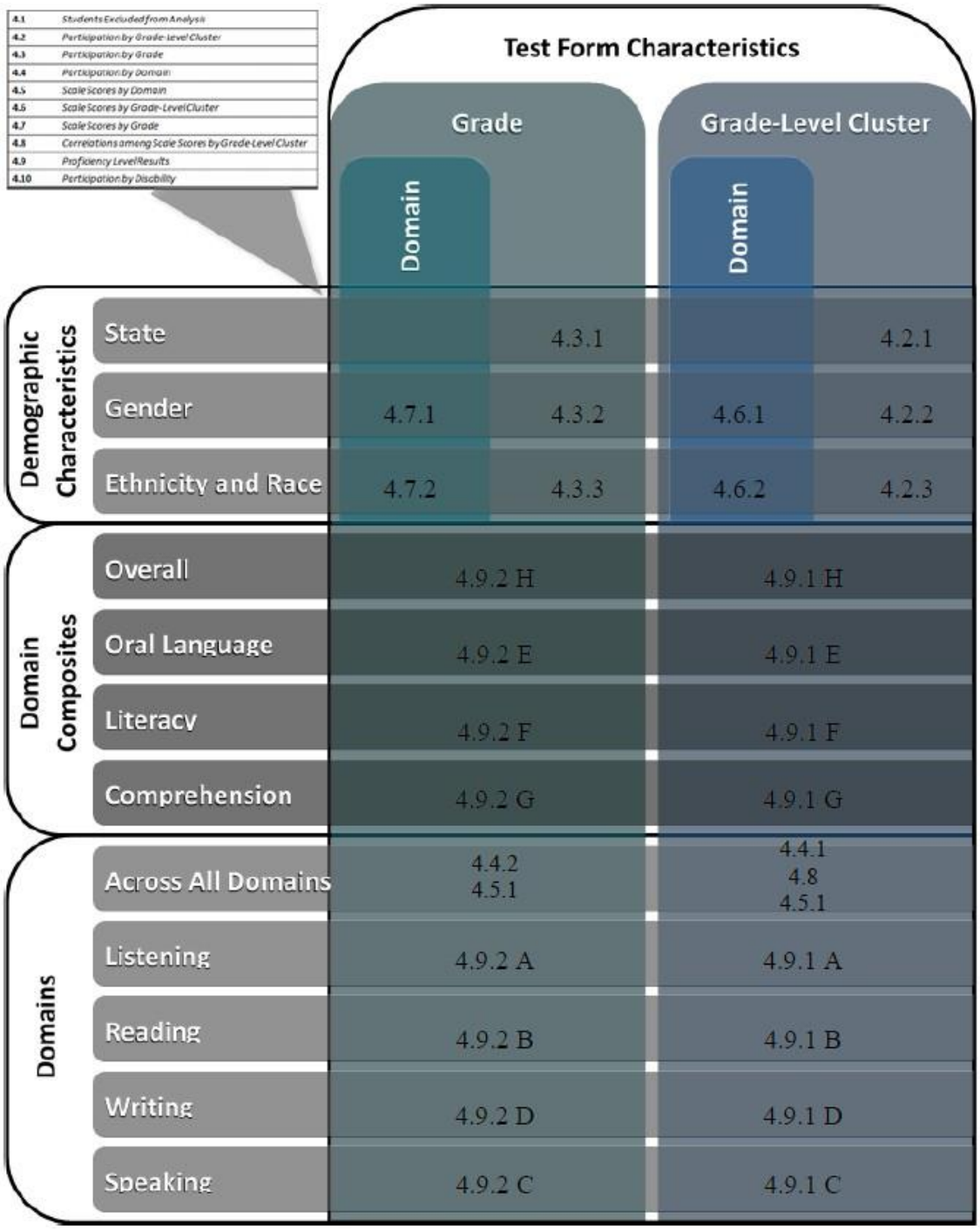


Figure 2.5.1 Chapter 4 Visual Guide to Tables and Figures

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*

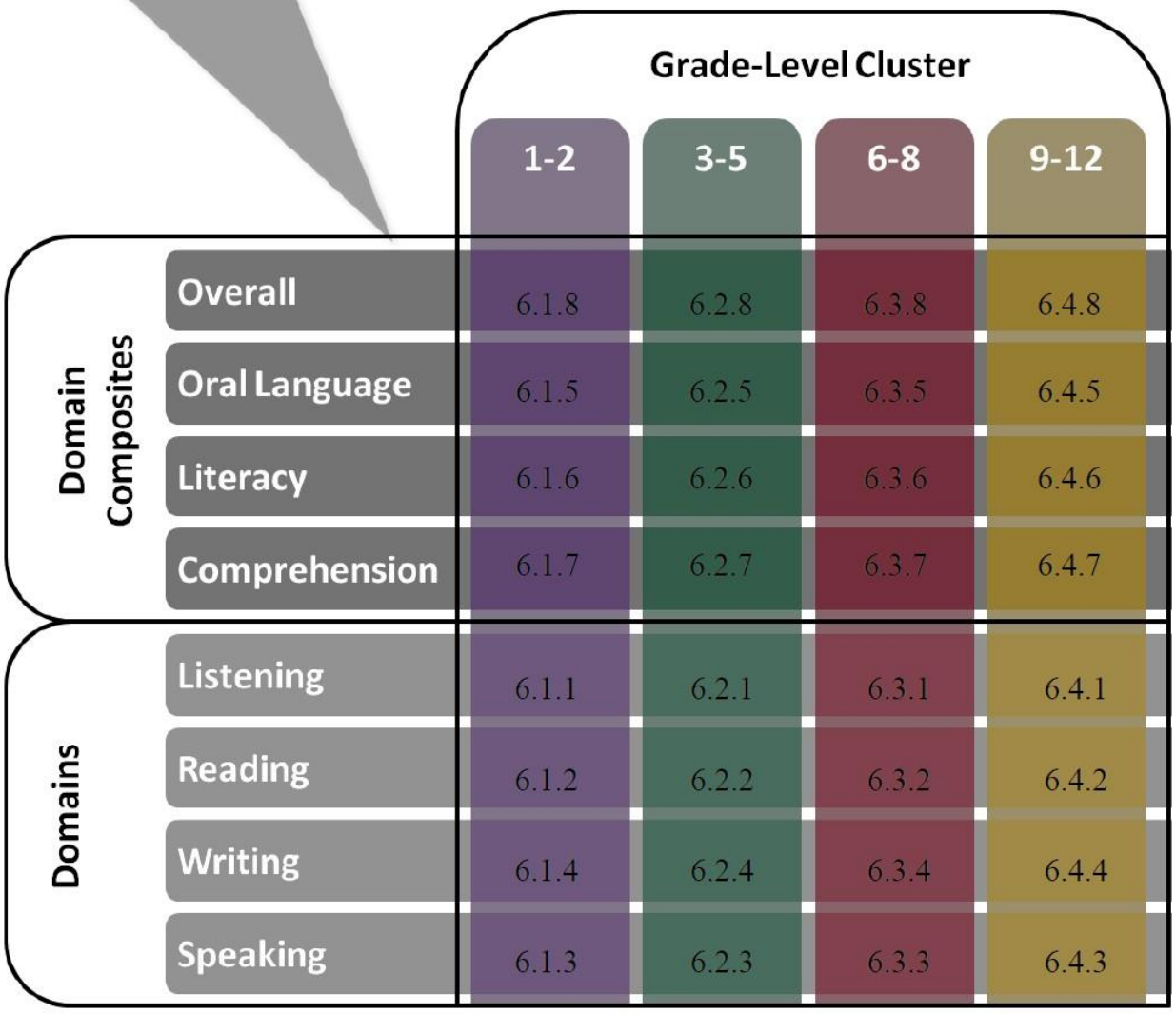


Figure 2.5.2 Chapter 6 Visual Guide to Tables and Figures

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 40 WIDA Consortium states/territories participated in the 2019-2020 Alternate ACCESS operational administration.

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, 40 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 40 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 level[3] 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 (SPD)
- 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

	Cluster				
Grade	**1-2**	**3-5**	**6-8**	**9-12**	**Total**
1		1	0	0	1
2		5	0	0	5
3	7		0	0	7
4	1		0	0	1
5	0		5	0	5
6	0	8		0	8
7	0	0		1	1
8	0	0		8	8
9	0	0	3		3
10	0	0	0		0
11	0	0	0		0
12	0	0	1		1
Total	8	14	9	9	40

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	10	17	29	37	93
AL	54	94	91	55	294
BI	4	12	19	5	40
CO	171	300	269	314	1,054
DC	2	8	12	19	41
DD	2	12	2	12	28
DE	8	11	10	5	34
FL	353	399	195	183	1,130
GA	286	424	377	288	1,375
HI	43	78	74	60	255
ID	23	61	52	34	170
IL	839	1,157	976	1,140	4,112
IN	156	270	300	372	1,098
KY	48	87	90	95	320
MA	347	451	370	423	1,591
MD	128	200	218	190	736
ME	15	12	18	11	56
MI	160	269	224	291	944
MN	233	301	192	260	986
MO	48	69	54	42	213
MP	0	1	0	0	1
MT	5	11	4	5	25
NC	247	492	385	418	1,542
ND	6	5	7	8	26
NH	7	9	12	4	32
NJ	66	49	42	24	181
NM	88	138	114	131	471
NV	168	258	311	320	1,057
OK	124	211	156	156	647
PA	227	428	290	365	1,310
RI	21	63	30	52	166
SC	98	86	86	81	351
SD	9	12	10	22	53
TN	34	75	83	72	264
UT	84	135	123	148	490
VA	433	532	495	534	1,994
VI	0	0	2	0	2
VT	8	1	5	6	20
WI	72	140	147	217	576
WY	11	12	11	20	54
Total	4,638	6,890	5,885	6,419	23,832

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	% within Cluster	
1-2	1,286	27.73	3,210	69.21	142	3.06	4,638
3-5	2,206	32.02	4,505	65.38	179	2.60	6,890
6-8	2,068	35.14	3,668	62.33	149	2.53	5,885
9-12	2,359	36.75	3,877	60.40	183	2.85	6,419
Total	7,919	33.23	15,260	64.03	653	2.74	23,832

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	2,675	57.68	1,621	34.95	342	7.37	4,638
3-5	4,156	60.32	2,145	31.13	589	8.55	6,890
6-8	3,846	65.35	1,634	27.77	405	6.88	5,885
9-12	3,893	60.65	1,970	30.69	556	8.66	6,419
Total	14,570	61.14	7,370	30.92	1,892	7.94	23,832

4.3 Participation by Grade

4.3.1 Participation by Grade by State

Table 4.3.1

Participation by Grade by State

State	Grade												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	Total
AK	2	8	7	5	5	12	8	9	7	2	6	22	93
AL	24	30	32	32	30	36	33	22	16	13	9	17	294
BI	0	4	4	3	5	9	4	6	4	0	0	1	40
CO	84	87	98	107	95	97	90	82	75	79	76	84	1,054
DC	1	1	3	1	4	7	3	2	14	0	2	3	41
DD	1	1	5	5	2	1	0	1	3	1	8	0	28
DE	5	3	3	2	6	5	2	3	3	1	1	0	34
FL	184	169	163	136	100	74	71	50	46	37	48	52	1,130
GA	152	134	124	154	146	130	137	110	97	79	59	53	1,375
HI	22	21	32	28	18	21	32	21	10	13	16	21	255
ID	12	11	22	24	15	21	19	12	7	14	8	5	170
IL	429	410	387	392	378	340	320	316	261	236	212	431	4,112
IN	80	76	83	80	107	97	89	114	99	82	67	124	1,098
KY	26	22	31	31	25	32	23	35	31	24	23	17	320
MA	196	151	153	161	137	159	113	98	117	91	103	112	1,591
MD	60	68	69	63	68	76	76	66	63	27	55	45	736
ME	8	7	5	3	4	7	4	7	4	3	2	2	56
MI	88	72	89	94	86	69	78	77	94	69	68	60	944
MN	119	114	92	112	97	64	70	58	81	47	47	85	986
MO	24	24	15	29	25	23	16	15	10	12	10	10	213
MP	0	0	1	0	0	0	0	0	0	0	0	0	1
MT	3	2	3	3	5	1	0	3	1	2	0	2	25
NC	132	115	157	163	172	142	121	122	103	93	88	134	1,542
ND	3	3	2	0	3	3	3	1	2	2	2	2	26
NH	3	4	3	5	1	5	2	5	1	0	1	2	32
NJ	42	24	18	17	14	22	9	11	4	6	8	6	181
NM	42	46	47	51	40	40	43	31	37	31	25	38	471
NV	81	87	90	92	76	104	101	106	90	91	62	77	1,057
OK	58	66	74	71	66	58	50	48	40	38	36	42	647
PA	103	124	138	162	128	108	86	96	98	88	78	101	1,310
RI	14	7	20	25	18	13	10	7	7	15	14	16	166
SC	50	48	33	33	20	30	30	26	18	23	23	17	351
SD	5	4	4	4	4	2	5	3	7	3	7	5	53
TN	10	24	19	31	25	31	18	34	26	20	11	15	264
UT	49	35	38	50	47	38	36	49	38	34	41	35	490
VA	230	203	182	156	194	169	170	156	139	111	121	163	1,994
VI	0	0	0	0	0	0	2	0	0	0	0	0	2
VT	2	6	1	0	0	1	2	2	1	2	1	2	20
WI	31	41	49	41	50	51	41	55	40	38	59	80	576
WY	4	7	6	1	5	1	4	6	9	5	4	2	54
Total	2,379	2,259	2,302	2,367	2,221	2,099	1,921	1,865	1,703	1,432	1,401	1,883	23,832

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	627	26.36	1,674	70.37	78	3.28	2,379
2	659	29.17	1,536	67.99	64	2.83	2,259
3	732	31.80	1,505	65.38	65	2.82	2,302
4	751	31.73	1,558	65.82	58	2.45	2,367
5	723	32.55	1,442	64.93	56	2.52	2,221
6	727	34.64	1,310	62.41	62	2.95	2,099
7	654	34.04	1,213	63.14	54	2.81	1,921
8	687	36.84	1,145	61.39	33	1.77	1,865
9	605	35.53	1,030	60.48	68	3.99	1,703
10	506	35.34	887	61.94	39	2.72	1,432
11	546	38.97	819	58.46	36	2.57	1,401
12	702	37.28	1,141	60.59	40	2.12	1,883
Total	7,919	33.23	15,260	64.03	653	2.74	23,832

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	1,361	57.21	858	36.07	160	6.73	2,379
2	1,314	58.17	763	33.78	182	8.06	2,259
3	1,385	60.17	719	31.23	198	8.60	2,302
4	1,405	59.36	747	31.56	215	9.08	2,367
5	1,366	61.50	679	30.57	176	7.92	2,221
6	1,372	65.36	578	27.54	149	7.10	2,099
7	1,268	66.01	526	27.38	127	6.61	1,921
8	1,206	64.66	530	28.42	129	6.92	1,865
9	1,019	59.84	533	31.30	151	8.87	1,703
10	875	61.10	443	30.94	114	7.96	1,432
11	839	59.89	433	30.91	129	9.21	1,401
12	1,160	61.60	561	29.79	162	8.60	1,883
Total	14,570	61.14	7,370	30.92	1,892	7.94	23,832

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	4,598	4,597	4,594	4,625
3-5	6,839	6,841	6,832	6,850
6-8	5,840	5,841	5,840	5,842
9-12	6,345	6,346	6,346	6,344
Total	23,622	23,625	23,612	23,661

4.4.2 Participation by Grade by Domain

Table 4.4.2
Participation by Grade by Domain

Grade	Domain			
	Listening	**Reading**	**Speaking**	**Writing**
1	2,355	2,355	2,351	2,368
2	2,243	2,242	2,243	2,257
3	2,284	2,285	2,283	2,299
4	2,349	2,351	2,347	2,348
5	2,206	2,205	2,202	2,203
6	2,086	2,088	2,089	2,089
7	1,904	1,903	1,902	1,902
8	1,850	1,850	1,849	1,851
9	1,686	1,687	1,687	1,689
10	1,418	1,418	1,419	1,419
11	1,379	1,380	1,381	1,380
12	1,862	1,861	1,859	1,856
Total	23,622	23,625	23,612	23,661

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	931.23	11.13	2,355	933.75	10.33	2,243	932.46	10.82	4,598
Reading	931.44	12.94	2,355	934.48	12.50	2,242	932.92	12.82	4,597
Speaking	931.60	14.51	2,351	934.46	14.34	2,243	933.00	14.50	4,594
Writing	926.56	11.24	2,368	929.08	11.37	2,257	927.79	11.37	4,625
Oral	931.67	12.06	2,351	934.37	11.66	2,241	932.99	11.95	4,592
Literacy	929.28	11.30	2,354	932.07	11.15	2,242	930.64	11.31	4,596
Comprehension	931.44	12.06	2,354	934.33	11.52	2,240	932.85	11.88	4,594
Overall	929.75	11.12	2,351	932.50	10.86	2,238	931.09	11.08	4,589

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	935.82	10.51	2,284	937.21	10.52	2,349	938.23	10.01	2,206	937.08	10.40	6,839
Reading	934.57	10.85	2,285	936.07	10.84	2,351	937.30	10.62	2,205	935.97	10.83	6,841
Speaking	935.07	13.51	2,283	936.31	13.22	2,347	937.02	13.14	2,202	936.12	13.31	6,832
Writing	930.66	11.58	2,299	932.13	11.85	2,348	933.62	12.13	2,203	932.12	11.91	6,850
Oral	935.58	11.16	2,281	936.88	11.12	2,345	937.75	10.85	2,201	936.73	11.08	6,827
Literacy	932.91	10.61	2,285	934.40	10.75	2,348	935.74	10.81	2,203	934.33	10.78	6,836
Comprehension	934.98	10.44	2,283	936.44	10.47	2,349	937.60	10.14	2,204	936.32	10.41	6,836
Overall	933.49	10.45	2,280	934.93	10.55	2,344	936.13	10.51	2,201	934.84	10.55	6,825

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	937.14	10.69	2,086	937.94	10.35	1,904	938.58	10.16	1,850	937.86	10.43	5,840
Reading	937.31	11.87	2,088	938.23	11.84	1,903	939.03	11.71	1,850	938.15	11.83	5,841
Speaking	936.08	13.17	2,089	936.74	13.07	1,902	937.03	12.98	1,849	936.60	13.08	5,840
Writing	932.30	10.59	2,089	932.65	10.80	1,902	933.41	11.13	1,851	932.76	10.84	5,842
Oral	936.96	11.45	2,085	937.72	11.23	1,902	938.18	11.01	1,848	937.59	11.25	5,835
Literacy	935.04	10.60	2,088	935.68	10.74	1,902	936.47	10.83	1,850	935.70	10.73	5,840
Comprehension	937.27	11.19	2,086	938.19	11.11	1,903	938.93	10.97	1,849	938.09	11.12	5,838
Overall	935.38	10.49	2,085	936.06	10.56	1,900	936.73	10.56	1,847	936.03	10.55	5,832

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	938.19	10.16	1,686	939.08	9.81	1,418	939.08	9.78	1,379	938.40	10.50	1,862	938.64	10.11	6,345
Reading	938.02	10.82	1,687	938.71	10.49	1,418	938.90	10.47	1,380	937.99	11.14	1,861	938.36	10.77	6,346
Speaking	936.28	12.24	1,687	936.82	12.04	1,419	937.04	11.93	1,381	936.21	12.27	1,859	936.55	12.14	6,346
Writing	934.10	11.04	1,689	934.26	11.01	1,419	934.80	11.07	1,380	934.05	11.28	1,856	934.27	11.11	6,344
Oral	937.39	10.52	1,685	938.09	10.24	1,418	938.21	10.22	1,379	937.46	10.74	1,859	937.75	10.46	6,341
Literacy	936.27	10.32	1,687	936.69	10.11	1,418	937.06	10.19	1,380	936.23	10.57	1,856	936.52	10.32	6,341
Comprehension	938.16	10.38	1,686	938.91	10.07	1,418	939.03	10.06	1,379	938.20	10.77	1,861	938.53	10.36	6,344
Overall	936.44	10.12	1,685	936.95	9.85	1,418	937.22	9.94	1,379	936.43	10.37	1,855	936.72	10.10	6,337

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	931.76	10.88	1,274	932.64	10.82	3,183	934.65	9.81	141
Reading	931.77	12.50	1,274	933.34	12.97	3,182	933.94	11.66	141
Speaking	932.08	14.43	1,273	933.35	14.52	3,180	933.24	14.21	141
Writing	926.85	11.11	1,282	928.16	11.50	3,201	927.91	10.37	142
Oral	932.18	11.96	1,272	933.26	11.96	3,179	934.19	11.14	141
Literacy	929.60	11.03	1,273	931.03	11.44	3,182	931.21	10.28	141
Comprehension	931.83	11.68	1,273	933.19	11.99	3,180	934.24	10.65	141
Overall	930.14	10.92	1,271	931.44	11.16	3,177	931.82	10.02	141

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	937.01	10.24	2,183	937.06	10.50	4,478	938.28	9.74	178
Reading	935.53	10.51	2,184	936.18	10.99	4,478	936.06	10.40	179
Speaking	935.58	13.45	2,185	936.38	13.27	4,468	936.47	12.45	179
Writing	931.31	11.71	2,190	932.57	12.02	4,481	930.75	10.68	179
Oral	936.42	11.07	2,182	936.85	11.11	4,467	937.46	10.39	178
Literacy	933.70	10.56	2,184	934.67	10.91	4,473	933.68	9.89	179
Comprehension	936.00	10.14	2,181	936.47	10.55	4,477	936.74	9.90	178
Overall	934.29	10.40	2,181	935.11	10.65	4,466	934.58	9.72	178

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	937.74	10.43	2,051	937.94	10.41	3,641	937.46	10.83	148
Reading	937.94	11.62	2,051	938.32	11.92	3,642	936.93	12.16	148
Speaking	936.30	13.13	2,048	936.86	13.00	3,644	934.07	13.94	148
Writing	932.54	11.03	2,051	933.01	10.74	3,643	929.76	10.18	148
Oral	937.38	11.24	2,048	937.77	11.22	3,639	936.11	11.79	148
Literacy	935.49	10.75	2,051	935.90	10.72	3,641	933.57	10.57	148
Comprehension	937.92	10.99	2,050	938.23	11.17	3,640	937.11	11.49	148
Overall	935.83	10.53	2,047	936.22	10.54	3,637	934.11	10.62	148

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	938.70	10.10	2,330	938.70	10.03	3,834	936.75	11.52	181
Reading	938.33	10.80	2,331	938.51	10.67	3,834	935.50	12.13	181
Speaking	936.55	12.06	2,330	936.68	12.11	3,834	933.76	13.23	182
Writing	934.36	11.26	2,327	934.34	10.97	3,835	931.70	11.78	182
Oral	937.78	10.43	2,328	937.84	10.40	3,832	935.38	11.79	181
Literacy	936.55	10.44	2,327	936.64	10.18	3,833	933.78	11.45	181
Comprehension	938.52	10.40	2,330	938.66	10.26	3,833	935.97	11.67	181
Overall	936.74	10.19	2,325	936.83	9.97	3,831	934.08	11.28	181

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	931.70	933.04	932.65	928.90	932.44	931.20	932.70	931.32
	Std. Dev.	10.63	13.75	14.36	12.07	11.81	12.07	12.48	11.57
	N	682	682	682	686	682	682	682	682
Non-Hispanic Pacific Islander	Mean	930.33	929.26	930.56	924.22	930.67	927.07	929.67	927.89
	Std. Dev.	10.13	12.64	13.20	9.66	10.59	10.46	11.68	10.05
	N	27	27	27	27	27	27	27	27
Non-Hispanic Black	Mean	931.81	932.36	933.82	927.29	933.07	930.08	932.27	930.77
	Std. Dev.	11.17	12.95	14.60	11.44	12.29	11.46	12.08	11.30
	N	366	366	366	367	366	366	366	366
Hispanic (Of Any Race)	Mean	932.57	932.81	932.68	927.33	932.89	930.37	932.80	930.87
	Std. Dev.	10.86	12.57	14.63	11.13	12.02	11.07	11.73	10.95
	N	2,649	2,649	2,645	2,667	2,643	2,648	2,646	2,641
Non-Hispanic American Indian	Mean	932.32	934.14	930.64	925.46	931.71	930.11	933.68	930.36
	Std. Dev.	12.29	12.22	15.47	11.27	13.28	10.93	11.81	11.22
	N	28	28	28	28	28	28	28	28
Non-Hispanic Multi-racial	Mean	934.13	935.91	935.17	929.19	934.92	933.74	935.43	933.78
	Std. Dev.	10.65	13.62	12.92	10.50	11.51	11.15	12.53	10.95
	N	24	23	24	26	24	23	23	23
Non-Hispanic White	Mean	931.96	932.48	932.72	927.55	932.60	930.24	932.40	930.70
	Std. Dev.	11.01	13.33	14.65	11.89	12.10	11.92	12.29	11.47
	N	483	483	483	484	483	483	483	483
Missing	Mean	934.52	934.78	935.89	930.38	935.50	932.87	934.77	933.40
	Std. Dev.	9.82	11.67	13.10	10.58	10.71	10.29	10.72	9.95
	N	339	339	339	340	339	339	339	339

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	936.10	935.49	935.19	932.57	935.78	934.30	935.70	934.50
	Std. Dev.	10.38	11.37	13.63	12.52	11.11	11.37	10.71	10.95
	N	865	865	864	864	864	864	865	864
Non-Hispanic Pacific Islander	Mean	934.52	934.24	936.12	931.44	935.44	933.06	934.28	933.60
	Std. Dev.	11.22	12.22	12.86	11.90	11.37	11.31	11.52	11.03
	N	50	50	50	50	50	50	50	50
Non-Hispanic Black	Mean	936.13	935.18	936.18	932.07	936.31	933.93	935.49	934.44
	Std. Dev.	10.27	10.99	13.29	12.60	10.97	11.09	10.46	10.74
	N	502	502	501	502	501	501	502	501
Hispanic (Of Any Race)	Mean	937.15	935.95	935.90	931.81	936.65	934.18	936.33	934.71
	Std. Dev.	10.47	10.75	13.42	11.71	11.18	10.65	10.40	10.50
	N	4,124	4,128	4,122	4,136	4,117	4,126	4,123	4,116
Non-Hispanic American Indian	Mean	939.71	939.13	938.61	933.73	939.22	936.78	939.35	937.15
	Std. Dev.	9.87	10.67	12.07	10.64	10.54	10.27	10.22	10.08
	N	41	40	41	41	41	40	40	40
Non-Hispanic Multi-racial	Mean	938.76	936.48	939.27	932.00	939.12	934.48	937.15	935.70
	Std. Dev.	7.62	7.36	10.28	9.49	8.03	7.85	7.11	7.60
	N	33	33	33	33	33	33	33	33
Non-Hispanic White	Mean	936.22	935.27	935.59	931.15	936.02	933.48	935.57	934.02
	Std. Dev.	10.72	11.36	13.82	12.22	11.50	11.21	10.90	11.01
	N	638	638	638	639	638	638	638	638
Missing	Mean	939.70	938.14	939.31	934.64	939.63	936.69	938.65	937.36
	Std. Dev.	9.20	9.44	11.14	11.23	9.54	9.78	9.03	9.35
	N	586	585	583	585	583	584	585	583

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	936.16	937.01	935.08	932.64	935.97	935.07	936.78	935.11
	Std. Dev.	10.98	12.32	13.40	11.11	11.62	11.05	11.60	10.88
	N	638	638	637	638	637	638	638	637
Non-Hispanic Pacific Islander	Mean	935.88	936.39	937.35	930.73	936.88	933.76	936.24	934.35
	Std. Dev.	11.73	11.96	13.36	10.31	11.68	10.41	11.66	10.51
	N	49	49	48	49	48	49	49	48
Non-Hispanic Black	Mean	937.14	937.09	936.76	932.42	937.31	935.00	937.14	935.46
	Std. Dev.	10.40	12.09	12.99	10.82	11.19	10.81	11.31	10.55
	N	404	404	404	404	404	404	404	404
Hispanic (Of Any Race)	Mean	938.08	938.29	936.48	932.66	937.65	935.71	938.26	936.06
	Std. Dev.	10.31	11.70	13.12	10.67	11.21	10.59	10.98	10.43
	N	3,812	3,812	3,813	3,813	3,809	3,811	3,810	3,806
Non-Hispanic American Indian	Mean	941.79	943.55	942.18	937.47	942.42	940.79	943.03	940.97
	Std. Dev.	6.45	7.09	8.93	8.25	6.73	7.26	6.51	6.83
	N	38	38	38	38	38	38	38	38
Non-Hispanic Multi-racial	Mean	938.81	938.52	935.81	932.26	937.67	935.67	938.67	936.00
	Std. Dev.	8.70	10.30	13.60	9.76	10.79	9.59	9.48	9.29
	N	27	27	27	27	27	27	27	27
Non-Hispanic White	Mean	937.09	937.53	936.66	932.42	937.24	935.22	937.42	935.55
	Std. Dev.	11.34	12.96	13.29	12.36	11.96	12.16	12.25	11.79
	N	472	472	472	472	472	472	472	472
Missing	Mean	939.89	940.19	939.24	934.53	939.97	937.60	940.11	938.06
	Std. Dev.	9.10	10.43	11.67	10.13	9.92	9.68	9.75	9.41
	N	400	401	401	401	400	401	400	400

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	937.48	937.60	935.39	934.15	936.61	936.09	937.66	936.07
	Std. Dev.	10.41	11.12	12.72	11.46	10.88	10.72	10.69	10.48
	N	830	830	831	831	830	830	830	830
Non-Hispanic Pacific Islander	Mean	940.51	939.82	939.26	934.54	940.03	937.41	940.13	938.03
	Std. Dev.	8.57	10.26	10.44	10.61	8.98	9.80	9.48	9.35
	N	39	39	39	39	39	39	39	39
Non-Hispanic Black	Mean	937.86	937.16	936.81	934.24	937.48	935.90	937.47	936.21
	Std. Dev.	10.49	11.13	11.94	11.43	10.52	10.70	10.75	10.38
	N	463	463	463	463	463	463	463	463
Hispanic (Of Any Race)	Mean	938.82	938.54	936.48	934.26	937.80	936.61	938.71	936.80
	Std. Dev.	10.03	10.74	12.22	10.98	10.45	10.23	10.31	10.04
	N	3,849	3,850	3,847	3,847	3,846	3,846	3,849	3,843
Non-Hispanic American Indian	Mean	939.38	940.44	938.23	936.13	938.94	938.48	940.23	938.46
	Std. Dev.	10.08	9.93	11.52	10.60	10.38	9.75	9.91	9.73
	N	52	52	52	52	52	52	52	52
Non-Hispanic Multi-racial	Mean	937.73	936.30	937.73	933.03	937.90	934.93	936.63	935.63
	Std. Dev.	10.57	11.53	11.42	10.62	9.72	10.01	11.11	9.82
	N	30	30	30	30	30	30	30	30
Non-Hispanic White	Mean	938.73	938.43	937.05	934.02	938.01	936.44	938.59	936.73
	Std. Dev.	10.17	10.65	11.92	11.69	10.40	10.47	10.32	10.17
	N	543	543	544	543	542	543	542	542
Missing	Mean	939.63	938.94	937.68	934.71	938.82	937.04	939.23	937.41
	Std. Dev.	9.68	10.19	10.95	10.67	9.83	9.85	9.79	9.58
	N	539	539	540	539	539	538	539	538

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	930.21	11.25	623	931.52	11.11	1,655	933.04	9.97	77	931.23	11.13	2,355
Reading	930.13	12.69	623	931.87	13.07	1,655	932.71	11.25	77	931.44	12.94	2,355
Speaking	930.71	14.43	622	931.92	14.54	1,652	931.82	14.21	77	931.60	14.51	2,351
Writing	925.22	10.81	624	927.01	11.40	1,666	927.47	10.15	78	926.56	11.24	2,368
Oral	930.73	12.12	622	931.98	12.06	1,652	932.62	11.18	77	931.67	12.06	2,351
Literacy	927.97	10.99	622	929.72	11.44	1,655	930.36	9.89	77	929.28	11.30	2,354
Comprehension	930.22	11.93	623	931.84	12.15	1,654	932.88	10.34	77	931.44	12.06	2,354
Overall	928.55	10.93	622	930.15	11.21	1,652	930.73	9.88	77	929.75	11.12	2,351

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	933.24	10.29	651	933.85	10.37	1,528	936.58	9.25	64	933.75	10.33	2,243
Reading	933.34	12.10	651	934.93	12.66	1,527	935.41	11.97	64	934.48	12.50	2,242
Speaking	933.39	14.31	651	934.89	14.34	1,528	934.95	14.01	64	934.46	14.34	2,243
Writing	928.39	11.17	658	929.40	11.47	1,535	928.44	10.60	64	929.08	11.37	2,257
Oral	933.56	11.63	650	934.65	11.69	1,527	936.08	10.78	64	934.37	11.66	2,241
Literacy	931.16	10.84	651	932.45	11.27	1,527	932.23	10.63	64	932.07	11.15	2,242
Comprehension	933.37	11.22	650	934.67	11.64	1,526	935.88	10.79	64	934.33	11.52	2,240
Overall	931.65	10.68	649	932.84	10.95	1,525	933.13	10.04	64	932.50	10.86	2,238

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	935.78	10.33	726	935.73	10.64	1,494	938.45	8.78	64	935.82	10.51	2,284
Reading	934.17	10.38	726	934.71	11.13	1,494	935.98	9.00	65	934.57	10.85	2,285
Speaking	934.57	13.76	727	935.27	13.39	1,491	936.28	13.02	65	935.07	13.50	2,283
Writing	929.73	11.15	732	931.13	11.83	1,502	930.11	9.78	65	930.66	11.58	2,299
Oral	935.30	11.16	726	935.64	11.19	1,491	937.41	10.27	64	935.58	11.16	2,281
Literacy	932.24	10.18	726	933.22	10.87	1,494	933.28	8.66	65	932.91	10.60	2,285
Comprehension	934.69	10.06	725	935.05	10.68	1,494	936.73	8.66	64	934.98	10.44	2,283
Overall	932.93	10.15	725	933.73	10.64	1,491	934.33	8.91	64	933.49	10.45	2,280

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	937.39	10.12	741	937.04	10.77	1,550	939.52	8.38	58	937.21	10.52	2,349
Reading	935.70	10.37	743	936.21	11.08	1,550	937.29	9.97	58	936.07	10.84	2,351
Speaking	935.81	13.17	743	936.46	13.34	1,546	938.71	9.91	58	936.31	13.22	2,347
Writing	931.49	11.77	743	932.46	11.91	1,547	931.59	10.64	58	932.13	11.84	2,348
Oral	936.70	10.90	741	936.88	11.31	1,546	939.21	8.29	58	936.88	11.12	2,345
Literacy	933.87	10.51	743	934.63	10.89	1,547	934.74	9.60	58	934.40	10.75	2,348
Comprehension	936.21	10.02	741	936.48	10.71	1,550	937.97	9.15	58	936.44	10.46	2,349
Overall	934.48	10.31	741	935.11	10.71	1,545	935.81	8.86	58	934.93	10.54	2,344

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	937.87	10.16	716	938.47	9.85	1,434	936.79	11.71	56	938.23	10.01	2,206
Reading	936.74	10.63	715	937.68	10.52	1,434	934.86	12.08	56	937.30	10.62	2,205
Speaking	936.37	13.36	715	937.44	12.98	1,431	934.39	13.68	56	937.02	13.14	2,202
Writing	932.73	12.00	715	934.18	12.16	1,432	930.63	11.63	56	933.62	12.12	2,203
Oral	937.25	11.06	715	938.08	10.67	1,430	935.70	12.05	56	937.75	10.85	2,201
Literacy	935.02	10.80	715	936.22	10.76	1,432	933.04	11.34	56	935.74	10.81	2,203
Comprehension	937.10	10.19	715	937.93	10.03	1,433	935.46	11.66	56	937.60	10.14	2,204
Overall	935.47	10.58	715	936.56	10.41	1,430	933.61	11.22	56	936.13	10.51	2,201

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	937.11	10.65	723	937.11	10.72	1,301	938.10	10.37	62	937.14	10.68	2,086
Reading	937.02	11.71	723	937.48	11.98	1,303	937.27	11.09	62	937.31	11.86	2,088
Speaking	935.71	13.40	722	936.33	13.04	1,305	935.11	12.81	62	936.08	13.16	2,089
Writing	931.71	10.68	723	932.69	10.56	1,304	930.98	9.70	62	932.30	10.59	2,089
Oral	936.76	11.58	722	937.07	11.39	1,301	936.97	11.10	62	936.96	11.45	2,085
Literacy	934.60	10.60	723	935.31	10.61	1,303	934.34	9.88	62	935.04	10.59	2,088
Comprehension	937.07	11.09	723	937.37	11.28	1,301	937.53	10.63	62	937.27	11.19	2,086
Overall	935.02	10.52	722	935.60	10.49	1,301	934.94	9.93	62	935.38	10.49	2,085

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	937.36	10.44	648	938.28	10.27	1,203	937.36	10.62	53	937.94	10.35	1,904
Reading	937.86	11.58	647	938.48	11.93	1,203	937.15	12.55	53	938.23	11.83	1,903
Speaking	936.50	12.95	647	936.93	13.10	1,202	935.51	13.57	53	936.74	13.06	1,902
Writing	932.47	10.94	647	932.86	10.74	1,202	930.15	9.85	53	932.65	10.80	1,902
Oral	937.29	11.16	647	937.99	11.24	1,202	936.81	11.49	53	937.72	11.22	1,902
Literacy	935.40	10.71	647	935.90	10.74	1,202	933.89	10.63	53	935.68	10.73	1,902
Comprehension	937.78	10.98	647	938.44	11.14	1,203	937.26	11.65	53	938.19	11.11	1,903
Overall	935.76	10.48	646	936.29	10.59	1,201	934.55	10.58	53	936.06	10.56	1,900

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	938.77	10.11	680	938.53	10.12	1,137	936.42	11.86	33	938.58	10.16	1,850
Reading	939.00	11.48	681	939.13	11.78	1,136	935.94	13.33	33	939.03	11.71	1,850
Speaking	936.75	13.00	679	937.41	12.82	1,137	929.79	15.63	33	937.03	12.98	1,849
Writing	933.50	11.38	681	933.54	10.92	1,137	926.82	10.94	33	933.41	11.13	1,851
Oral	938.13	10.90	679	938.35	10.97	1,136	933.36	13.05	33	938.18	11.01	1,848
Literacy	936.50	10.85	681	936.59	10.77	1,136	931.61	11.45	33	936.47	10.83	1,850
Comprehension	938.97	10.80	680	938.98	11.01	1,136	936.09	12.68	33	938.93	10.97	1,849
Overall	936.75	10.52	679	936.86	10.51	1,135	931.88	11.57	33	936.73	10.55	1,847

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	937.97	10.39	597	938.36	10.04	1,021	937.54	9.86	68	938.19	10.16	1,686
Reading	937.72	11.11	597	938.28	10.66	1,022	936.79	10.39	68	938.02	10.82	1,687
Speaking	935.97	12.42	596	936.62	12.10	1,023	933.94	12.29	68	936.28	12.23	1,687
Writing	934.10	11.43	597	934.24	10.75	1,024	931.84	11.48	68	934.10	11.04	1,689
Oral	937.14	10.73	596	937.64	10.39	1,021	935.93	10.44	68	937.39	10.52	1,685
Literacy	936.13	10.73	597	936.47	10.05	1,022	934.51	10.32	68	936.27	10.31	1,687
Comprehension	937.88	10.67	597	938.39	10.22	1,021	937.16	10.01	68	938.16	10.38	1,686
Overall	936.24	10.48	596	936.66	9.89	1,021	934.75	10.07	68	936.44	10.12	1,685

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	938.73	9.93	502	939.32	9.60	878	938.11	12.21	38	939.08	9.80	1,418
Reading	938.26	11.05	502	939.03	10.07	878	937.39	11.94	38	938.71	10.49	1,418
Speaking	936.33	12.24	502	937.12	11.84	878	936.18	13.57	39	936.82	12.04	1,419
Writing	933.87	11.12	502	934.47	10.91	878	934.41	11.50	39	934.26	11.00	1,419
Oral	937.68	10.44	502	938.37	10.00	878	937.11	12.27	38	938.09	10.24	1,418
Literacy	936.25	10.47	502	936.97	9.82	878	936.13	11.43	38	936.69	10.10	1,418
Comprehension	938.49	10.51	502	939.21	9.70	878	937.66	11.69	38	938.91	10.06	1,418
Overall	936.54	10.21	502	937.22	9.55	878	936.24	11.42	38	936.95	9.84	1,418

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	939.51	9.42	535	938.87	9.92	808	937.47	11.41	36	939.08	9.78	1,379
Reading	939.27	10.14	536	938.74	10.61	808	936.97	11.52	36	938.90	10.46	1,380
Speaking	937.64	11.40	537	936.72	12.20	808	935.19	12.72	36	937.04	11.93	1,381
Writing	935.15	11.04	536	934.68	11.08	808	932.50	10.92	36	934.80	11.07	1,380
Oral	938.72	9.70	535	937.95	10.45	808	936.50	11.77	36	938.21	10.21	1,379
Literacy	937.44	9.97	536	936.90	10.27	808	934.89	10.99	36	937.06	10.18	1,380
Comprehension	939.41	9.75	535	938.86	10.18	808	937.17	11.44	36	939.03	10.06	1,379
Overall	937.61	9.65	535	937.04	10.05	808	935.22	11.08	36	937.22	9.93	1,379

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	938.67	10.42	696	938.41	10.41	1,127	933.38	12.89	39	938.40	10.49	1,862
Reading	938.17	10.77	696	938.16	11.14	1,126	930.03	13.94	39	937.99	11.13	1,861
Speaking	936.37	12.07	695	936.34	12.26	1,125	929.72	13.98	39	936.21	12.26	1,859
Writing	934.32	11.36	692	934.09	11.12	1,125	928.03	12.41	39	934.05	11.28	1,856
Oral	937.66	10.66	695	937.53	10.66	1,125	931.72	12.73	39	937.46	10.74	1,859
Literacy	936.46	10.46	692	936.34	10.48	1,125	929.18	12.48	39	936.23	10.57	1,856
Comprehension	938.40	10.51	696	938.32	10.74	1,126	931.15	13.16	39	938.20	10.76	1,861
Overall	936.65	10.28	692	936.53	10.28	1,124	929.74	12.18	39	936.43	10.37	1,855

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	930.62	931.37	931.44	927.06	931.29	929.44	931.21	929.75
	Std. Dev.	10.81	13.40	14.55	11.59	11.99	11.60	12.29	11.30
	N	394	394	394	396	394	394	394	394
Non-Hispanic Pacific Islander	Mean	929.47	927.00	930.71	922.35	930.29	925.00	927.82	926.29
	Std. Dev.	11.28	11.39	13.84	9.62	11.42	9.94	11.18	10.00
	N	17	17	17	17	17	17	17	17
Non-Hispanic Black	Mean	930.58	930.95	932.51	926.87	931.79	929.15	930.92	929.76
	Std. Dev.	11.57	13.59	14.84	11.41	12.52	11.72	12.64	11.54
	N	182	182	182	182	182	182	182	182
Hispanic (Of Any Race)	Mean	931.34	931.23	931.04	925.96	931.45	928.90	931.33	929.40
	Std. Dev.	11.20	12.66	14.65	11.05	12.16	11.13	11.90	11.05
	N	1,345	1,345	1,341	1,354	1,341	1,344	1,344	1,341
Non-Hispanic American Indian	Mean	930.40	935.20	928.30	924.10	929.50	930.00	933.90	929.60
	Std. Dev.	14.02	12.29	14.89	8.67	14.17	9.98	11.97	10.53
	N	10	10	10	10	10	10	10	10
Non-Hispanic Multi-racial	Mean	931.15	932.54	933.15	928.14	932.38	931.38	932.23	931.15
	Std. Dev.	12.53	14.99	14.06	11.72	13.14	12.54	14.11	12.33
	N	13	13	13	14	13	13	13	13
Non-Hispanic White	Mean	931.00	931.61	932.11	926.86	931.83	929.49	931.50	929.96
	Std. Dev.	11.35	13.81	14.34	11.85	12.05	12.10	12.70	11.60
	N	236	236	236	236	236	236	236	236
Missing	Mean	933.10	933.82	935.11	930.01	934.37	932.19	933.68	932.57
	Std. Dev.	9.78	11.61	12.48	10.18	10.22	9.79	10.66	9.50
	N	158	158	158	159	158	158	158	158

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	933.17	935.33	934.31	931.40	934.01	933.60	934.74	933.47
	Std. Dev.	10.19	13.90	13.92	12.26	11.38	12.28	12.45	11.59
	N	288	288	288	290	288	288	288	288
Non-Hispanic Pacific Islander	Mean	931.80	933.10	930.30	927.40	931.30	930.60	932.80	930.60
	Std. Dev.	7.57	13.69	12.02	8.87	8.98	10.39	11.86	9.54
	N	10	10	10	10	10	10	10	10
Non-Hispanic Black	Mean	933.03	933.76	935.12	927.70	934.34	930.99	933.61	931.76
	Std. Dev.	10.63	12.13	14.24	11.45	11.93	11.12	11.34	10.96
	N	184	184	184	185	184	184	184	184
Hispanic (Of Any Race)	Mean	933.85	934.44	934.37	928.74	934.38	931.89	934.33	932.38
	Std. Dev.	10.34	12.26	14.42	11.04	11.67	10.80	11.36	10.64
	N	1,304	1,304	1,304	1,313	1,302	1,304	1,302	1,300
Non-Hispanic American Indian	Mean	933.39	933.56	931.94	926.22	932.94	930.17	933.56	930.78
	Std. Dev.	11.08	12.14	15.63	12.42	12.60	11.43	11.72	11.56
	N	18	18	18	18	18	18	18	18
Non-Hispanic Multi-racial	Mean	937.64	940.30	937.55	930.42	937.91	936.80	939.60	937.20
	Std. Dev.	6.26	10.04	10.97	8.71	8.27	8.05	8.48	7.57
	N	11	10	11	12	11	10	10	10
Non-Hispanic White	Mean	932.88	933.30	933.29	928.20	933.34	930.96	933.26	931.40
	Std. Dev.	10.58	12.80	14.91	11.89	12.09	11.69	11.81	11.30
	N	247	247	247	248	247	247	247	247
Missing	Mean	935.75	935.62	936.58	930.71	936.48	933.46	935.73	934.12
	Std. Dev.	9.69	11.66	13.58	10.92	11.03	10.68	10.68	10.27
	N	181	181	181	181	181	181	181	181

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	935.21	934.62	934.61	931.90	935.07	933.54	934.81	933.76
	Std. Dev.	10.53	11.13	13.24	12.19	10.77	11.14	10.57	10.70
	N	294	294	294	294	294	294	294	294
Non-Hispanic Pacific Islander	Mean	934.17	932.17	938.50	933.17	936.42	932.92	932.67	933.83
	Std. Dev.	9.87	11.44	7.99	8.48	8.18	9.31	10.14	8.41
	N	12	12	12	12	12	12	12	12
Non-Hispanic Black	Mean	934.60	933.50	934.54	930.07	934.71	932.10	933.87	932.70
	Std. Dev.	9.87	10.68	13.97	12.31	11.09	10.76	10.04	10.51
	N	161	161	161	162	161	161	161	161
Hispanic (Of Any Race)	Mean	936.01	934.65	935.04	930.44	935.65	932.84	935.09	933.48
	Std. Dev.	10.61	10.81	13.55	11.46	11.25	10.53	10.47	10.43
	N	1,373	1,375	1,373	1,385	1,371	1,375	1,373	1,371
Non-Hispanic American Indian	Mean	936.59	934.19	935.41	929.41	936.06	932.00	934.81	932.81
	Std. Dev.	11.33	12.42	13.93	11.72	12.20	11.65	11.73	11.66
	N	17	16	17	17	17	16	16	16
Non-Hispanic Multi-racial	Mean	937.00	937.13	941.13	933.00	939.13	935.25	937.13	936.50
	Std. Dev.	9.18	7.29	5.44	4.18	6.99	5.31	7.62	5.83
	N	8	8	8	8	8	8	8	8
Non-Hispanic White	Mean	934.38	933.50	933.55	929.37	934.09	931.69	933.78	932.19
	Std. Dev.	10.68	11.57	14.42	11.52	11.74	10.91	11.05	10.86
	N	222	222	222	223	222	222	222	222
Missing	Mean	938.06	936.14	937.69	932.16	938.00	934.43	936.78	935.28
	Std. Dev.	9.59	9.62	12.06	10.94	10.20	9.66	9.30	9.43
	N	197	197	196	198	196	197	197	196

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	935.35	934.84	934.59	931.98	935.10	933.68	935.03	933.86
	Std. Dev.	10.61	11.95	14.11	12.51	11.57	11.53	11.20	11.19
	N	297	297	297	297	297	297	297	297
Non-Hispanic Pacific Islander	Mean	934.40	933.95	934.85	930.55	934.75	932.55	934.05	933.05
	Std. Dev.	12.04	11.97	13.14	12.01	12.12	11.02	11.88	11.11
	N	20	20	20	20	20	20	20	20
Non-Hispanic Black	Mean	936.38	935.34	936.41	931.58	936.60	933.82	935.68	934.41
	Std. Dev.	10.17	10.80	12.73	11.96	10.47	10.49	10.31	10.17
	N	187	187	186	186	186	186	187	186
Hispanic (Of Any Race)	Mean	937.42	936.19	936.12	931.79	936.88	934.28	936.58	934.86
	Std. Dev.	10.49	10.60	13.23	11.61	11.12	10.54	10.30	10.40
	N	1,392	1,394	1,392	1,393	1,390	1,393	1,392	1,389
Non-Hispanic American Indian	Mean	940.18	940.27	939.36	937.00	939.82	939.00	940.36	938.82
	Std. Dev.	10.15	10.15	11.35	9.94	10.39	9.95	10.10	9.73
	N	11	11	11	11	11	11	11	11
Non-Hispanic Multi-racial	Mean	939.71	935.86	937.07	931.93	938.50	934.21	937.00	935.21
	Std. Dev.	5.05	5.64	11.97	9.85	7.73	7.32	5.21	7.12
	N	14	14	14	14	14	14	14	14
Non-Hispanic White	Mean	936.26	935.32	936.43	931.84	936.45	933.84	935.60	934.42
	Std. Dev.	11.38	11.38	13.63	12.42	11.77	11.45	11.12	11.33
	N	213	213	213	213	213	213	213	213
Missing	Mean	940.01	938.43	939.69	935.29	939.97	937.18	938.91	937.79
	Std. Dev.	9.30	9.93	11.31	11.35	9.74	10.15	9.43	9.67
	N	215	215	214	214	214	214	215	214

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	937.88	937.14	936.47	933.93	937.29	935.81	937.39	936.00
	Std. Dev.	9.74	10.78	13.43	12.78	10.80	11.30	10.09	10.81
	N	274	274	273	273	273	273	274	273
Non-Hispanic Pacific Islander	Mean	934.89	935.94	935.94	931.28	935.56	933.72	935.61	934.06
	Std. Dev.	11.11	12.74	14.81	13.51	12.25	12.72	11.81	12.37
	N	18	18	18	18	18	18	18	18
Non-Hispanic Black	Mean	937.43	936.73	937.62	934.77	937.63	935.97	936.94	936.29
	Std. Dev.	10.61	11.30	13.04	13.18	11.21	11.76	10.82	11.32
	N	154	154	154	154	154	154	154	154
Hispanic (Of Any Race)	Mean	938.02	937.03	936.54	933.24	937.41	935.42	937.33	935.81
	Std. Dev.	10.20	10.72	13.44	11.89	11.09	10.73	10.29	10.54
	N	1,359	1,359	1,357	1,358	1,356	1,358	1,358	1,356
Non-Hispanic American Indian	Mean	943.38	944.23	942.15	936.62	942.85	940.77	944.08	941.08
	Std. Dev.	5.05	4.02	8.41	7.24	6.15	5.04	4.10	4.95
	N	13	13	13	13	13	13	13	13
Non-Hispanic Multi-racial	Mean	938.82	936.82	940.73	931.36	939.91	934.27	937.36	935.73
	Std. Dev.	8.79	9.07	10.13	11.52	9.00	9.77	8.66	9.12
	N	11	11	11	11	11	11	11	11
Non-Hispanic White	Mean	938.18	937.16	936.93	932.40	937.66	935.08	937.50	935.62
	Std. Dev.	9.64	10.78	13.08	12.54	10.60	11.01	10.15	10.52
	N	203	203	203	203	203	203	203	203
Missing	Mean	941.18	940.04	940.66	936.68	941.08	938.66	940.46	939.17
	Std. Dev.	8.27	8.08	9.51	10.88	8.16	8.88	7.72	8.36
	N	174	173	173	173	173	173	173	173

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	934.65	935.92	934.00	932.11	934.65	934.28	935.54	934.13
	Std. Dev.	11.44	12.57	14.25	11.13	12.29	11.23	11.91	11.18
	N	210	210	210	210	210	210	210	210
Non-Hispanic Pacific Islander	Mean	935.31	937.00	939.15	931.08	937.54	934.23	936.38	934.92
	Std. Dev.	10.82	11.58	12.65	10.00	11.57	10.06	11.14	10.14
	N	13	13	13	13	13	13	13	13
Non-Hispanic Black	Mean	936.07	936.11	936.46	932.19	936.59	934.39	936.13	934.82
	Std. Dev.	10.84	12.48	12.93	11.20	11.52	11.28	11.70	10.98
	N	153	153	153	153	153	153	153	153
Hispanic (Of Any Race)	Mean	937.41	937.38	935.89	932.07	937.00	934.95	937.41	935.34
	Std. Dev.	10.57	11.75	13.17	10.33	11.37	10.38	11.07	10.33
	N	1,360	1,361	1,362	1,362	1,359	1,361	1,360	1,359
Non-Hispanic American Indian	Mean	942.33	942.93	940.00	937.67	941.53	940.47	942.67	940.60
	Std. Dev.	4.25	5.22	9.40	8.72	6.51	6.63	4.61	6.24
	N	15	15	15	15	15	15	15	15
Non-Hispanic Multi-racial	Mean	936.00	936.08	931.75	930.42	934.17	933.50	936.08	933.42
	Std. Dev.	10.07	11.23	15.76	9.96	12.32	10.38	10.66	10.23
	N	12	12	12	12	12	12	12	12
Non-Hispanic White	Mean	936.74	936.87	936.59	931.77	937.03	934.55	936.84	935.04
	Std. Dev.	11.26	12.53	13.00	12.03	11.82	11.66	11.87	11.32
	N	175	175	175	175	175	175	175	175
Missing	Mean	939.48	939.95	939.41	935.11	939.84	937.79	939.80	938.09
	Std. Dev.	9.19	10.23	11.09	9.31	9.74	9.28	9.63	9.08
	N	148	149	149	149	148	149	148	148

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	935.54	936.03	934.48	931.70	935.34	934.07	935.94	934.28
	Std. Dev.	11.07	12.86	13.71	11.22	11.78	11.34	12.02	11.15
	N	210	210	209	210	209	210	210	209
Non-Hispanic Pacific Islander	Mean	936.41	936.71	937.76	930.65	937.47	933.82	936.71	934.82
	Std. Dev.	11.25	13.07	12.34	9.51	11.58	11.01	12.31	10.96
	N	17	17	17	17	17	17	17	17
Non-Hispanic Black	Mean	938.57	938.49	937.57	932.50	938.47	935.77	938.57	936.34
	Std. Dev.	9.10	10.55	12.79	9.76	10.33	9.30	9.82	9.15
	N	129	129	129	129	129	129	129	129
Hispanic (Of Any Race)	Mean	937.99	938.24	936.51	932.44	937.62	935.57	938.20	935.95
	Std. Dev.	10.37	11.81	13.15	10.76	11.30	10.72	11.09	10.58
	N	1,259	1,258	1,258	1,257	1,258	1,257	1,258	1,256
Non-Hispanic American Indian	Mean	938.89	945.22	943.78	938.44	941.78	942.11	943.44	941.78
	Std. Dev.	9.23	6.60	6.14	8.11	5.79	6.82	7.23	6.37
	N	9	9	9	9	9	9	9	9
Non-Hispanic Multi-racial	Mean	941.40	940.10	939.60	933.30	940.90	937.00	940.60	938.00
	Std. Dev.	6.97	10.46	11.20	6.97	8.85	8.41	8.92	7.81
	N	10	10	10	10	10	10	10	10
Non-Hispanic White	Mean	938.27	938.60	937.79	934.05	938.39	936.60	938.53	936.83
	Std. Dev.	10.76	12.51	12.68	11.60	11.25	11.66	11.81	11.30
	N	146	146	146	146	146	146	146	146
Missing	Mean	940.43	940.72	939.90	934.78	940.59	937.98	940.67	938.56
	Std. Dev.	8.87	10.02	11.37	10.53	9.74	9.67	9.36	9.31
	N	124	124	124	124	124	124	124	124

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	938.22	939.02	936.70	934.05	937.84	936.80	938.79	936.85
	Std. Dev.	10.09	11.25	12.01	10.83	10.50	10.36	10.58	10.09
	N	218	218	218	218	218	218	218	218
Non-Hispanic Pacific Islander	Mean	935.79	935.68	935.67	930.58	935.83	933.37	935.74	933.50
	Std. Dev.	12.71	11.12	14.52	11.17	11.79	10.07	11.39	10.30
	N	19	19	18	19	18	19	19	18
Non-Hispanic Black	Mean	936.97	936.84	936.30	932.62	936.99	934.96	936.91	935.33
	Std. Dev.	10.94	12.95	13.24	11.39	11.56	11.60	12.12	11.28
	N	122	122	122	122	122	122	122	122
Hispanic (Of Any Race)	Mean	938.94	939.37	937.13	933.58	938.41	936.73	939.28	936.99
	Std. Dev.	9.88	11.44	13.01	10.90	10.87	10.59	10.68	10.33
	N	1,193	1,193	1,193	1,194	1,192	1,193	1,192	1,191
Non-Hispanic American Indian	Mean	943.07	943.14	943.50	936.64	943.79	940.29	943.14	940.86
	Std. Dev.	5.61	8.77	9.43	7.71	7.28	8.04	7.64	7.64
	N	14	14	14	14	14	14	14	14
Non-Hispanic Multi-racial	Mean	940.40	941.20	938.00	934.60	939.60	938.20	941.00	938.20
	Std. Dev.	5.75	4.96	8.90	12.77	7.31	8.68	4.94	7.98
	N	5	5	5	5	5	5	5	5
Non-Hispanic White	Mean	936.36	937.26	935.65	931.60	936.37	934.66	937.01	934.91
	Std. Dev.	11.86	13.78	14.09	13.28	12.67	13.05	12.99	12.66
	N	151	151	151	151	151	151	151	151
Missing	Mean	939.84	939.95	938.41	933.59	939.52	937.02	939.91	937.55
	Std. Dev.	9.20	11.0	12.53	10.56	10.2	10.12	10.2	9.85
	N	128	128	128	128	128	128	128	128

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	937.27	937.37	935.15	934.47	936.38	936.14	937.42	936.04
	Std. Dev.	9.86	10.60	12.47	11.05	10.51	10.36	10.17	10.13
	N	215	215	216	216	215	215	215	215
Non-Hispanic Pacific Islander	Mean	941.78	942.11	943.67	932.33	942.78	937.33	942.11	938.89
	Std. Dev.	4.21	5.19	2.49	7.32	3.12	5.29	4.61	4.20
	N	9	9	9	9	9	9	9	9
Non-Hispanic Black	Mean	938.57	938.66	937.20	934.90	938.04	936.99	938.73	937.12
	Std. Dev.	10.04	10.41	12.04	11.10	10.16	10.13	10.11	9.90
	N	119	119	119	119	119	119	119	119
Hispanic (Of Any Race)	Mean	938.12	938.03	936.11	933.94	937.27	936.19	938.15	936.35
	Std. Dev.	10.43	11.14	12.51	11.09	10.78	10.49	10.67	10.33
	N	1,009	1,009	1,008	1,010	1,008	1,009	1,009	1,008
Non-Hispanic American Indian	Mean	938.90	939.70	939.55	933.35	939.35	936.70	939.60	937.35
	Std. Dev.	11.28	11.53	10.33	11.61	10.66	11.21	11.47	10.80
	N	20	20	20	20	20	20	20	20
Non-Hispanic Multi-racial	Mean	937.45	935.36	933.55	930.64	935.73	933.18	935.82	933.73
	Std. Dev.	10.17	10.53	14.61	13.26	10.80	10.49	10.28	10.39
	N	11	11	11	11	11	11	11	11
Non-Hispanic White	Mean	938.35	937.76	936.80	934.12	937.69	936.17	938.01	936.46
	Std. Dev.	9.75	10.37	11.44	11.01	9.99	10.02	9.94	9.75
	N	156	157	157	157	156	157	156	156
Missing	Mean	939.29	938.42	937.12	934.39	938.40	936.62	938.80	936.97
	Std. Dev.	9.09	9.67	10.92	10.45	9.48	9.47	9.21	9.18
	N	147	147	147	147	147	147	147	147

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	936.88	936.63	934.39	931.92	935.79	934.48	936.77	934.70
	Std. Dev.	10.76	11.39	13.47	11.64	11.43	10.97	10.96	10.75
	N	171	171	171	171	171	171	171	171
Non-Hispanic Pacific Islander	Mean	938.50	936.25	936.17	931.33	937.50	934.08	937.08	934.83
	Std. Dev.	10.63	13.32	12.13	12.37	11.09	12.01	12.09	11.53
	N	12	12	12	12	12	12	12	12
Non-Hispanic Black	Mean	937.67	936.58	936.61	933.57	937.28	935.29	937.01	935.71
	Std. Dev.	10.68	11.32	12.29	10.90	10.93	10.61	10.92	10.39
	N	109	109	109	109	109	109	109	109
Hispanic (Of Any Race)	Mean	939.59	939.25	937.12	934.69	938.50	937.17	939.45	937.42
	Std. Dev.	9.41	10.25	11.83	10.74	9.87	9.81	9.77	9.54
	N	865	865	865	865	865	865	865	865
Non-Hispanic American Indian	Mean	939.56	940.67	937.22	937.67	938.50	939.39	940.44	938.94
	Std. Dev.	8.79	7.87	12.91	9.13	10.34	7.59	8.02	8.28
	N	18	18	18	18	18	18	18	18
Non-Hispanic Multi-racial	Mean	944.20	943.00	942.00	938.60	943.20	941.00	943.40	941.60
	Std. Dev.	3.49	3.90	3.79	5.92	3.43	4.60	3.26	4.18
	N	5	5	5	5	5	5	5	5
Non-Hispanic White	Mean	938.19	938.57	936.51	933.54	937.48	936.25	938.52	936.44
	Std. Dev.	10.27	10.21	12.06	11.81	10.61	10.25	10.06	10.03
	N	127	127	127	127	127	127	127	127
Missing	Mean	940.63	939.77	938.52	935.53	939.70	937.86	940.11	938.28
	Std. Dev.	9.29	9.75	10.44	10.70	9.32	9.80	9.38	9.42
	N	111	111	112	112	111	111	111	111

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	938.04	938.02	936.28	935.08	937.34	936.76	938.11	936.77
	Std. Dev.	10.04	11.15	12.12	10.62	10.34	10.35	10.52	10.07
	N	184	184	184	184	184	184	184	184
Non-Hispanic Pacific Islander	Mean	942.89	943.11	943.67	939.33	943.44	941.44	943.11	941.89
	Std. Dev.	4.33	4.75	2.05	5.85	2.71	4.99	4.56	4.28
	N	9	9	9	9	9	9	9	9
Non-Hispanic Black	Mean	939.10	937.77	937.88	936.02	938.63	937.09	938.27	937.37
	Std. Dev.	9.36	10.65	10.83	11.82	9.41	10.79	10.06	10.13
	N	112	112	112	112	112	112	112	112
Hispanic (Of Any Race)	Mean	938.92	938.78	936.53	934.32	937.87	936.76	938.89	936.89
	Std. Dev.	9.99	10.62	12.40	11.17	10.56	10.30	10.26	10.12
	N	827	828	828	828	827	828	827	827
Non-Hispanic American Indian	Mean	928.50	927.00	927.50	928.00	928.00	927.50	927.50	927.50
	Std. Dev.	18.50	17.00	17.50	18.00	18.00	17.50	17.50	17.50
	N	2	2	2	2	2	2	2	2
Non-Hispanic Multi-racial	Mean	941.67	940.33	942.50	936.17	942.17	938.67	940.50	939.50
	Std. Dev.	3.77	3.68	2.69	3.58	2.79	2.98	3.55	2.69
	N	6	6	6	6	6	6	6	6
Non-Hispanic White	Mean	940.16	939.92	938.67	934.97	939.54	937.65	940.07	938.03
	Std. Dev.	9.34	10.21	11.00	11.61	9.60	10.26	9.75	9.80
	N	116	116	117	116	116	116	116	116
Missing	Mean	940.48	940.89	938.73	936.11	939.77	938.68	940.84	938.85
	Std. Dev.	8.54	7.93	9.97	9.80	8.74	8.39	7.91	8.24
	N	123	123	123	123	123	123	123	123

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	937.65	938.15	935.61	934.68	936.81	936.63	938.12	936.50
	Std. Dev.	10.85	11.28	12.79	12.05	11.15	10.99	11.00	10.77
	N	260	260	260	260	260	260	260	260
Non-Hispanic Pacific Islander	Mean	939.56	939.00	934.56	936.22	937.22	937.89	939.22	937.56
	Std. Dev.	10.84	11.66	13.85	12.33	11.53	11.81	11.24	11.67
	N	9	9	9	9	9	9	9	9
Non-Hispanic Black	Mean	936.21	935.68	935.63	932.56	936.07	934.32	935.94	934.71
	Std. Dev.	11.46	11.81	12.37	11.57	11.26	11.01	11.56	10.83
	N	123	123	123	123	123	123	123	123
Hispanic (Of Any Race)	Mean	938.78	938.30	936.27	934.17	937.67	936.44	938.53	936.65
	Std. Dev.	10.10	10.78	12.11	10.91	10.49	10.25	10.39	10.08
	N	1,148	1,148	1,146	1,144	1,146	1,144	1,148	1,143
Non-Hispanic American Indian	Mean	941.75	943.58	939.33	939.83	940.75	941.92	943.08	941.42
	Std. Dev.	5.42	4.91	8.44	6.49	6.25	5.16	4.82	5.54
	N	12	12	12	12	12	12	12	12
Non-Hispanic Multi-racial	Mean	931.13	930.38	937.25	930.50	934.38	930.75	930.63	931.63
	Std. Dev.	13.32	15.80	11.31	10.40	11.43	12.21	14.99	11.94
	N	8	8	8	8	8	8	8	8
Non-Hispanic White	Mean	938.47	937.84	936.49	933.57	937.60	935.93	938.10	936.22
	Std. Dev.	11.02	11.51	12.88	12.30	11.14	11.22	11.23	10.92
	N	144	143	143	143	143	143	143	143
Missing	Mean	938.57	937.33	936.80	933.33	937.84	935.57	937.77	936.08
	Std. Dev.	11.08	12.06	11.90	11.30	11.10	10.97	11.53	10.76
	N	158	158	158	157	158	157	158	157

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	0.847**	0.750**	0.681**
	N	4,598	4,594	4,592	4,596
Reading	Pearson Correlation		1	0.725**	0.730**
	N		4,597	4,591	4,596
Writing	Pearson Correlation			1	0.712**
	N			4,594	4,594
Speaking	Pearson Correlation				1
	N				4,625

**. 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	0.878**	0.771**	0.714**
	N	6,839	6,836	6,827	6,832
Reading	Pearson Correlation		1	0.782**	0.786**
	N		6,841	6,831	6,836
Writing	Pearson Correlation			1	0.739**
	N			6,832	6,831
Speaking	Pearson Correlation				1
	N				6,850

**. 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	0.873**	0.783**	0.731**
	N	5,840	5,838	5,835	5,838
Reading	Pearson Correlation		1	0.789**	0.773**
	N		5,841	5,835	5,840
Writing	Pearson Correlation			1	0.741**
	N			5,840	5,836
Speaking	Pearson Correlation				1
	N				5,842

**. 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	0.886**	0.775**	0.742**
	N	6,345	6,344	6,341	6,339
Reading	Pearson Correlation		1	0.778**	0.773**
	N		6,346	6,343	6,341
Writing	Pearson Correlation			1	0.737**
	N			6,346	6,341
Speaking	Pearson Correlation				1
	N				6,344

**. 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

	Listening Proficiency Range										
	A1		A2		A3		P1		P2		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	960	20.88	609	13.24	918	19.97	955	20.77	1,156	25.14	4,598
3-5	902	13.19	598	8.74	926	13.54	1,473	21.54	2,940	42.99	6,839
6-8	691	11.83	504	8.63	695	11.90	738	12.64	3,212	55.00	5,840
9-12	647	10.20	443	6.98	737	11.62	1,331	20.98	3,187	50.23	6,345
Total	3,200	13.55	2,154	9.12	3,276	13.87	4,497	19.04	10,495	44.43	23,622

Table 4.9.1 B

Proficiency Level by Cluster: Reading

	Reading Proficiency Range										
	A1		A2		A3		P1		P2		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1,061	23.08	694	15.10	860	18.71	930	20.23	1,052	22.88	4,597
3-5	984	14.38	828	12.10	960	14.03	1,566	22.89	2,503	36.59	6,841
6-8	786	13.46	395	6.76	607	10.39	1,032	17.67	3,021	51.72	5,841
9-12	710	11.19	503	7.93	687	10.83	1,167	18.39	3,279	51.67	6,346
Total	3,541	14.99	2,420	10.24	3,114	13.18	4,695	19.87	9,855	41.71	23,625

Table 4.9.1 C

Proficiency Level by Cluster: Speaking

	Speaking Proficiency Range										
	A1		A2		A3		P1		P2		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1,347	29.32	196	4.27	618	13.45	1,463	31.85	970	21.11	4,594
3-5	1,445	21.15	330	4.83	531	7.77	2,143	31.37	2,383	34.88	6,832
6-8	1,119	19.16	182	3.12	675	11.56	1,620	27.74	2,244	38.42	5,840
9-12	1,141	17.98	166	2.62	641	10.10	1,596	25.15	2,802	44.15	6,346
Total	5,052	21.40	874	3.70	2,465	10.44	6,822	28.89	8,399	35.57	23,612

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	1,517	32.80	999	21.60	1,223	26.44	759	16.41	95	2.05	32	0.69	4,625
3-5	1,493	21.80	1,362	19.88	1,653	24.13	1,315	19.20	824	12.03	203	2.96	6,850
6-8	1,017	17.41	1,534	26.26	1,010	17.29	1,922	32.90	134	2.29	225	3.85	5,842
9-12	963	15.18	1,415	22.30	1,076	16.96	2,362	37.23	155	2.44	373	5.88	6,344
Total	4,990	21.09	5,310	22.44	4,962	20.97	6,358	26.87	1,208	5.11	833	3.52	23,661

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	1,225	26.68	369	8.04	692	15.07	1,254	27.31	1,052	22.91	4,592
3-5	1,187	17.39	461	6.75	811	11.88	1,770	25.93	2,598	38.05	6,827
6-8	954	16.35	347	5.95	700	12.00	1,167	20.00	2,667	45.71	5,835
9-12	911	14.37	347	5.47	746	11.76	1,662	26.21	2,675	42.19	6,341
Total	4,277	18.13	1,524	6.46	2,949	12.50	5,853	24.81	8,992	38.11	23,595

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	1,236	26.89	921	20.04	1,163	25.30	770	16.75	506	11.01	4,596
3-5	1,201	17.57	1,106	16.18	1,479	21.64	1,606	23.49	1,444	21.12	6,836
6-8	863	14.78	652	11.16	1,298	22.23	1,770	30.31	1,257	21.52	5,840
9-12	801	12.63	726	11.45	1,295	20.42	1,872	29.52	1,647	25.97	6,341
Total	4,101	17.37	3,405	14.42	5,235	22.17	6,018	25.49	4,854	20.56	23,613

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	1,022	22.25	645	14.04	828	18.02	1,151	25.05	948	20.64	4,594
3-5	949	13.88	759	11.10	916	13.40	1,431	20.93	2,781	40.68	6,836
6-8	762	13.05	405	6.94	593	10.16	1,126	19.29	2,952	50.57	5,838
9-12	682	10.75	474	7.47	680	10.72	1,245	19.62	3,263	51.43	6,344
Total	3,415	14.46	2,283	9.67	3,017	12.78	4,953	20.98	9,944	42.11	23,612

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	1,170	25.50	701	15.28	1,195	26.04	947	20.64	576	12.55	4,589
3-5	1,144	16.76	751	11.00	1,537	22.52	1,760	25.79	1,633	23.93	6,825
6-8	852	14.61	500	8.57	1,184	20.30	1,732	29.70	1,564	26.82	5,832
9-12	787	12.42	560	8.84	1,249	19.71	1,843	29.08	1,898	29.95	6,337
Total	3,953	16.76	2,512	10.65	5,165	21.90	6,282	26.64	5,671	24.05	23,583

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	576	24.46	339	14.39	487	20.68	465	19.75	488	20.72	2,355
2	384	17.12	270	12.04	431	19.22	490	21.85	668	29.78	2,243
3	345	15.11	243	10.64	336	14.71	548	23.99	812	35.55	2,284
4	311	13.24	194	8.26	316	13.45	493	20.99	1,035	44.06	2,349
5	246	11.15	161	7.30	274	12.42	432	19.58	1,093	49.55	2,206
6	267	12.80	224	10.74	258	12.37	258	12.37	1,079	51.73	2,086
7	227	11.92	148	7.77	238	12.50	247	12.97	1,044	54.83	1,904
8	197	10.65	132	7.14	199	10.76	233	12.59	1,089	58.86	1,850
9	179	10.62	135	8.01	214	12.69	354	21.00	804	47.69	1,686
10	134	9.45	91	6.42	162	11.42	282	19.89	749	52.82	1,418
11	126	9.14	93	6.74	144	10.44	306	22.19	710	51.49	1,379
12	208	11.17	124	6.66	217	11.65	389	20.89	924	49.62	1,862
Total	3,200	13.55	2,154	9.12	3,276	13.87	4,497	19.04	10,495	44.43	23,622

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	623	26.45	397	16.86	429	18.22	461	19.58	445	18.90	2,355
2	438	19.54	297	13.25	431	19.22	469	20.92	607	27.07	2,242
3	371	16.24	327	14.31	380	16.63	547	23.94	660	28.88	2,285
4	335	14.25	274	11.65	308	13.10	558	23.73	876	37.26	2,351
5	278	12.61	227	10.29	272	12.34	461	20.91	967	43.85	2,205
6	296	14.18	176	8.43	236	11.30	378	18.10	1,002	47.99	2,088
7	260	13.66	124	6.52	193	10.14	333	17.50	993	52.18	1,903
8	230	12.43	95	5.14	178	9.62	321	17.35	1,026	55.46	1,850
9	191	11.32	138	8.18	192	11.38	327	19.38	839	49.73	1,687
10	152	10.72	111	7.83	151	10.65	247	17.42	757	53.39	1,418
11	143	10.36	106	7.68	129	9.35	257	18.62	745	53.99	1,380
12	224	12.04	148	7.95	215	11.55	336	18.05	938	50.40	1,861
Total	3,541	14.99	2,420	10.24	3,114	13.18	4,695	19.87	9,855	41.71	23,625

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	767	32.62	112	4.76	346	14.72	742	31.56	384	16.33	2,351
2	580	25.86	84	3.74	272	12.13	721	32.14	586	26.13	2,243
3	531	23.26	116	5.08	207	9.07	749	32.81	680	29.79	2,283
4	483	20.58	122	5.20	175	7.46	734	31.27	833	35.49	2,347
5	431	19.57	92	4.18	149	6.77	660	29.97	870	39.51	2,202
6	422	20.20	70	3.35	265	12.69	576	27.57	756	36.19	2,089
7	365	19.19	57	3.00	213	11.20	517	27.18	750	39.43	1,902
8	332	17.96	55	2.97	197	10.65	527	28.50	738	39.91	1,849
9	306	18.14	48	2.85	181	10.73	440	26.08	712	42.21	1,687
10	251	17.69	31	2.18	143	10.08	342	24.10	652	45.95	1,419
11	233	16.87	33	2.39	126	9.12	352	25.49	637	46.13	1,381
12	351	18.88	54	2.90	191	10.27	462	24.85	801	43.09	1,859
Total	5,052	21.40	874	3.70	2,465	10.44	6,822	28.89	8,399	35.57	23,612

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	871	36.78	511	21.58	614	25.93	327	13.81	30	1.27	15	0.63	2,368
2	646	28.62	488	21.62	609	26.98	432	19.14	65	2.88	17	0.75	2,257
3	566	24.62	516	22.44	566	24.62	404	17.57	195	8.48	52	2.26	2,299
4	505	21.51	448	19.08	603	25.68	446	18.99	278	11.84	68	2.90	2,348
5	422	19.16	398	18.07	484	21.97	465	21.11	351	15.93	83	3.77	2,203
6	364	17.42	597	28.58	369	17.66	653	31.26	42	2.01	64	3.06	2,089
7	340	17.88	510	26.81	315	16.56	618	32.49	50	2.63	69	3.63	1,902
8	313	16.91	427	23.07	326	17.61	651	35.17	42	2.27	92	4.97	1,851
9	258	15.28	386	22.85	282	16.70	637	37.71	35	2.07	91	5.39	1,689
10	222	15.64	324	22.83	233	16.42	521	36.72	31	2.18	88	6.20	1,419
11	193	13.99	295	21.38	230	16.67	538	38.99	42	3.04	82	5.94	1,380
12	290	15.63	410	22.09	331	17.83	666	35.88	47	2.53	112	6.03	1,856
Total	4,990	21.09	5,310	22.44	4,962	20.97	6,358	26.87	1,208	5.11	833	3.52	23,66

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	703	29.90	214	9.10	377	16.04	638	27.14	419	17.82	2,351
2	522	23.29	155	6.92	315	14.06	616	27.49	633	28.25	2,241
3	447	19.60	174	7.63	312	13.68	637	27.93	711	31.17	2,281
4	403	17.19	149	6.35	272	11.60	607	25.88	914	38.98	2,345
5	337	15.31	138	6.27	227	10.31	526	23.90	973	44.21	2,201
6	370	17.75	128	6.14	270	12.95	417	20.00	900	43.17	2,085
7	304	15.98	120	6.31	221	11.62	377	19.82	880	46.27	1,902
8	280	15.15	99	5.36	209	11.31	373	20.18	887	48.00	1,848
9	249	14.78	101	5.99	204	12.11	462	27.42	669	39.70	1,685
10	200	14.10	72	5.08	167	11.78	336	23.70	643	45.35	1,418
11	185	13.42	77	5.58	135	9.79	381	27.63	601	43.58	1,379
12	277	14.90	97	5.22	240	12.91	483	25.98	762	40.99	1,859
Total	4,277	18.13	1,524	6.46	2,949	12.50	5,853	24.81	8,992	38.11	23,595

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	730	31.01	491	20.86	590	25.06	346	14.70	197	8.37	2,354
2	506	22.57	430	19.18	573	25.56	424	18.91	309	13.78	2,242
3	462	20.22	424	18.56	530	23.19	521	22.80	348	15.23	2,285
4	399	16.99	375	15.97	514	21.89	569	24.23	491	20.91	2,348
5	340	15.43	307	13.94	435	19.75	516	23.42	605	27.46	2,203
6	327	15.66	257	12.31	477	22.84	646	30.94	381	18.25	2,088
7	285	14.98	203	10.67	432	22.71	584	30.70	398	20.93	1,902
8	251	13.57	192	10.38	389	21.03	540	29.19	478	25.84	1,850
9	212	12.57	208	12.33	352	20.87	508	30.11	407	24.13	1,687
10	174	12.27	157	11.07	298	21.02	418	29.48	371	26.16	1,418
11	161	11.67	147	10.65	268	19.42	415	30.07	389	28.19	1,380
12	254	13.69	214	11.53	377	20.31	531	28.61	480	25.86	1,856
Total	4,101	17.37	3,405	14.42	5,235	22.17	6,018	25.49	4,854	20.56	23,613

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	606	25.74	365	15.51	429	18.22	554	23.53	400	16.99	2,354
2	416	18.57	280	12.50	399	17.81	597	26.65	548	24.46	2,240
3	361	15.81	317	13.89	334	14.63	518	22.69	753	32.98	2,283
4	321	13.67	249	10.60	301	12.81	513	21.84	965	41.08	2,349
5	267	12.11	193	8.76	281	12.75	400	18.15	1,063	48.23	2,204
6	294	14.09	172	8.25	227	10.88	419	20.09	974	46.69	2,086
7	251	13.19	123	6.46	197	10.35	366	19.23	966	50.76	1,903
8	217	11.74	110	5.95	169	9.14	341	18.44	1,012	54.73	1,849
9	188	11.15	131	7.77	186	11.03	349	20.70	832	49.35	1,686
10	145	10.23	101	7.12	144	10.16	269	18.97	759	53.53	1,418
11	136	9.86	96	6.96	135	9.79	275	19.94	737	53.44	1,379
12	213	11.45	146	7.85	215	11.55	352	18.91	935	50.24	1,861
Total	3,415	14.46	2,283	9.67	3,017	12.78	4,953	20.98	9,944	42.11	23,612

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	687	29.22	390	16.59	619	26.33	431	18.33	224	9.53	2,351
2	483	21.58	311	13.90	576	25.74	516	23.06	352	15.73	2,238
3	435	19.08	296	12.98	562	24.65	599	26.27	388	17.02	2,280
4	382	16.30	253	10.79	526	22.44	628	26.79	555	23.68	2,344
5	327	14.86	202	9.18	449	20.40	533	24.22	690	31.35	2,201
6	319	15.30	210	10.07	442	21.20	633	30.36	481	23.07	2,085
7	280	14.74	163	8.58	374	19.68	567	29.84	516	27.16	1,900
8	253	13.70	127	6.88	368	19.92	532	28.80	567	30.70	1,847
9	210	12.46	154	9.14	354	21.01	493	29.26	474	28.13	1,685
10	166	11.71	132	9.31	277	19.53	410	28.91	433	30.54	1,418
11	165	11.97	108	7.83	249	18.06	410	29.73	447	32.41	1,379
12	246	13.26	166	8.95	369	19.89	530	28.57	544	29.33	1,855
Total	3,953	16.76	2,512	10.65	5,165	21.90	6,282	26.64	5,671	24.05	23,583

4.10 Participation by Disability

4.10.1 Participation by Disability

Table 4.10.1
Participation by Disability

		Secondary Disability														
		NSD	AS	DB	DD	HI	ID	MD	OHI	OI	SED	SLD	SLI	TBI	VI	Total
Primary Disability	NPD	2,612	4	1	0	0	8	4	3	0	1	6	51	2	4	2,696
	AS	4,566	13	0	42	16	546	58	89	3	10	58	1,415	2	13	6,831
	DB	9	1	0	0	0	2	0	2	1	0	0	4	0	0	19
	DD	773	25	2	3	17	49	9	46	15	3	27	298	2	16	1,285
	HI	41	2	0	4	0	12	2	2	1	0	1	12	0	0	77
	ID	6,227	259	11	80	142	10	83	492	262	39	82	2,213	9	120	10,029
	MD	370	14	3	8	8	26	54	35	13	3	16	190	2	8	750
	OHI	638	22	6	15	22	92	12	13	18	3	18	193	1	27	1,080
	OI	55	1	1	2	0	38	3	8	0	0	5	26	0	8	147
	SED	30	1	0	2	1	9	0	4	1	0	1	8	0	0	57
	SLD	342	6	0	5	3	9	3	20	2	1	1	115	0	2	509
	SLI	126	9	1	4	0	8	1	7	0	0	10	3	0	2	171
	TBI	83	1	0	2	0	12	1	4	1	0	1	18	0	6	129
	VI	13	1	0	2	2	10	6	7	2	2	1	5	0	1	52
Total		15,885	359	25	169	211	831	236	732	319	62	227	4,551	18	207	23,832

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
ITD	Infant/Toddler with a Disability
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 1.0 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 501 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	(Ability Measure in Logits*7.913)+925.056
Reading	(Ability Measure in Logits*6.026)+925.788
Speaking	(Ability Measure in Logits*4.433)+924.531
Writing	(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 501 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
- Cronbach's coefficient alpha (as a measure of internal consistency)
- The classical standard error of measurement (SEM) in terms of ***raw scores***

[4] In Series 501, 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 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

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

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., 501).

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

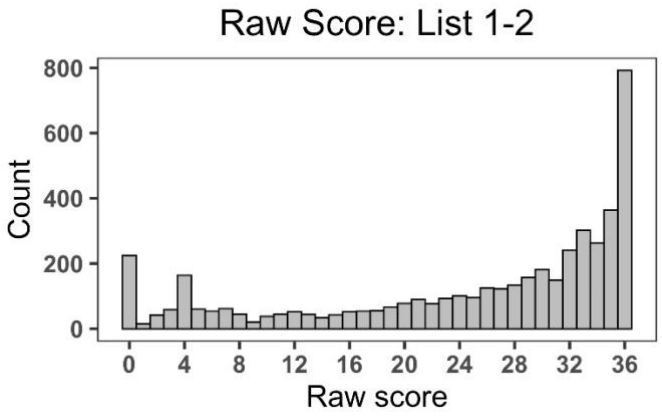


Table 6.1.1A
Raw Score Descriptive Statistics: List 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,355	0	36	23.64	11.84
2	2,243	0	36	26.30	11.03
Total	4,598	0	36	24.94	11.53

Figure 6.1.1B

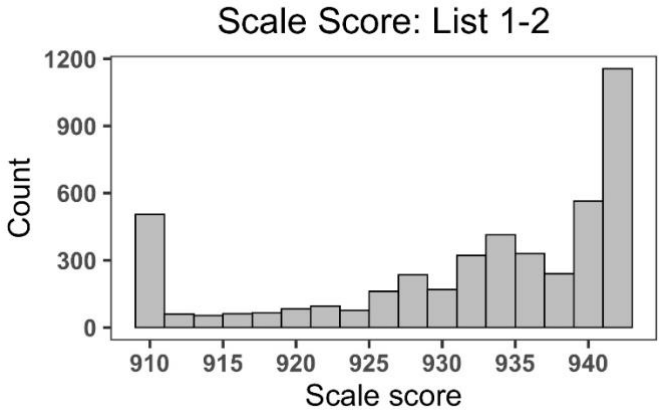


Table 6.1.1B
Scale Score Descriptive Statistics: List 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,355	910	943	931.23	11.13
2	2,243	910	943	933.75	10.33
Total	4,598	910	943	932.46	10.82

Figure 6.1.1C

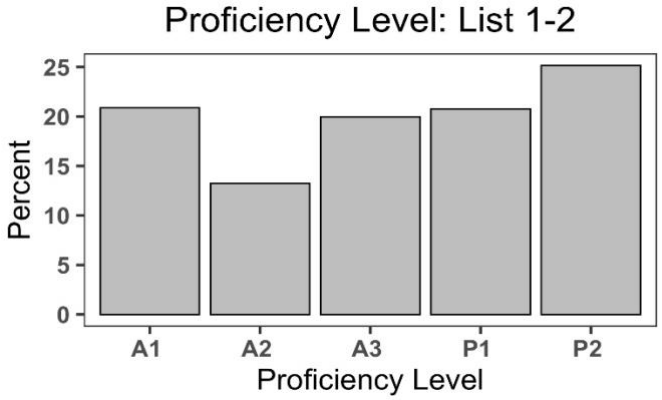


Table 6.1.1C

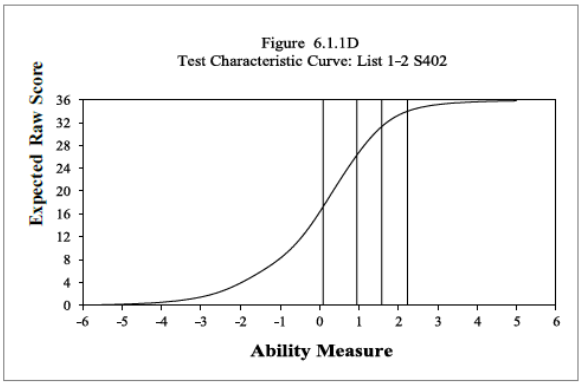
Proficiency Level Distribution: List 1-2

	Grade 1		Grade 2		Total	
Level	Count	Percent	Count	Percent	Count	Percent
A1	576	24.46	384	17.12	960	20.88
A2	339	14.39	270	12.04	609	13.24
A3	487	20.68	431	19.22	918	19.97
P1	465	19.75	490	21.85	955	20.77
P2	488	20.72	668	29.78	1,156	25.14
Total	2,355	100.00	2,243	100.00	4,598	100.00

Table 6.1.1D

Equating Summary: List 1-2

No equating summary is presented because the Alternate ACCESS Series 501 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.



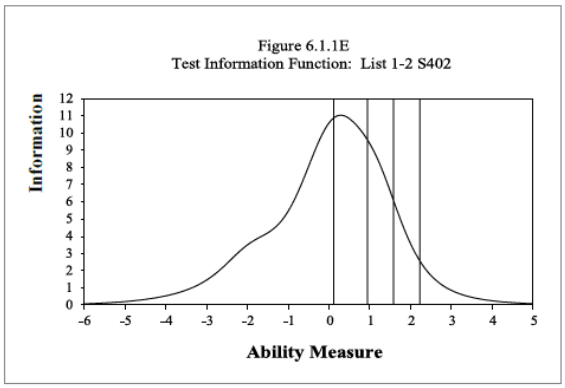


Table 6.1.1E

Reliability: List 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
4,598	9	0.938	2.685

Table 6.1.1F

Item Analysis Summary: List 1-2

Table 6.1.1G

Complete Item Analysis: List 1-2

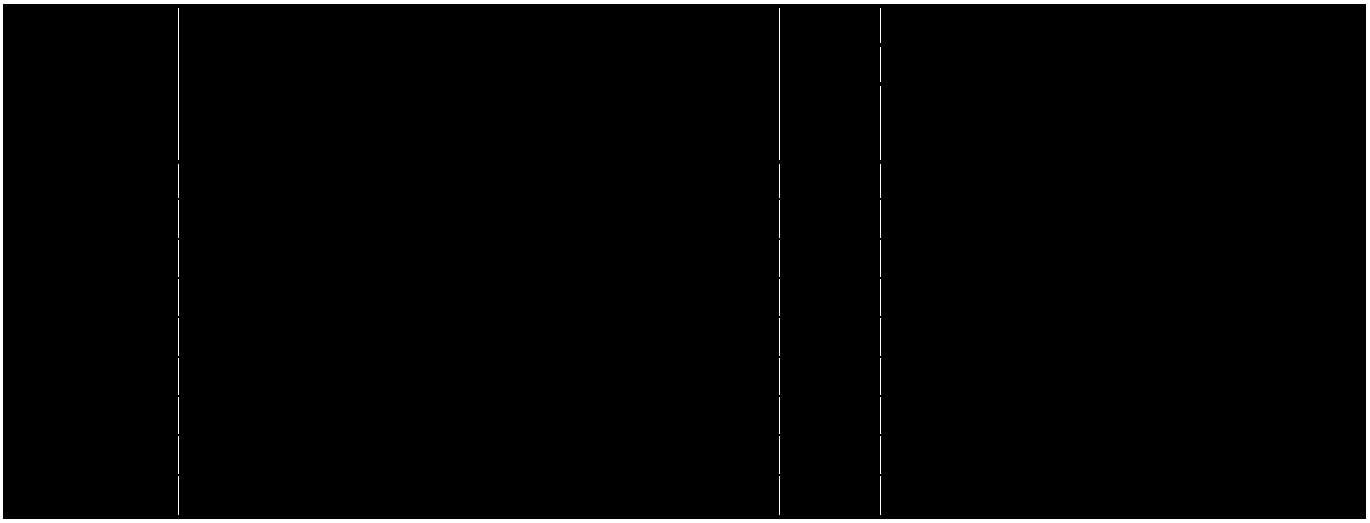


Table 6.1.1H

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 6.1.1I

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	5.65	5.65	A1	4.10	4.10
1	A1	0.38	6.03	A1	0.27	4.37
2	A1	1.02	7.05	A1	0.80	5.17
3	A1	1.57	8.62	A1	0.98	6.15
4	A1	4.54	13.16	A1	2.54	8.69
5	A1	1.44	14.61	A1	1.16	9.85
6	A1	1.23	15.84	A1	1.11	10.97
7	A1	1.49	17.32	A1	1.20	12.17
8	A1	1.10	18.43	A1	0.85	13.02
9	A1	0.47	18.90	A1	0.45	13.46
10	A1	1.06	19.96	A1	0.58	14.04
11	A1	1.10	21.06	A1	0.85	14.89
12	A1	1.44	22.51	A1	0.80	15.69
13	A1	1.10	23.61	A1	0.80	16.50
14	A1	0.85	24.46	A1	0.62	17.12
15	A2	0.93	25.39	A2	0.94	18.06
16	A2	1.19	26.58	A2	1.07	19.13
17	A2	1.23	27.81	A2	1.11	20.24
18	A2	1.44	29.26	A2	0.94	21.18
19	A2	1.36	30.62	A2	1.56	22.74
20	A2	2.21	32.82	A2	1.16	23.90
21	A2	1.83	34.65	A2	2.10	25.99
22	A2	2.04	36.69	A2	1.29	27.28
23	A2	2.17	38.85	A2	1.87	29.16
24	A3	2.25	41.10	A3	2.14	31.30
25	A3	2.38	43.48	A3	1.78	33.08
26	A3	2.46	45.94	A3	2.99	36.07
27	A3	2.93	48.87	A3	2.41	38.48
28	A3	2.85	51.72	A3	2.99	41.46
29	A3	3.44	55.16	A3	3.39	44.85
30	A3	4.37	59.53	A3	3.52	48.37
31	P1	3.52	63.06	P1	2.94	51.32
32	P1	4.50	67.56	P1	6.02	57.33
33	P1	6.50	74.06	P1	6.64	63.98
34	P1	5.22	79.28	P1	6.24	70.22
35	P2	6.79	86.07	P2	9.09	79.31
36	P2	13.93	100.00	P2	20.69	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

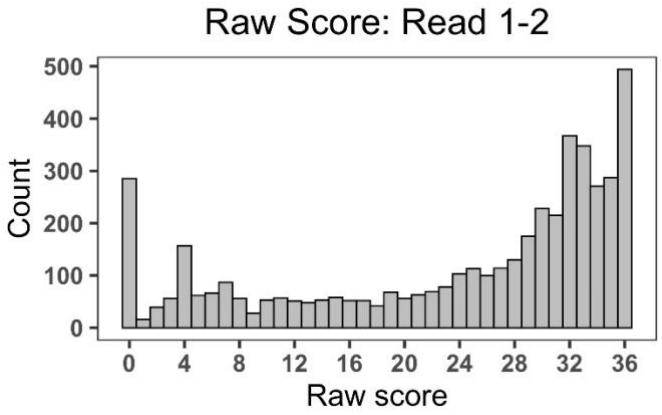


Table 6.1.2A
Raw Score Descriptive Statistics: Read 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,355	0	36	22.58	12.10
2	2,242	0	36	25.30	11.19
Total	4,597	0	36	23.91	11.74

Figure 6.1.2B

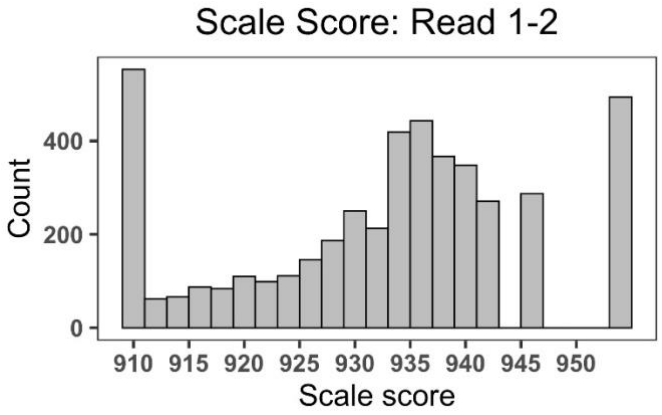


Table 6.1.2B
Scale Score Descriptive Statistics: Read 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,355	910	954	931.44	12.94
2	2,242	910	954	934.48	12.50
Total	4,597	910	954	932.92	12.82

Figure 6.1.2C

Table 6.1.2C

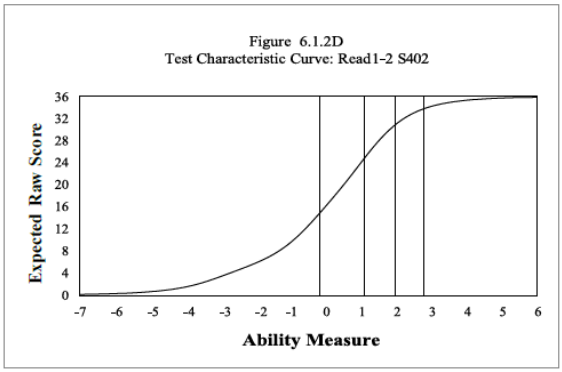
Proficiency Level Distribution: Read 1-2

	Grade 1		Grade 2		Total	
Level	Count	Percent	Count	Percent	Count	Percent
A1	623	26.45	438	19.54	1,061	23.08
A2	397	16.86	297	13.25	694	15.10
A3	429	18.22	431	19.22	860	18.71
P1	461	19.58	469	20.92	930	20.23
P2	445	18.90	607	27.07	1,052	22.88
Total	2,355	100.00	2,242	100.00	4,597	100.00

Table 6.1.2D

Equating Summary: Read 1-2

No equating summary is presented because the Alternate ACCESS Series 501 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).



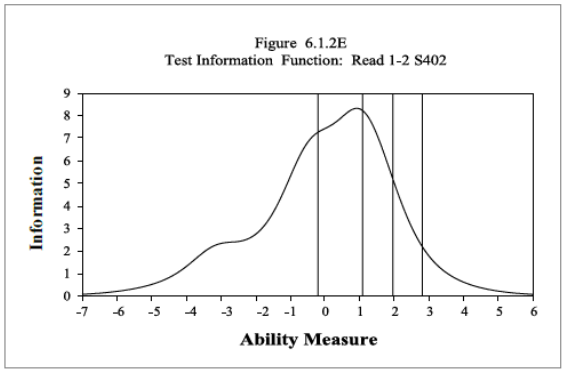


Table 6.1.2E

Reliability: Read 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
4,597	9	0.947	2.962

Table 6.1.2F

Item Analysis Summary: Read 1-2

Table 6.1.2G

Complete Item Analysis: Read 1-2

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 6.1.2I

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	7.52	7.52	A1	4.82	4.82
1	A1	0.30	7.81	A1	0.40	5.22
2	A1	1.19	9.00	A1	0.49	5.71
3	A1	1.61	10.62	A1	0.80	6.51
4	A1	3.99	14.61	A1	2.81	9.32
5	A1	1.40	16.01	A1	1.29	10.62
6	A1	1.70	17.71	A1	1.16	11.78
7	A1	2.12	19.83	A1	1.65	13.43
8	A1	1.49	21.32	A1	0.94	14.36
9	A1	0.68	22.00	A1	0.54	14.90
10	A1	1.19	23.18	A1	1.12	16.01
11	A1	1.23	24.42	A1	1.25	17.26
12	A1	1.23	25.65	A1	0.98	18.24
13	A1	0.81	26.45	A1	1.29	19.54
14	A2	1.23	27.69	A2	1.07	20.61
15	A2	1.49	29.17	A2	1.03	21.63
16	A2	1.40	30.57	A2	0.85	22.48
17	A2	1.15	31.72	A2	1.12	23.60
18	A2	1.10	32.82	A2	0.71	24.31
19	A2	1.70	34.52	A2	1.25	25.56
20	A2	1.49	36.01	A2	0.94	26.49
21	A2	1.15	37.15	A2	1.61	28.10
22	A2	1.66	38.81	A2	1.34	29.44
23	A2	1.78	40.59	A2	1.61	31.04
24	A2	2.72	43.31	A2	1.74	32.78
25	A3	2.63	45.94	A3	2.27	35.06
26	A3	2.00	47.94	A3	2.36	37.42
27	A3	2.25	50.19	A3	2.72	40.14
28	A3	2.72	52.91	A3	2.94	43.09
29	A3	3.91	56.82	A3	3.70	46.79
30	A3	4.71	61.53	A3	5.22	52.01
31	P1	4.50	66.03	P1	4.86	56.87
32	P1	7.98	74.01	P1	7.98	64.85
33	P1	7.09	81.10	P1	8.07	72.93
34	P2	4.71	85.82	P2	7.14	80.06
35	P2	5.14	90.96	P2	7.40	87.47
36	P2	9.04	100.00	P2	12.53	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

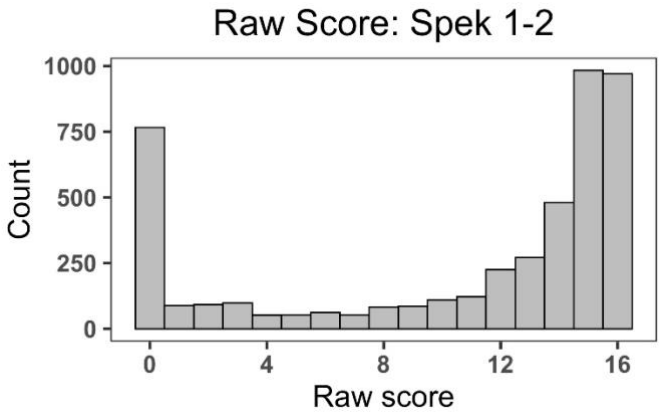


Table 6.1.3A
Raw Score Descriptive Statistics: Spek 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,351	0	16	10.12	6.13
2	2,243	0	16	11.18	5.92
Total	4,594	0	16	10.64	6.05

Figure 6.1.3B

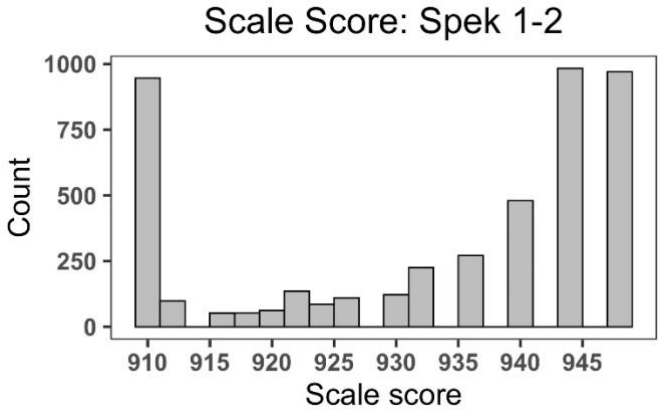


Table 6.1.3B
Scale Score Descriptive Statistics: Spek 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,351	910	948	931.60	14.51
2	2,243	910	948	934.46	14.34
Total	4,594	910	948	933.00	14.50

Figure 6.1.3C

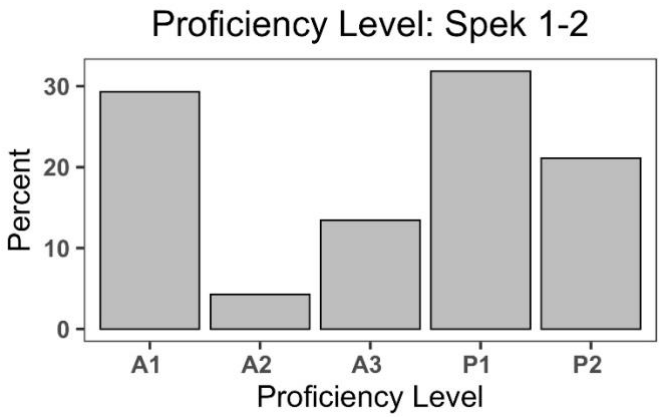


Table 6.1.3C

Proficiency Level Distribution: Spek 1-2

	Grade 1		Grade 2		Total	
Level	Count	Percent	Count	Percent	Count	Percent
A1	767	32.62	580	25.86	1,347	29.32
A2	112	4.76	84	3.74	196	4.27
A3	346	14.72	272	12.13	618	13.45
P1	742	31.56	721	32.14	1,463	31.85
P2	384	16.33	586	26.13	970	21.11
Total	2,351	100.00	2,243	100.00	4,594	100.00

Table 6.1.3D

Equating Summary: Spek 1-2

No equating summary is presented because the Alternate ACCESS Series 501 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).

Figure 6.1.3D
Test Characteristic Curve: Spek 1-2 S402

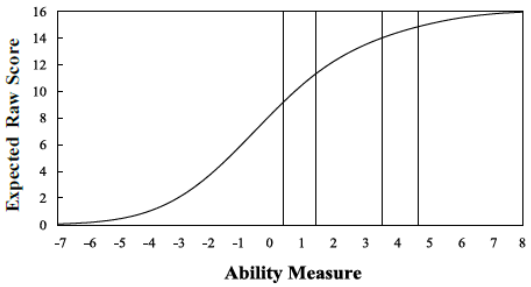


Figure 6.1.3E
Test Information Function: Spek 1-2 S402

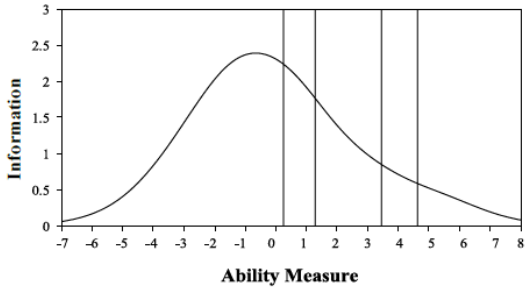


Table 6.1.3E

Reliability: Spek 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
4,594	8	0.962	2.817

Table 6.1.3F

Item Analysis Summary: Spek 1-2

Table 6.1.3G

Complete Item Analysis: Spek 1-2

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 6.1.3I

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	18.84	18.84	A1	14.40	14.40
1	A1	1.79	20.63	A1	2.05	16.45
2	A1	2.08	22.71	A1	1.92	18.37
3	A1	2.17	24.88	A1	2.10	20.46
4	A1	1.28	26.16	A1	0.98	21.44
5	A1	1.28	27.44	A1	1.03	22.47
6	A1	1.70	29.14	A1	0.98	23.45
7	A1	1.23	30.37	A1	1.07	24.52
8	A1	2.25	32.62	A1	1.34	25.86
9	A2	1.96	34.58	A2	1.78	27.64
10	A2	2.81	37.39	A2	1.96	29.60
11	A3	2.47	39.86	A3	2.85	32.46
12	A3	5.57	45.43	A3	4.19	36.65
13	A3	6.68	52.11	A3	5.08	41.73
14	P1	10.46	62.57	P1	10.43	52.16
15	P1	21.10	83.67	P1	21.71	73.87
16	P2	16.33	100.00	P2	26.13	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

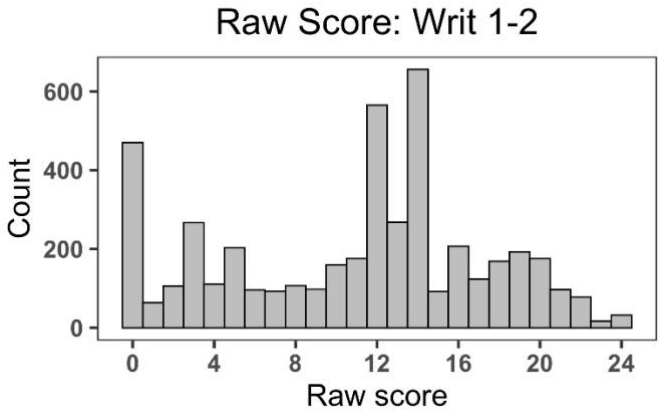


Table 6.1.4A
Raw Score Descriptive Statistics: Writ 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,368	0	24	10.22	6.33
2	2,257	0	24	11.62	6.33
Total	4,625	0	24	10.90	6.37

Figure 6.1.4B

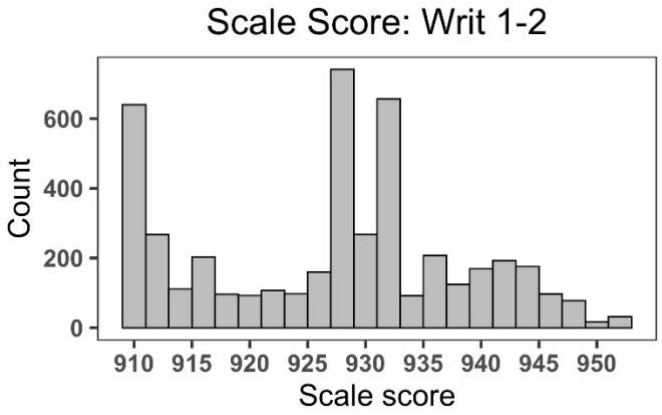


Table 6.1.4B
Scale Score Descriptive Statistics: Writ 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,368	910	953	926.56	11.24
2	2,257	910	953	929.08	11.37
Total	4,625	910	953	927.79	11.37

Figure 6.1.4C

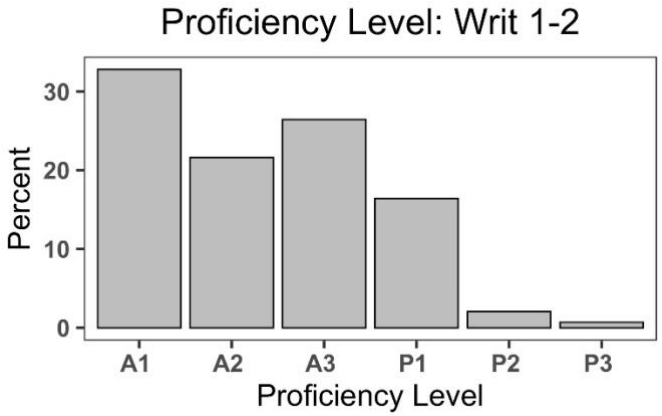


Table 6.1.4C

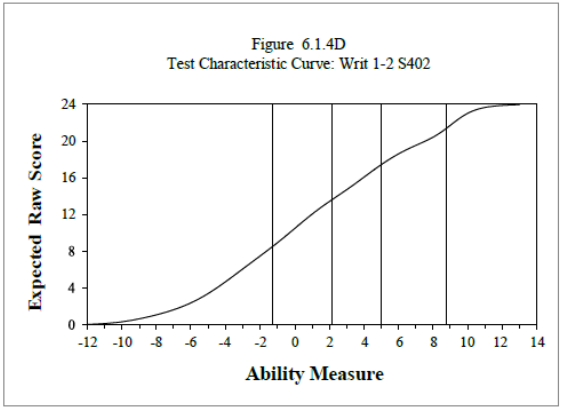
Proficiency Level Distribution: Writ 1-2

	Grade 1		Grade 2		Total	
Level	Count	Percent	Count	Percent	Count	Percent
A1	871	36.78	646	28.62	1,517	32.80
A2	511	21.58	488	21.62	999	21.60
A3	614	25.93	609	26.98	1,223	26.44
P1	327	13.81	432	19.14	759	16.41
P2	30	1.27	65	2.88	95	2.05
P3	15	0.63	17	0.75	32	0.69
Total	2,368	100.00	2,257	100.00	4,625	100.00

Table 6.1.4D

Equating Summary: Writ 1-2

No equating summary is presented because the Alternate ACCESS Series 501 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).



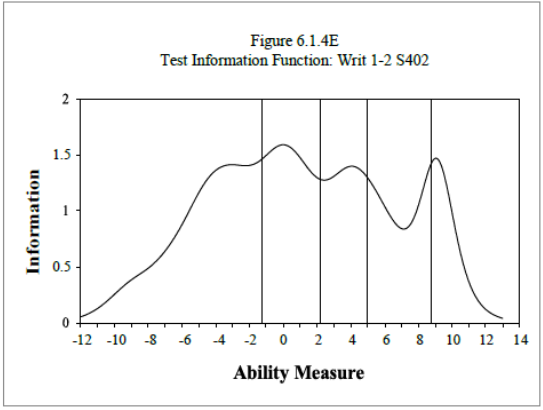


Table 6.1.4E

Reliability: Writ 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
4,625	10	0.929	3.025

Table 6.1.4F

Item Analysis Summary: Writ 1-2

Table 6.1.4G

Complete Item Analysis: Writ 1-2

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	11.57	11.57	A1	8.68	8.68
1	A1	1.73	13.30	A1	1.02	9.70
2	A1	2.58	15.88	A1	1.99	11.70
3	A1	6.29	22.17	A1	5.23	16.93
4	A1	2.87	25.04	A1	1.91	18.83
5	A1	4.56	29.60	A1	4.21	23.04
6	A1	2.41	32.01	A1	1.73	24.77
7	A1	2.36	34.38	A1	1.64	26.41
8	A1	2.41	36.78	A1	2.22	28.62
9	A2	2.36	39.15	A2	1.86	30.48
10	A2	3.55	42.69	A2	3.37	33.85
11	A2	3.51	46.20	A2	4.12	37.97
12	A2	12.16	58.36	A2	12.27	50.24
13	A3	6.33	64.70	A3	5.23	55.47
14	A3	13.81	78.51	A3	14.58	70.05
15	A3	2.07	80.57	A3	1.91	71.95
16	A3	3.72	84.29	A3	5.27	77.23
17	P1	2.58	86.87	P1	2.79	80.02
18	P1	2.79	89.65	P1	4.56	84.58
19	P1	3.55	93.20	P1	4.83	89.41
20	P1	2.91	96.11	P1	4.74	94.15
21	P1	1.98	98.10	P1	2.22	96.37
22	P2	1.18	99.28	P2	2.22	98.58
23	P2	0.08	99.37	P2	0.66	99.25
24	P3	0.63	100.00	P3	0.75	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

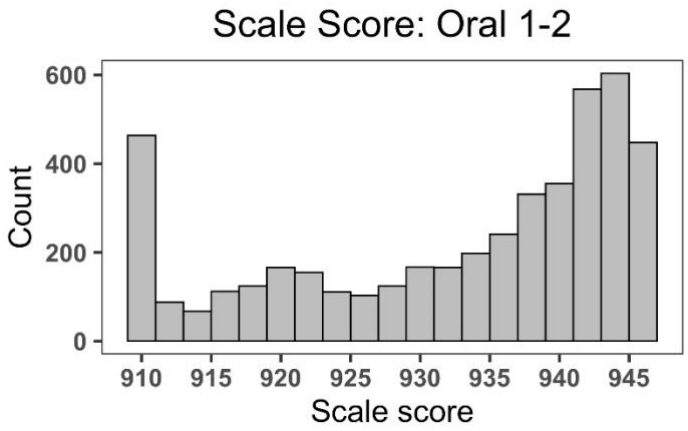


Table 6.1.5A

Scale Score Descriptive Statistics: Oral 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,351	910	946	931.67	12.06
2	2,241	910	946	934.37	11.66
Total	4,592	910	946	932.99	11.95

Figure 6.1.5B

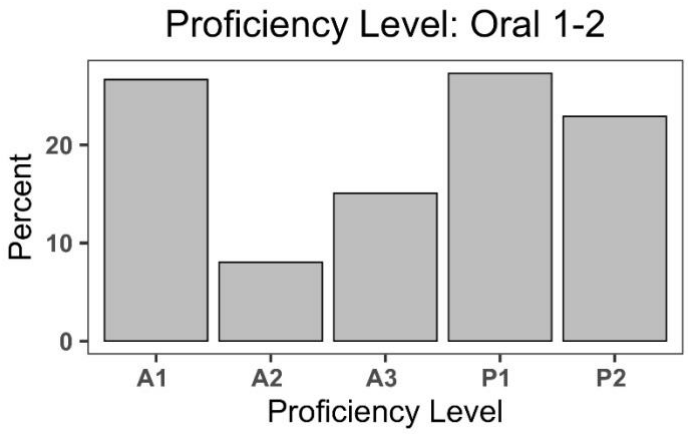


Table 6.1.5C

Proficiency Level Distribution: Oral 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	703	29.90	522	23.29	1,225	26.68
A2	214	9.10	155	6.92	369	8.04
A3	377	16.04	315	14.06	692	15.07
P1	638	27.14	616	27.49	1,254	27.31
P2	419	17.82	633	28.25	1,052	22.91
Total	2,351	100.00	2,241	100.00	4,592	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	117.087	0.938
Speaking	0.5	210.324	0.962
Oral		142.775	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

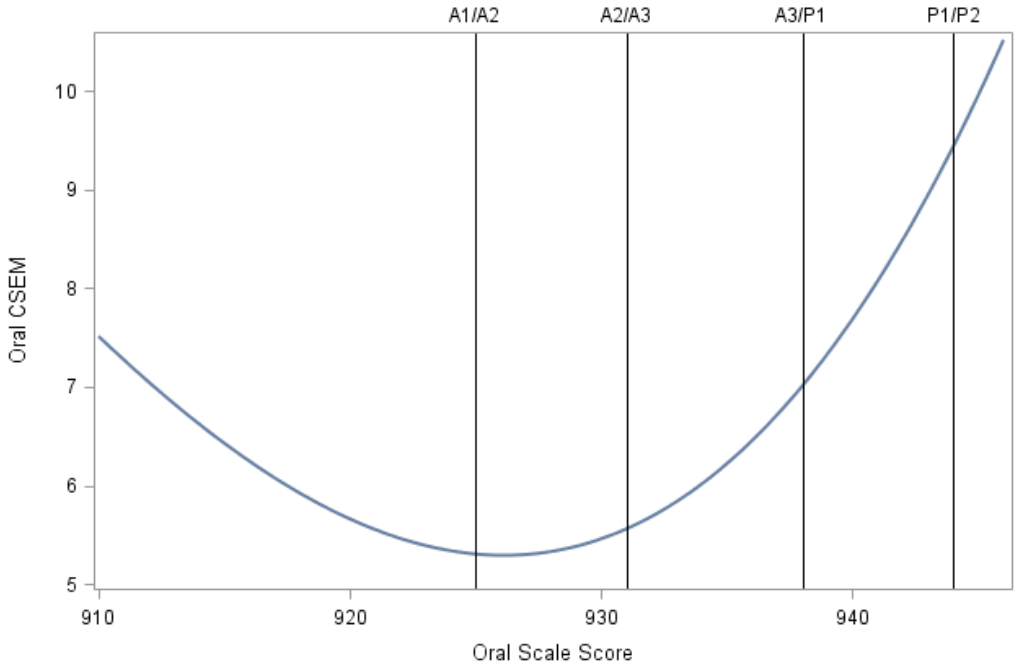
Table 6.1.5J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

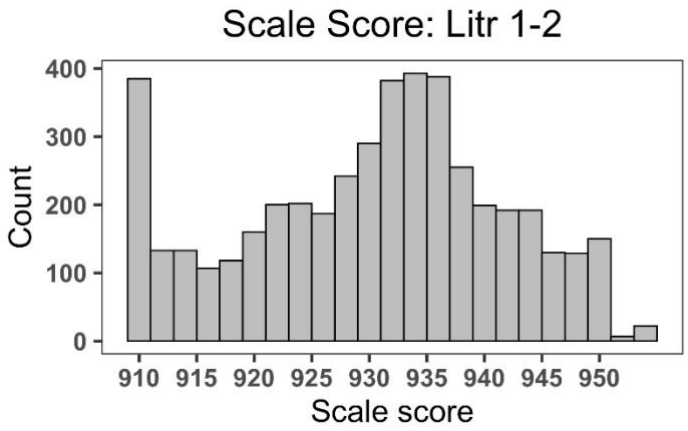


Table 6.1.6A
Scale Score Descriptive Statistics: Litr1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,354	910	954	929.28	11.30
2	2,242	910	954	932.07	11.15
Total	4,596	910	954	930.64	11.31

Figure 6.1.6B

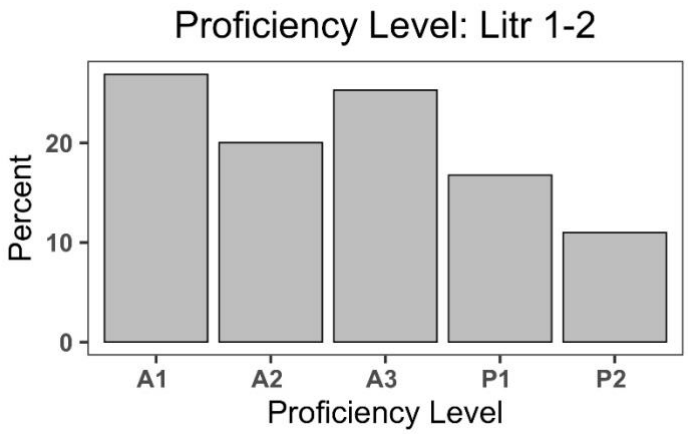


Table 6.1.6C

Proficiency Level Distribution: Litr 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	730	31.01	506	22.57	1,236	26.89
A2	491	20.86	430	19.18	921	20.04
A3	590	25.06	573	25.56	1,163	25.30
P1	346	14.70	424	18.91	770	16.75
P2	197	8.37	309	13.78	506	11.01
Total	2,354	100.00	2,242	100.00	4,596	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	164.618	0.947
Writing	0.5	129.285	0.929
Literacy		128.207	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

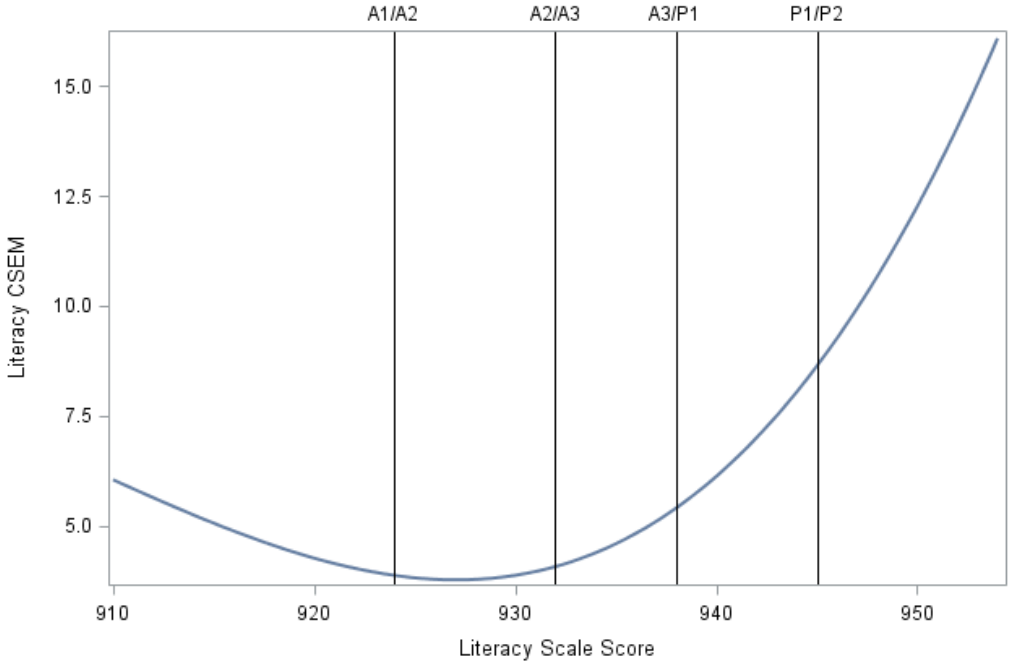
Table 6.1.6I
n/a

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

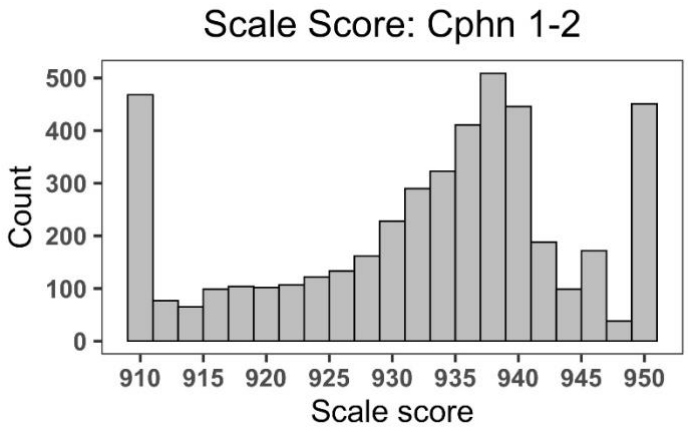


Table 6.1.7A

Scale Score Descriptive Statistics: Cphn 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,354	910	951	931.44	12.06
2	2,240	910	951	934.33	11.52
Total	4,594	910	951	932.85	11.88

Figure 6.1.7B

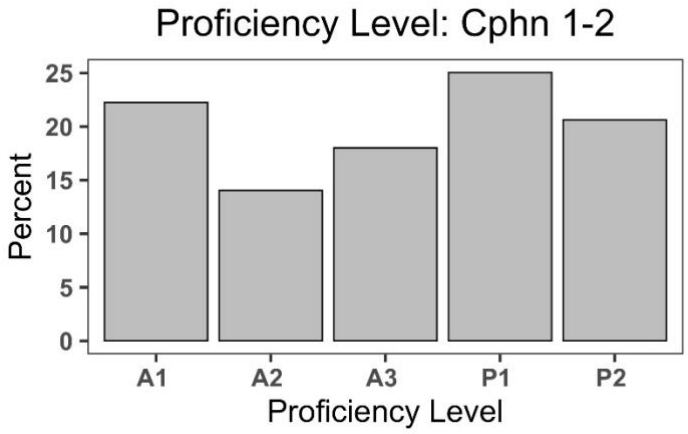


Table 6.1.7C

Proficiency Level Distribution: Cphn 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	606	25.74	416	18.57	1,022	22.25
A2	365	15.51	280	12.50	645	14.04
A3	429	18.22	399	17.81	828	18.02
P1	554	23.53	597	26.65	1,151	25.05
P2	400	16.99	548	24.46	948	20.64
Total	2,354	100.00	2,240	100.00	4,594	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	117.087	0.938
Reading	0.7	164.618	0.947
Comprehension		141.375	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

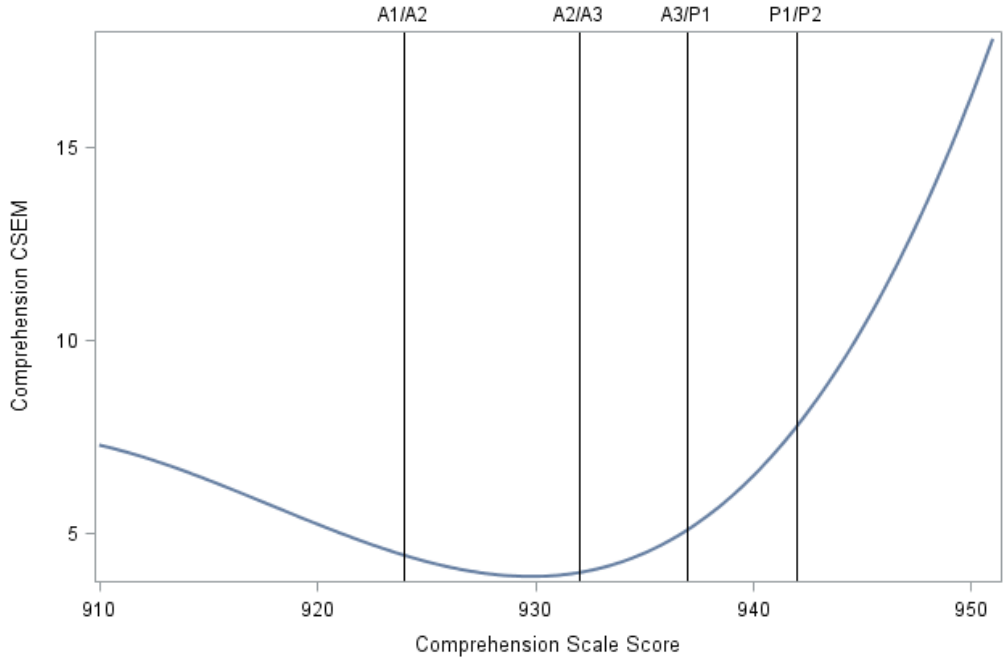
Table 6.1.7J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

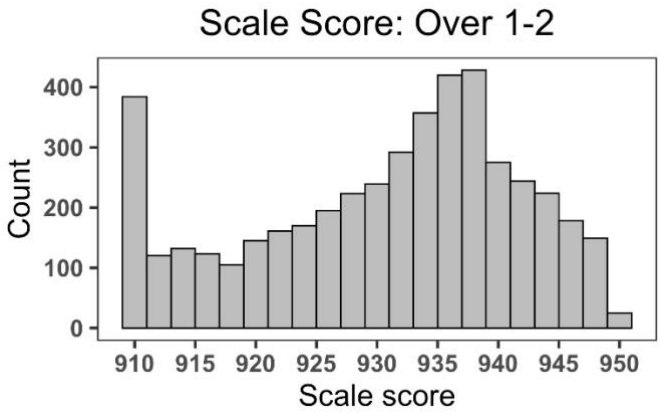


Table 6.1.8A

Scale Score Descriptive Statistics: Over 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2,351	910	951	929.75	11.12
2	2,238	910	951	932.50	10.86
Total	4,589	910	951	931.09	11.08

Figure 6.1.8B

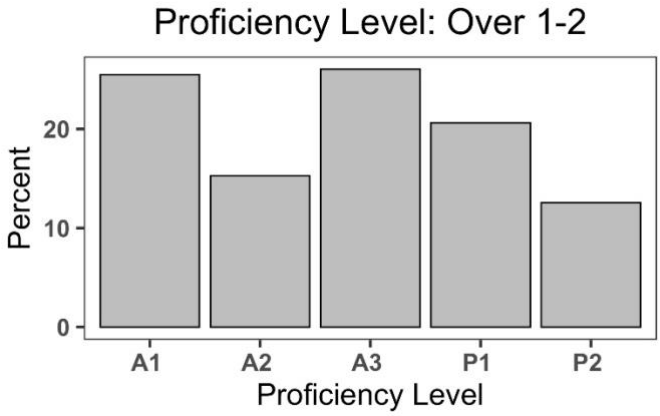


Table 6.1.8C

Proficiency Level Distribution: Over 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	687	29.22	483	21.58	1,170	25.50
A2	390	16.59	311	13.90	701	15.28
A3	619	26.33	576	25.74	1,195	26.04
P1	431	18.33	516	23.06	947	20.64
P2	224	9.53	352	15.73	576	12.55
Total	2,351	100.00	2,238	100.00	4,589	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	117.087	0.938
Reading	0.35	164.618	0.947
Speaking	0.15	210.324	0.962
Writing	0.35	129.285	0.929
Overall Composite		122.790	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

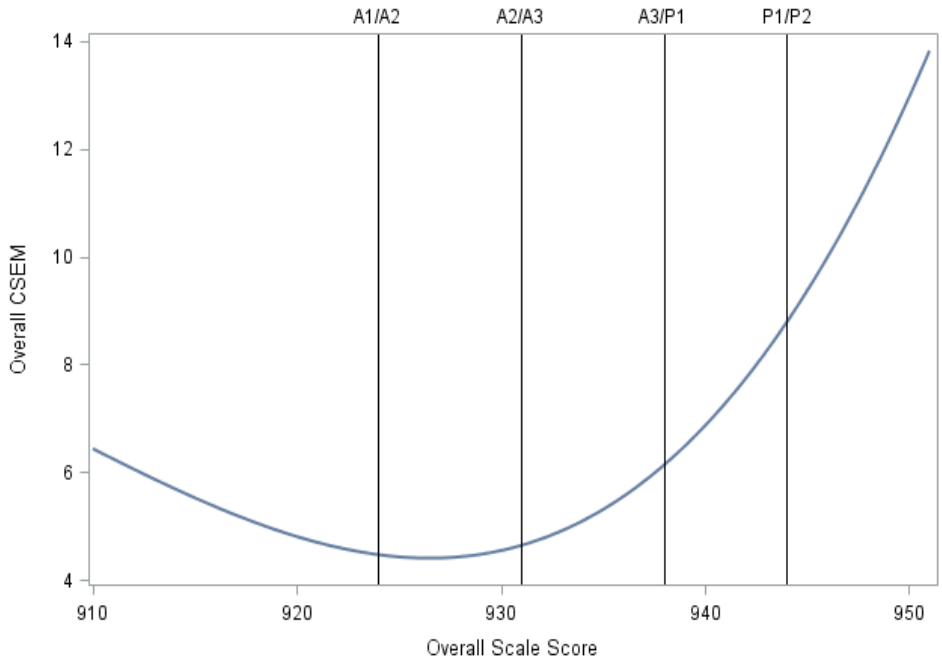
Table 6.1.8I
n/a

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

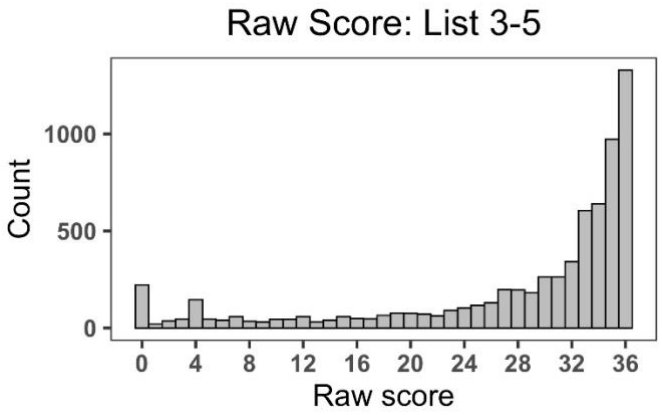


Table 6.2.1A
Raw Score Descriptive Statistics: List 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,284	0	36	26.98	10.44
4	2,349	0	36	28.15	10.30
5	2,206	0	36	29.08	9.69
Total	6,839	0	36	28.06	10.19

Figure 6.2.1B

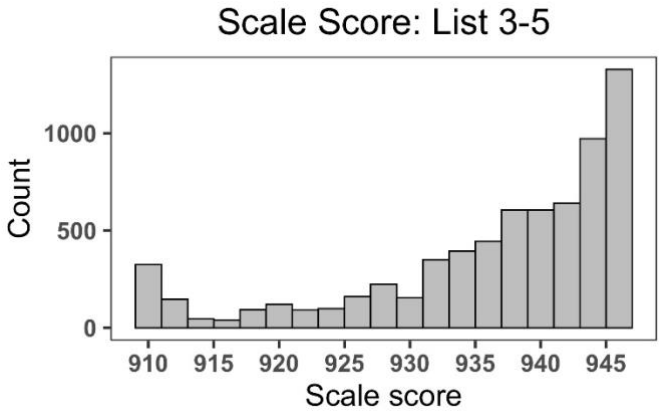


Table 6.2.1B
Scale Score Descriptive Statistics: List 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,284	910	947	935.82	10.51
4	2,349	910	947	937.21	10.52
5	2,206	910	947	938.23	10.01
Total	6,839	910	947	937.08	10.40

Figure 6.2.1C

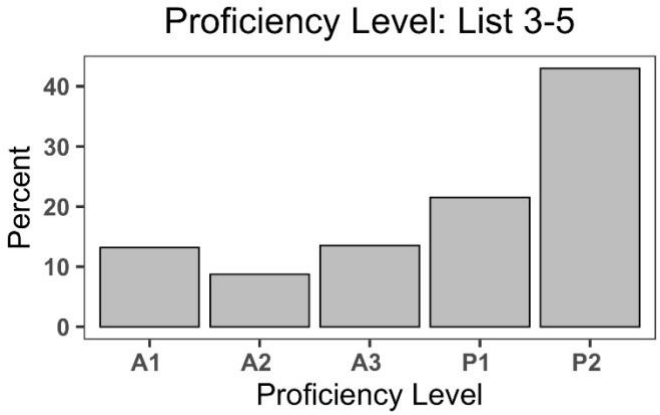


Table 6.2.1C

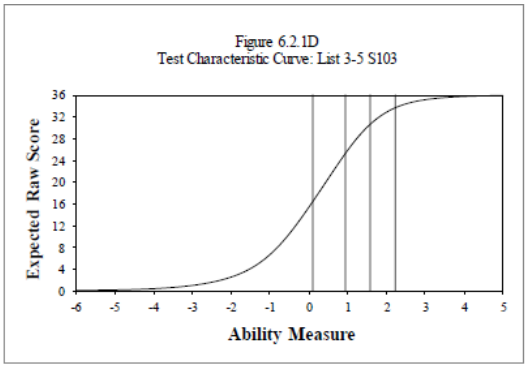
Proficiency Level Distribution: List 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	345	15.11	311	13.24	246	11.15	902	13.19
A2	243	10.64	194	8.26	161	7.30	598	8.74
A3	336	14.71	316	13.45	274	12.42	926	13.54
P1	548	23.99	493	20.99	432	19.58	1,473	21.54
P2	812	35.55	1,035	44.06	1,093	49.55	2,940	42.99
Total	2,284	100.00	2,349	100.00	2,206	100.00	6,839	100.00

Table 6.2.1D

Equating Summary: List 3-5

No equating summary is presented because the Alternate ACCESS Series 501 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.



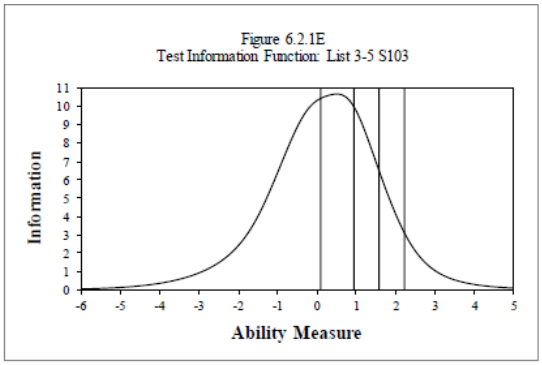


Table 6.2.1E

Reliability: List 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
6,839	9	0.938	2.598

Table 6.2.1F

Item Analysis Summary: List 3-5

[redacted]

Table 6.2.1G

Complete Item Analysis: List 3-5

[redacted]

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 6.2.1I

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	3.15	3.15	A1	3.79	3.79	A1	2.77	2.77
1	A1	0.39	3.55	A1	0.34	4.13	A1	0.14	2.90
2	A1	0.48	4.03	A1	0.60	4.73	A1	0.54	3.45
3	A1	1.01	5.04	A1	0.47	5.19	A1	0.54	3.99
4	A1	2.54	7.57	A1	2.04	7.24	A1	1.81	5.80
5	A1	0.88	8.45	A1	0.60	7.83	A1	0.54	6.35
6	A1	0.61	9.06	A1	0.47	8.30	A1	0.68	7.03
7	A1	0.96	10.03	A1	0.94	9.24	A1	0.68	7.71
8	A1	0.79	10.81	A1	0.38	9.62	A1	0.32	8.02
9	A1	0.44	11.25	A1	0.64	10.26	A1	0.27	8.30
10	A1	0.74	12.00	A1	0.64	10.90	A1	0.59	8.88
11	A1	0.74	12.74	A1	0.72	11.62	A1	0.50	9.38
12	A1	1.18	13.92	A1	0.64	12.26	A1	0.77	10.15
13	A1	0.61	14.54	A1	0.34	12.60	A1	0.45	10.61
14	A1	0.57	15.11	A1	0.64	13.24	A1	0.54	11.15
15	A2	1.05	16.16	A2	0.89	14.13	A2	0.63	11.79
16	A2	0.96	17.12	A2	0.55	14.69	A2	0.63	12.42
17	A2	0.83	17.95	A2	0.55	15.24	A2	0.68	13.10
18	A2	1.27	19.22	A2	0.89	16.13	A2	0.68	13.78
19	A2	1.36	20.58	A2	0.94	17.07	A2	1.09	14.87
20	A2	1.49	22.07	A2	0.85	17.92	A2	0.95	15.82
21	A2	1.05	23.12	A2	1.19	19.11	A2	0.91	16.73
22	A2	0.92	24.04	A2	0.98	20.09	A2	0.86	17.59
23	A2	1.71	25.74	A2	1.40	21.50	A2	0.86	18.45
24	A3	1.80	27.54	A3	1.28	22.78	A3	1.45	19.90
25	A3	2.10	29.64	A3	1.66	24.44	A3	1.36	21.26
26	A3	2.06	31.70	A3	2.21	26.65	A3	1.41	22.67
27	A3	2.76	34.46	A3	2.89	29.54	A3	3.04	25.70
28	A3	3.20	37.65	A3	2.60	32.14	A3	2.81	28.51
29	A3	2.80	40.46	A3	2.81	34.95	A3	2.36	30.87
30	P1	4.64	45.10	P1	3.66	38.61	P1	3.22	34.09
31	P1	4.47	49.56	P1	3.70	42.32	P1	3.35	37.44
32	P1	5.30	54.86	P1	5.19	47.51	P1	4.49	41.93
33	P1	9.59	64.45	P1	8.43	55.94	P1	8.52	50.45
34	P2	8.76	73.20	P2	8.90	64.84	P2	10.47	60.92
35	P2	12.87	86.08	P2	13.88	78.71	P2	15.96	76.88
36	P2	13.92	100.00	P2	21.29	100.00	P2	23.12	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

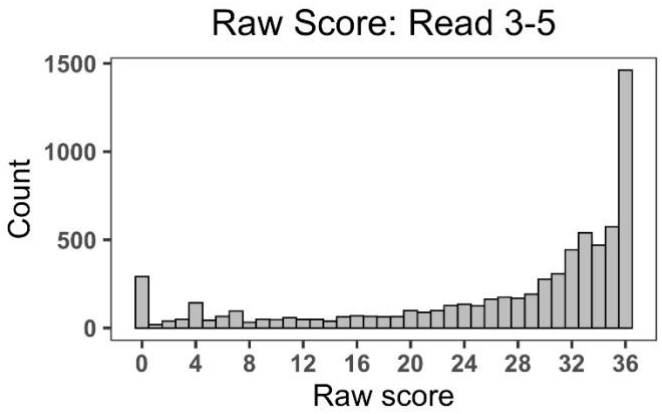


Table 6.2.2A
Raw Score Descriptive Statistics: Read 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,285	0	36	25.68	10.94
4	2,351	0	36	27.05	10.77
5	2,205	0	36	28.09	10.39
Total	6,841	0	36	26.93	10.75

Figure 6.2.2B

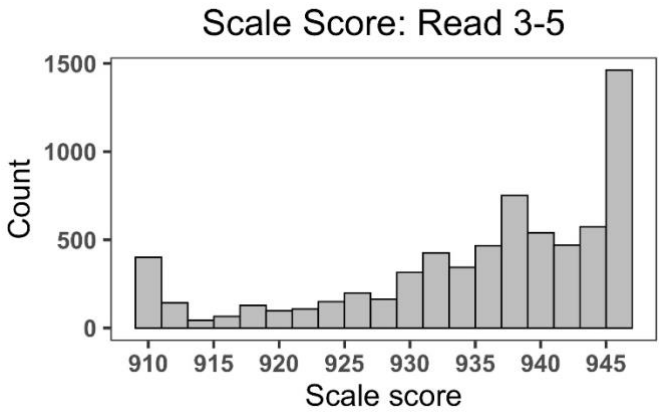


Table 6.2.2B
Scale Score Descriptive Statistics: Read 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,285	910	947	934.57	10.85
4	2,351	910	947	936.07	10.84
5	2,205	910	947	937.30	10.62
Total	6,841	910	947	935.97	10.83

Figure 6.2.2B

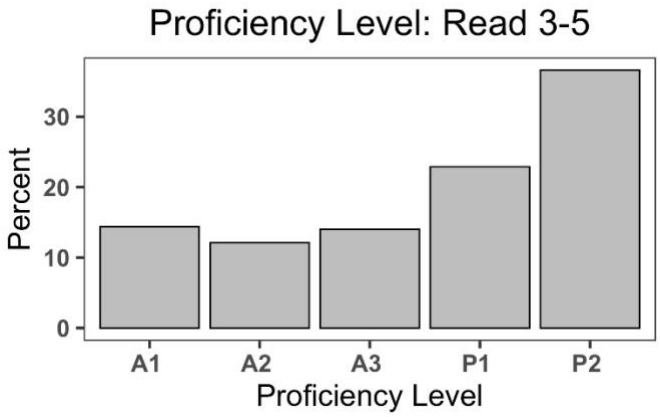


Table 6.2.2C

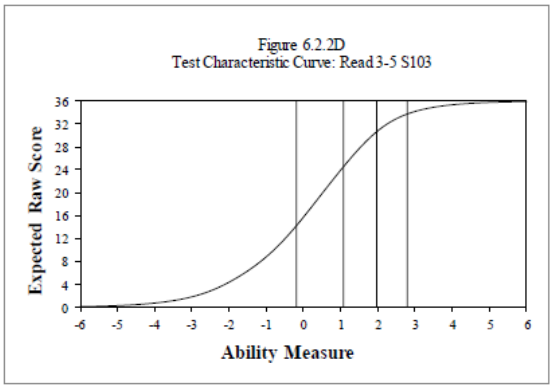
Proficiency Level Distribution: Read 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	371	16.24	335	14.25	278	12.61	984	14.38
A2	327	14.31	274	11.65	227	10.29	828	12.10
A3	380	16.63	308	13.10	272	12.34	960	14.03
P1	547	23.94	558	23.73	461	20.91	1,566	22.89
P2	660	28.88	876	37.26	967	43.85	2,503	36.59
Total	2,285	100.00	2,351	100.00	2,205	100.00	6,841	100.00

Table 6.2.2D

Equating Summary: Read 3-5

No equating summary is presented because the Alternate ACCESS Series 501 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.



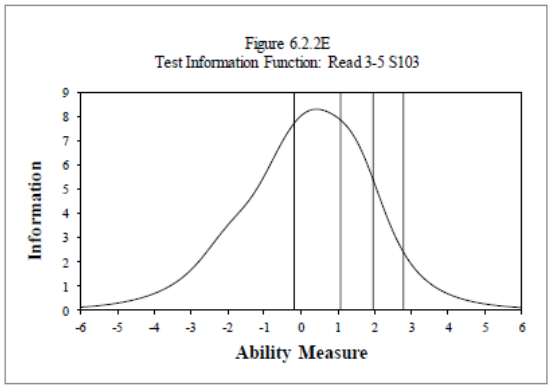


Table 6.2.2E

Reliability: Read 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
6,841	9	0.950	2.417

Table 6.2.2F

Item Analysis Summary: Read 3-5

Table 6.2.2G

Complete Item Analysis: Read 3-5

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	4.42	4.42	A1	4.68	4.68	A1	3.63	3.63
1	A1	0.26	4.68	A1	0.43	5.10	A1	0.14	3.76
2	A1	0.74	5.43	A1	0.51	5.61	A1	0.50	4.26
3	A1	1.01	6.43	A1	0.55	6.17	A1	0.63	4.90
4	A1	2.23	8.67	A1	2.21	8.38	A1	1.77	6.67
5	A1	0.57	9.23	A1	0.51	8.89	A1	0.86	7.53
6	A1	1.36	10.59	A1	0.72	9.61	A1	0.82	8.34
7	A1	1.62	12.21	A1	1.36	10.97	A1	1.22	9.57
8	A1	0.70	12.91	A1	0.47	11.44	A1	0.23	9.80
9	A1	0.70	13.61	A1	0.60	12.04	A1	0.91	10.70
10	A1	0.70	14.31	A1	0.72	12.76	A1	0.63	11.34
11	A1	1.05	15.36	A1	0.85	13.61	A1	0.63	11.97
12	A1	0.88	16.24	A1	0.64	14.25	A1	0.63	12.61
13	A2	1.09	17.33	A2	0.64	14.89	A2	0.41	13.02
14	A2	0.74	18.07	A2	0.51	15.40	A2	0.41	13.42
15	A2	1.53	19.61	A2	0.68	16.08	A2	0.54	13.97
16	A2	1.23	20.83	A2	0.89	16.97	A2	0.95	14.92
17	A2	1.05	21.88	A2	0.98	17.95	A2	0.86	15.78
18	A2	1.14	23.02	A2	0.94	18.89	A2	0.68	16.46
19	A2	1.01	24.03	A2	0.85	19.74	A2	1.00	17.46
20	A2	1.62	25.65	A2	1.36	21.10	A2	1.36	18.82
21	A2	1.66	27.31	A2	1.19	22.29	A2	1.04	19.86
22	A2	1.44	28.75	A2	1.62	23.90	A2	1.27	21.13
23	A2	1.79	30.55	A2	2.00	25.90	A2	1.77	22.90
24	A3	2.80	33.35	A3	1.49	27.39	A3	1.63	24.54
25	A3	2.01	35.36	A3	2.00	29.39	A3	1.50	26.03
26	A3	2.67	38.03	A3	2.76	32.16	A3	1.72	27.76
27	A3	2.58	40.61	A3	2.34	34.50	A3	2.77	30.52
28	A3	3.15	43.76	A3	2.08	36.58	A3	2.18	32.70
29	A3	3.41	47.18	A3	2.42	39.00	A3	2.54	35.24
30	P1	4.16	51.33	P1	4.55	43.56	P1	3.36	38.59
31	P1	4.99	56.32	P1	4.42	47.98	P1	4.08	42.68
32	P1	6.35	62.67	P1	7.40	55.38	P1	5.62	48.30
33	P1	8.45	71.12	P1	7.36	62.74	P1	7.85	56.15
34	P2	6.00	77.11	P2	7.49	70.23	P2	7.07	63.22
35	P2	6.91	84.03	P2	8.97	79.20	P2	9.25	72.47
36	P2	15.97	100.00	P2	20.80	100.00	P2	27.53	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

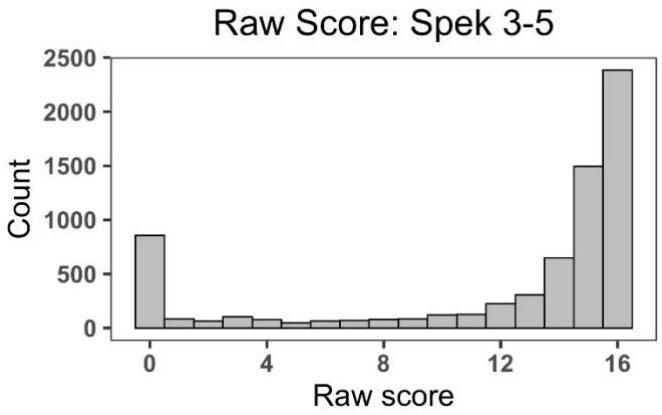


Table 6.2.3A
Raw Score Descriptive Statistics: Spek 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,283	0	16	11.65	5.76
4	2,347	0	16	12.11	5.58
5	2,202	0	16	12.35	5.53
Total	6,832	0	16	12.03	5.63

Figure 6.2.3B

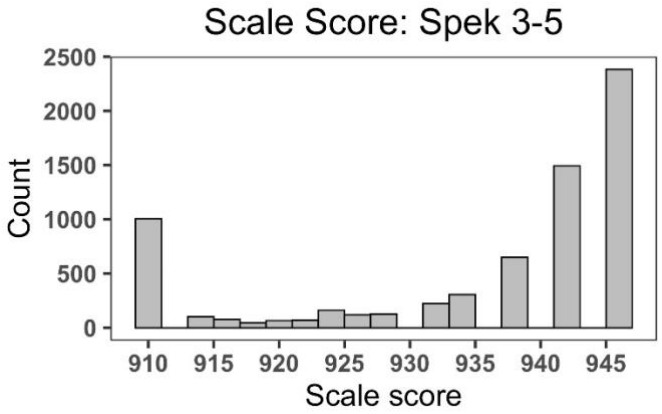


Table 6.2.3B
Scale Score Descriptive Statistics: Spek 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,283	910	947	935.07	13.50
4	2,347	910	947	936.31	13.22
5	2,202	910	947	937.02	13.14
Total	6,832	910	947	936.12	13.31

Figure 6.2.3C

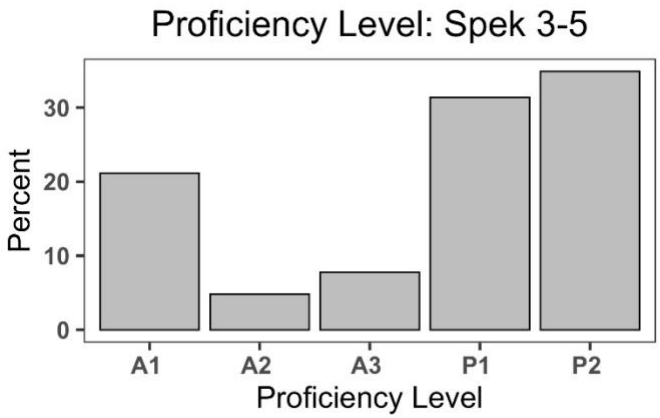


Table 6.2.3C

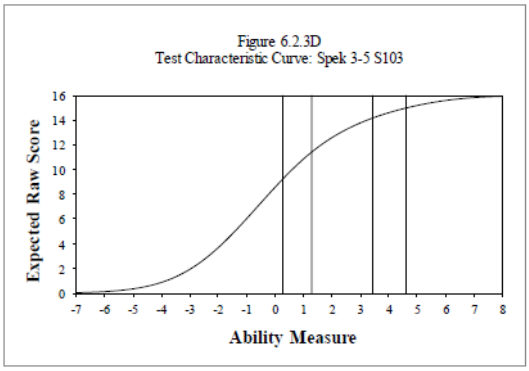
Proficiency Level Distribution: Spek 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	531	23.26	483	20.58	431	19.57	1,445	21.15
A2	116	5.08	122	5.20	92	4.18	330	4.83
A3	207	9.07	175	7.46	149	6.77	531	7.77
P1	749	32.81	734	31.27	660	29.97	2,143	31.37
P2	680	29.79	833	35.49	870	39.51	2,383	34.88
Total	2,283	100.00	2,347	100.00	2,202	100.00	6,832	100.00

Table 6.2.3D

Equating Summary: Spek 3-5

No equating summary is presented because the Alternate ACCESS Series 501 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).



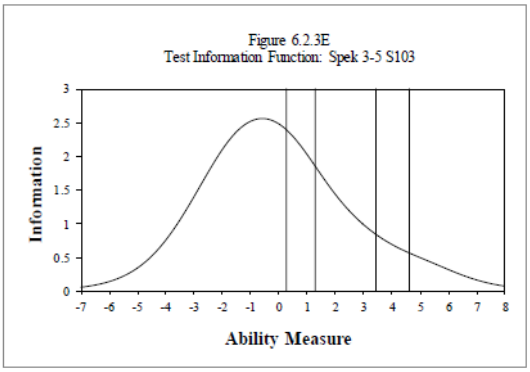


Table 6.2.3E

Reliability: Spek 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
6,832	8	0.966	2.441

Table 6.2.3F

Item Analysis Summary: Spek 3-5

Table 6.2.3G

Complete Item Analysis: Spek 3-5

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	13.67	13.67	A1	12.27	12.27	A1	11.63	11.63
1	A1	1.40	15.07	A1	1.07	13.34	A1	1.23	12.85
2	A1	1.01	16.08	A1	0.85	14.19	A1	1.00	13.85
3	A1	1.58	17.65	A1	1.75	15.94	A1	1.18	15.03
4	A1	1.40	19.05	A1	1.02	16.96	A1	0.95	15.99
5	A1	0.74	19.80	A1	0.68	17.64	A1	0.64	16.62
6	A1	1.23	21.02	A1	0.68	18.32	A1	1.00	17.62
7	A1	1.01	22.03	A1	0.98	19.30	A1	1.04	18.66
8	A1	1.23	23.26	A1	1.28	20.58	A1	0.91	19.57
9	A2	1.31	24.57	A2	1.28	21.86	A2	1.09	20.66
10	A2	1.71	26.28	A2	1.96	23.82	A2	1.63	22.30
11	A2	2.06	28.34	A2	1.96	25.78	A2	1.45	23.75
12	A3	4.16	32.50	A3	2.85	28.63	A3	2.82	26.57
13	A3	4.91	37.41	A3	4.60	33.23	A3	3.95	30.52
14	P1	10.69	48.09	P1	9.42	42.65	P1	8.36	38.87
15	P1	22.12	70.21	P1	21.86	64.51	P1	21.62	60.49
16	P2	29.79	100.00	P2	35.49	100.00	P2	39.51	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

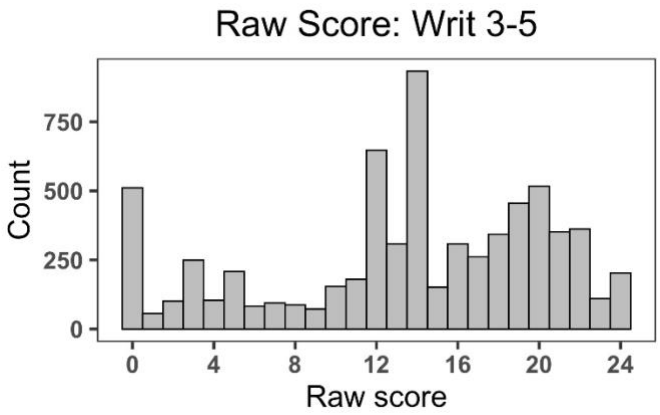


Table 6.2.4A
Raw Score Descriptive Statistics: Writ 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,299	0	24	12.78	6.62
4	2,348	0	24	13.57	6.71
5	2,203	0	24	14.36	6.79
Total	6,850	0	24	13.56	6.74

Figure 6.2.4B

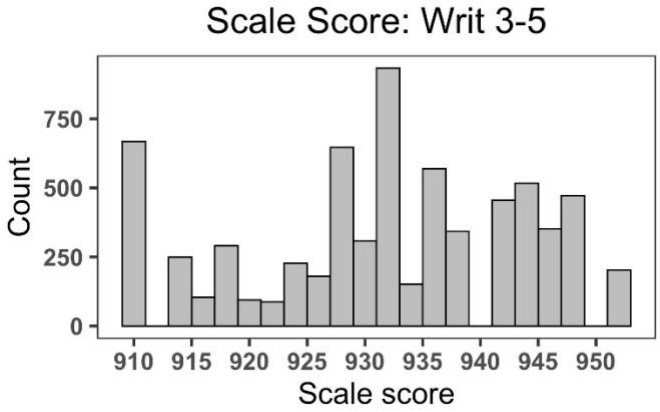


Table 6.2.4B
Scale Score Descriptive Statistics: Writ 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,299	910	953	930.66	11.58
4	2,348	910	953	932.13	11.84
5	2,203	910	953	933.62	12.12
Total	6,850	910	953	932.12	11.91

Figure 6.2.4C

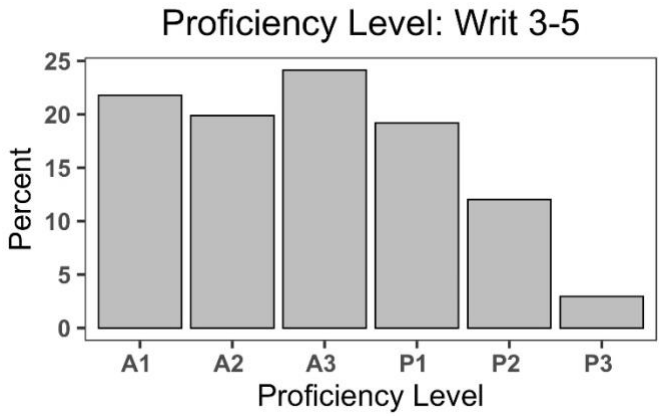


Table 6.2.4C

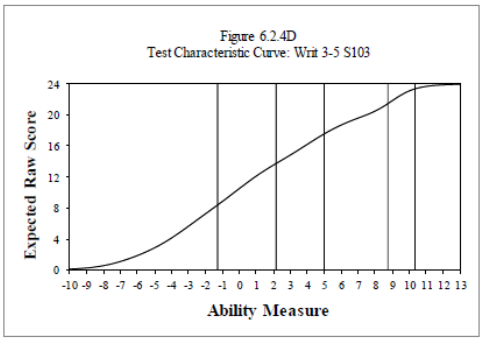
Proficiency Level Distribution: Writ 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	566	24.62	505	21.51	422	19.16	1,493	21.80
A2	516	22.44	448	19.08	398	18.07	1,362	19.88
A3	566	24.62	603	25.68	484	21.97	1,653	24.13
P1	404	17.57	446	18.99	465	21.11	1,315	19.20
P2	195	8.48	278	11.84	351	15.93	824	12.03
P3	52	2.26	68	2.90	83	3.77	203	2.96
Total	2,299	100.00	2,348	100.00	2,203	100.00	6,850	100.00

Table 6.2.4D

Equating Summary: Writ 3-5

No equating summary is presented because the Alternate ACCESS Series 501 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).



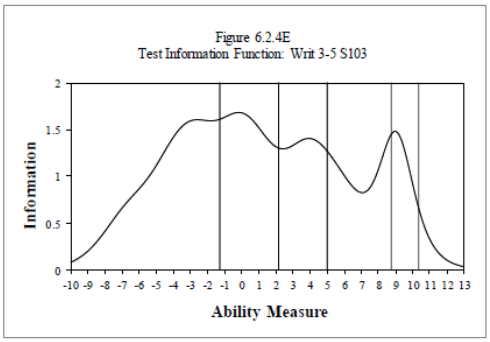


Table 6.2.4E

Reliability: Writ 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
6,850	10	0.923	3.302

Table 6.2.4F

Item Analysis Summary: Writ 3-5

Table 6.2.4G

Complete Item Analysis: Writ 3-5

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	8.09	8.09	A1	7.79	7.79	A1	6.45	6.45
1	A1	1.13	9.22	A1	0.55	8.35	A1	0.77	7.22
2	A1	1.70	10.92	A1	1.49	9.84	A1	1.23	8.44
3	A1	3.52	14.44	A1	3.71	13.54	A1	3.68	12.12
4	A1	1.65	16.09	A1	1.49	15.03	A1	1.41	13.53
5	A1	3.35	19.44	A1	2.64	17.67	A1	3.13	16.66
6	A1	1.74	21.18	A1	0.94	18.61	A1	0.95	17.61
7	A1	1.91	23.10	A1	1.24	19.85	A1	0.95	18.57
8	A1	1.52	24.62	A1	1.66	21.51	A1	0.59	19.16
9	A2	1.13	25.75	A2	1.15	22.66	A2	0.91	20.06
10	A2	2.57	28.32	A2	2.43	25.09	A2	1.72	21.79
11	A2	3.22	31.54	A2	2.51	27.60	A2	2.13	23.92
12	A2	10.57	42.11	A2	8.73	36.33	A2	9.03	32.96
13	A2	4.96	47.06	A2	4.26	40.59	A2	4.27	37.22
14	A3	13.79	60.85	A3	14.35	54.94	A3	12.66	49.89
15	A3	2.09	62.94	A3	2.56	57.50	A3	1.95	51.84
16	A3	4.74	67.68	A3	4.73	62.22	A3	3.99	55.83
17	A3	4.00	71.68	A3	4.05	66.27	A3	3.36	59.19
18	P1	4.96	76.64	P1	5.11	71.38	P1	4.95	64.14
19	P1	7.26	83.91	P1	6.26	77.64	P1	6.40	70.54
20	P1	5.35	89.26	P1	7.62	85.26	P1	9.76	80.30
21	P2	3.65	92.91	P2	5.24	90.50	P2	6.58	86.88
22	P2	3.78	96.69	P2	5.37	95.87	P2	6.76	93.65
23	P2	1.04	97.74	P2	1.24	97.10	P2	2.59	96.23
24	P3	2.26	100.00	P3	2.90	100.00	P3	3.77	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

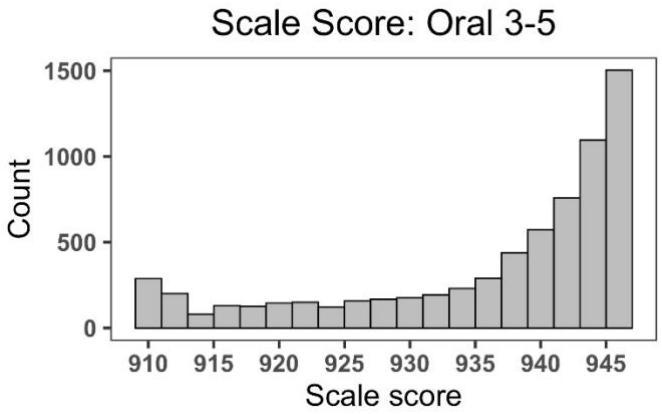


Table 6.2.5A

Scale Score Descriptive Statistics: Oral 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,281	910	947	935.58	11.16
4	2,345	910	947	936.88	11.12
5	2,201	910	947	937.75	10.85
Total	6,827	910	947	936.73	11.08

Figure 6.2.5B

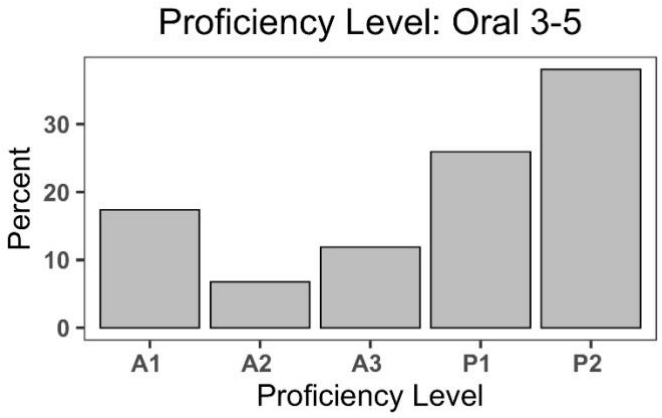


Table 6.2.5C

Proficiency Level Distribution: Oral 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	447	19.60	403	17.19	337	15.31	1,187	17.39
A2	174	7.63	149	6.35	138	6.27	461	6.75
A3	312	13.68	272	11.60	227	10.31	811	11.88
P1	637	27.93	607	25.88	526	23.90	1,770	25.93
P2	711	31.17	914	38.98	973	44.21	2,598	38.05
Total	2,281	100.00	2,345	100.00	2,201	100.00	6,827	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	107.806	0.938
Speaking	0.5	177.295	0.966
Oral		122.808	0.974

*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

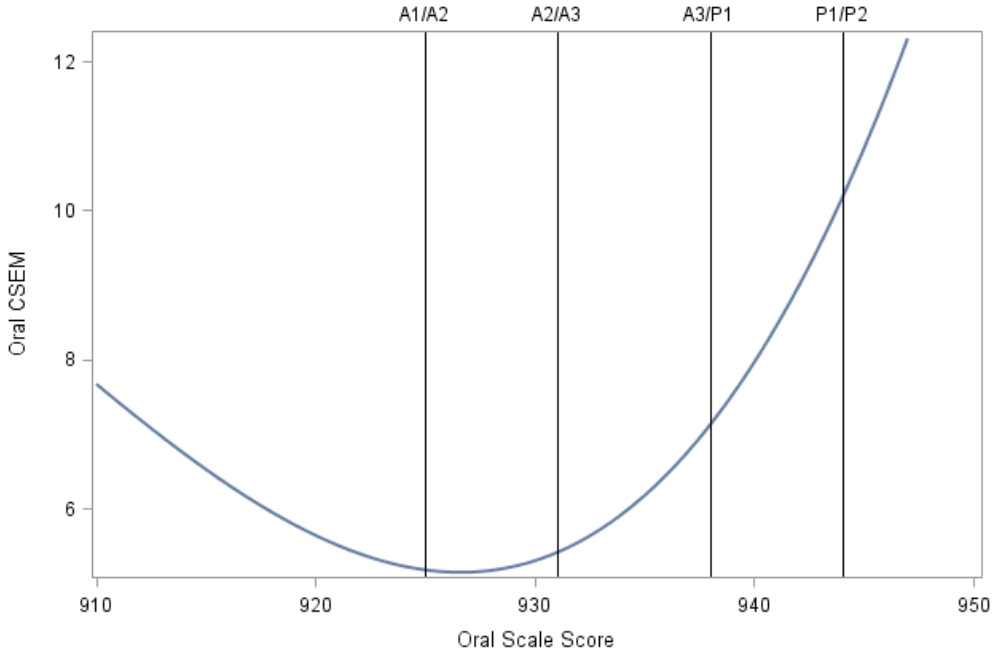
Table 6.2.5I
n/a

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		**Accuracy**		
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

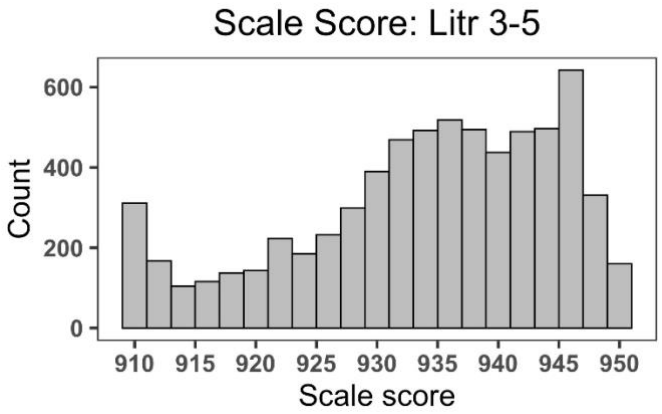


Table 6.2.6A

Scale Score Descriptive Statistics: Litr 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,285	910	950	932.91	10.60
4	2,348	910	950	934.40	10.75
5	2,203	910	950	935.74	10.81
Total	6,836	910	950	934.33	10.78

Figure 6.2.6B

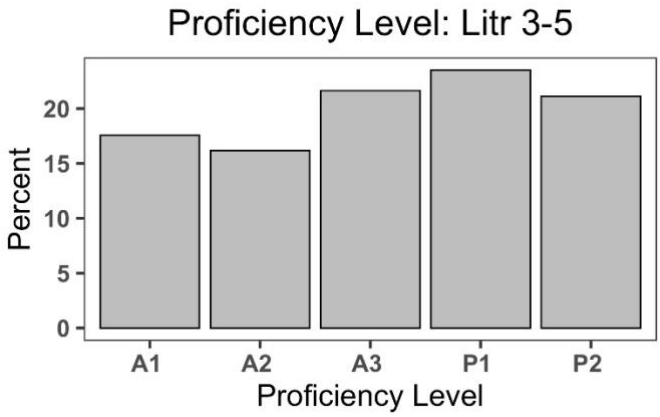


Table 6.2.6C

Proficiency Level Distribution: Litr 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	462	20.22	399	16.99	340	15.43	1,201	17.57
A2	424	18.56	375	15.97	307	13.94	1,106	16.18
A3	530	23.19	514	21.89	435	19.75	1,479	21.64
P1	521	22.80	569	24.23	516	23.42	1,606	23.49
P2	348	15.23	491	20.91	605	27.46	1,444	21.12
Total	2,285	100.00	2,348	100.00	2,203	100.00	6,836	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	117.115	0.950
Writing	0.5	141.767	0.923
Literacy		116.205	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

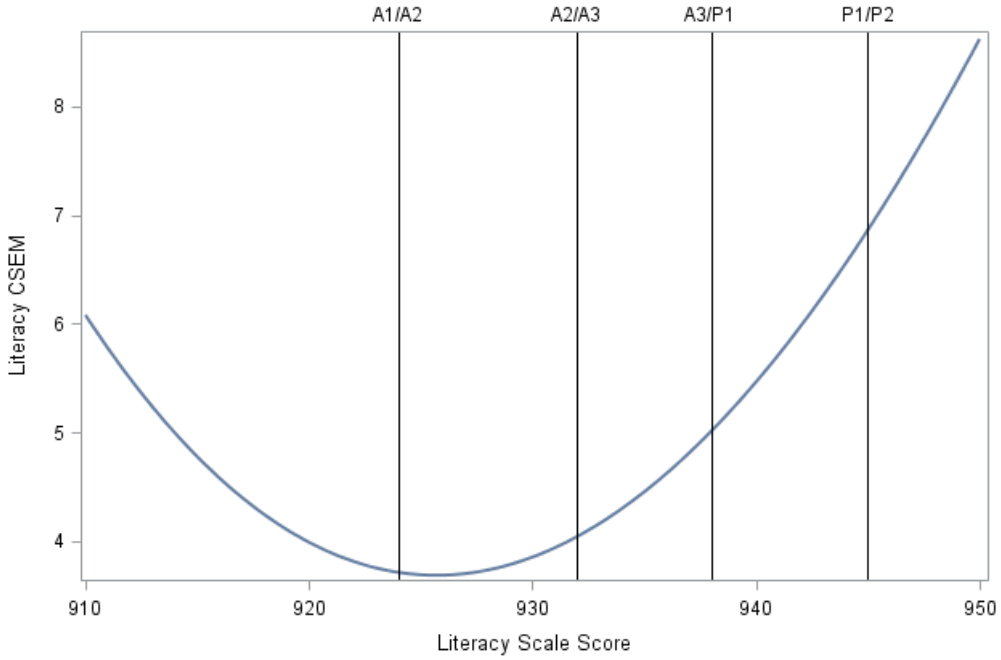
Table 6.2.6I
n/a

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

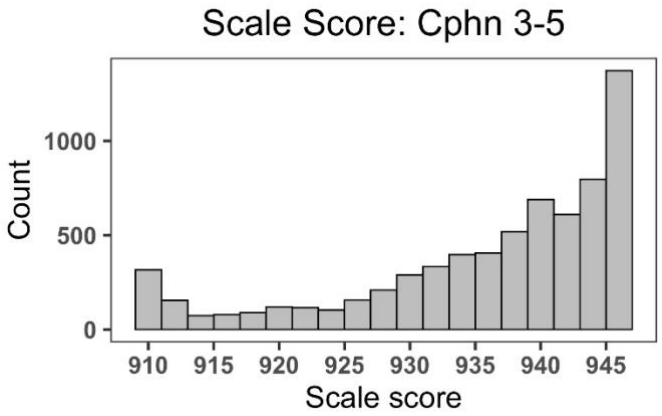


Table 6.2.7A

Scale Score Descriptive Statistics: Cphn 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,283	910	947	934.98	10.44
4	2,349	910	947	936.44	10.46
5	2,204	910	947	937.60	10.14
Total	6,836	910	947	936.32	10.41

Figure 6.2.7B

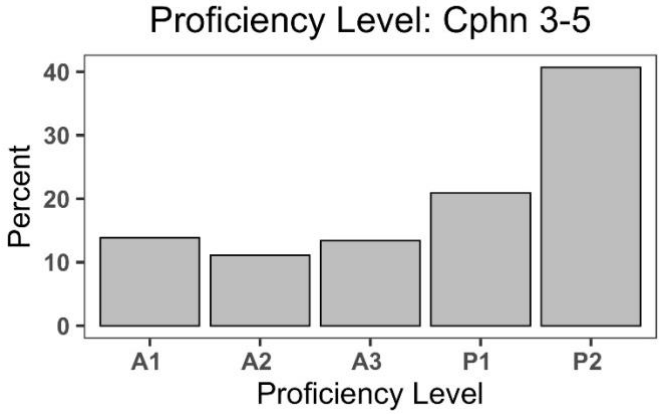


Table 6.2.7C

Proficiency Level Distribution: Cphn 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	361	15.81	321	13.67	267	12.11	949	13.88
A2	317	13.89	249	10.60	193	8.76	759	11.10
A3	334	14.63	301	12.81	281	12.75	916	13.40
P1	518	22.69	513	21.84	400	18.15	1,431	20.93
P2	753	32.98	965	41.08	1,063	48.23	2,781	40.68
Total	2,283	100.00	2,349	100.00	2,204	100.00	6,836	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	107.806	0.938
Reading	0.7	117.115	0.950
Comprehension		108.121	0.968

*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

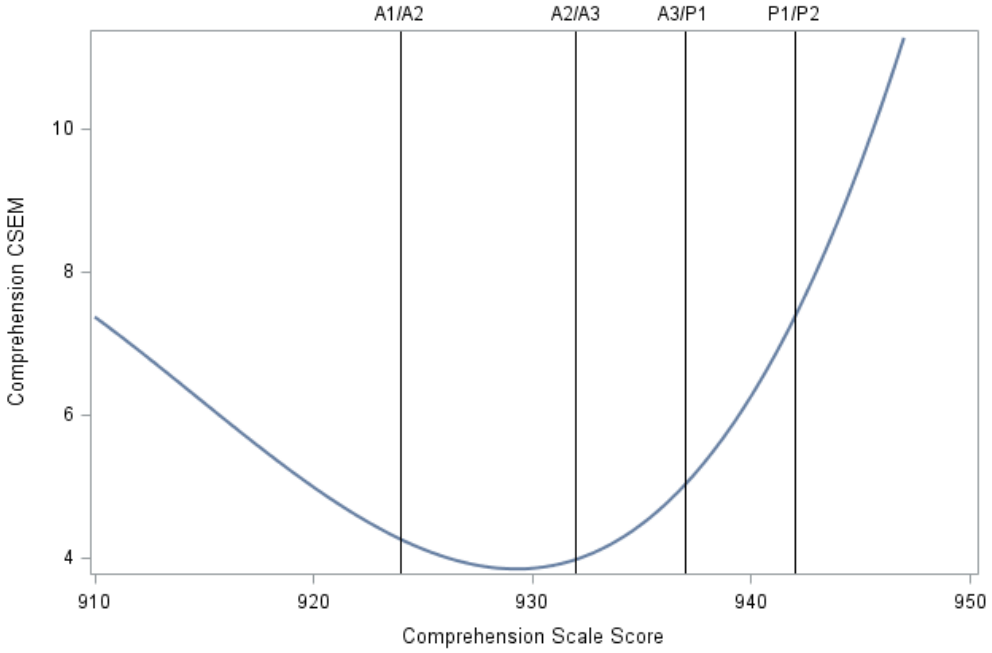
Table 6.2.7J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

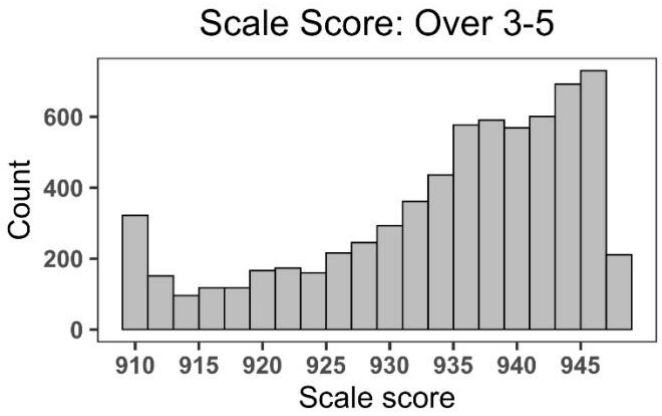


Table 6.2.8A

Scale Score Descriptive Statistics: Over 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2,280	910	949	933.49	10.45
4	2,344	910	949	934.93	10.54
5	2,201	910	949	936.13	10.51
Total	6,825	910	949	934.84	10.55

Figure 6.2.8A

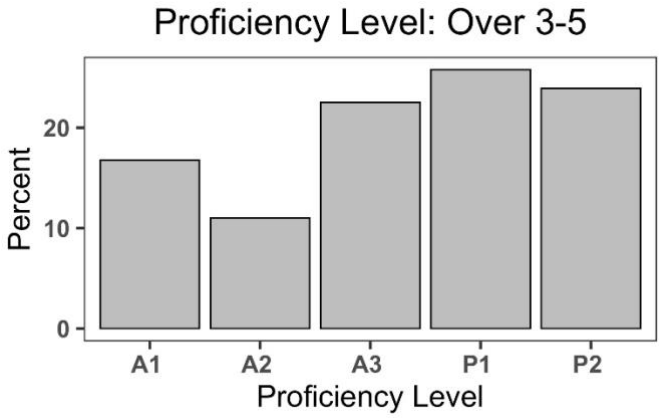


Table 6.2.8C

Proficiency Level Distribution: Over 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	435	19.08	382	16.30	327	14.86	1,144	16.76
A2	296	12.98	253	10.79	202	9.18	751	11.00
A3	562	24.65	526	22.44	449	20.40	1,537	22.52
P1	599	26.27	628	26.79	533	24.22	1,760	25.79
P2	388	17.02	555	23.68	690	31.35	1,633	23.93
Total	2,280	100.00	2,344	100.00	2,201	100.00	6,825	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	107.806	0.938
Reading	0.35	117.115	0.950
Speaking	0.15	177.295	0.966
Writing	0.35	141.767	0.923
Overall Composite		111.413	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

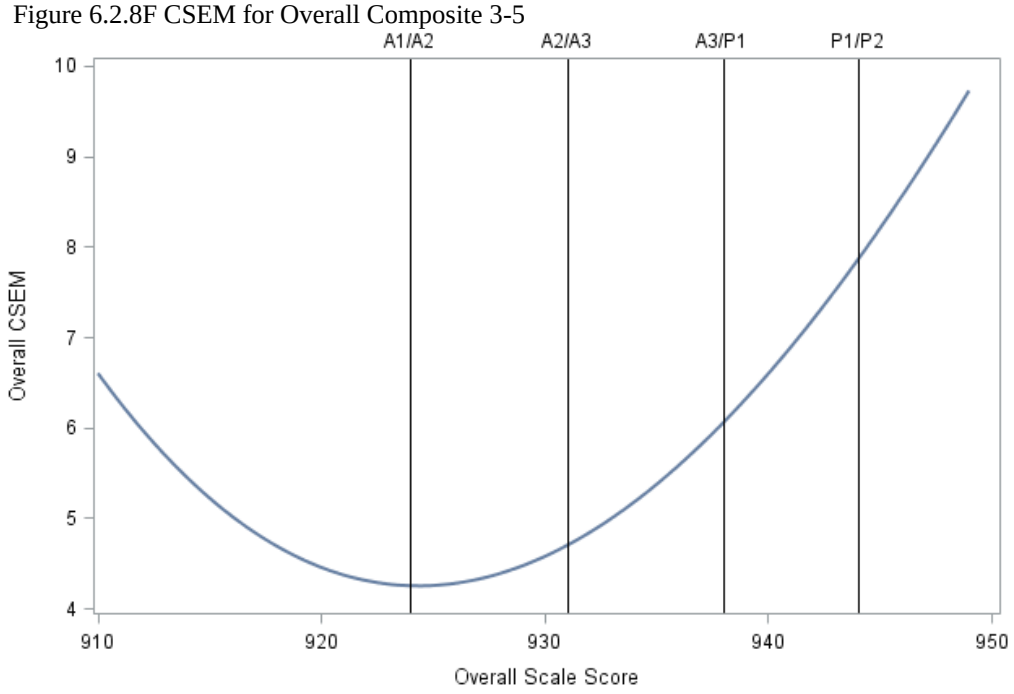
Table 6.2.8J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

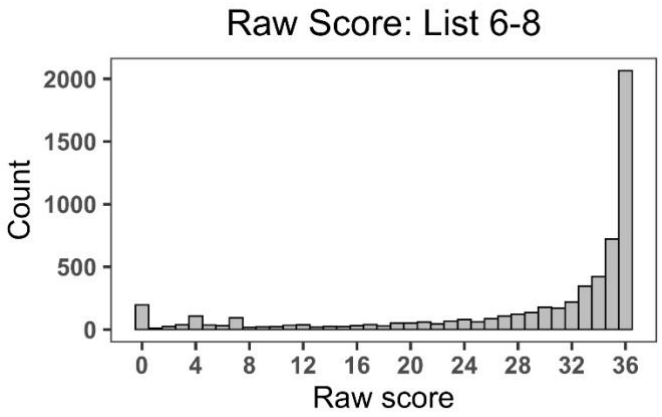


Table 6.3.1A
Raw Score Descriptive Statistics: List 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,086	0	36	28.58	10.50
7	1,904	0	36	29.32	10.17
8	1,850	0	36	29.89	9.93
Total	5,840	0	36	29.24	10.23

Figure 6.3.1B

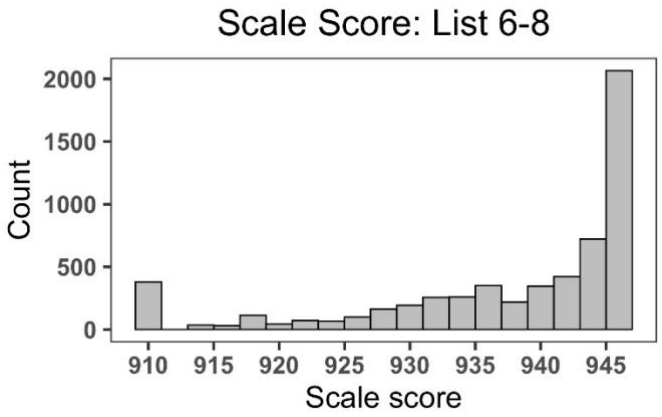


Table 6.3.1B
Scale Score Descriptive Statistics: List 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,086	910	946	937.14	10.68
7	1,904	910	946	937.94	10.35
8	1,850	910	946	938.58	10.16
Total	5,840	910	946	937.86	10.43

Figure 6.3.1C

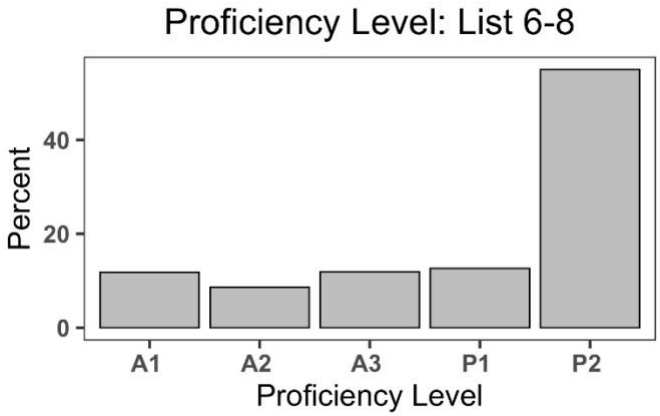


Table 6.3.1C

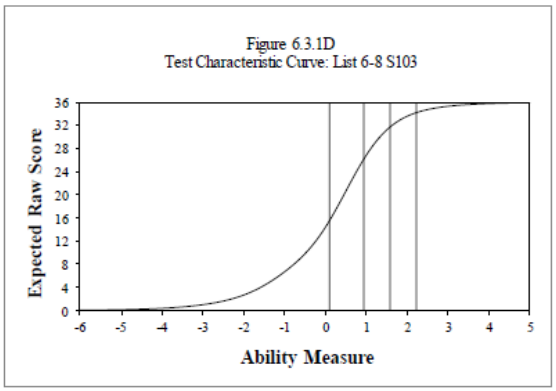
Proficiency Level Distribution: List 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	267	12.80	227	11.92	197	10.65	691	11.83
A2	224	10.74	148	7.77	132	7.14	504	8.63
A3	258	12.37	238	12.50	199	10.76	695	11.90
P1	258	12.37	247	12.97	233	12.59	738	12.64
P2	1,079	51.73	1,044	54.83	1,089	58.86	3,212	55.00
Total	2,086	100.00	1,904	100.00	1,850	100.00	5,840	100.00

Table 6.3.1D

Equating Summary: List 6-8

No equating summary is presented because the Alternate ACCESS Series 501 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.



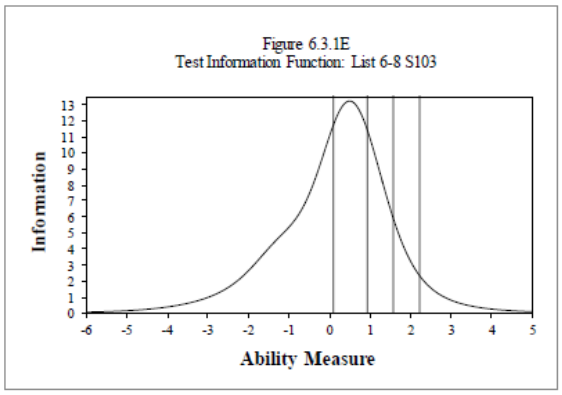


Table 6.3.1E

Reliability: List 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
5,840	9	0.951	2.298

Table 6.3.1F

Item Analysis Summary: List 6-8

Table 6.3.1G

Complete Item Analysis: List 6-8

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 6.3.1I

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	3.21	3.21	A1	3.36	3.36	A1	3.51	3.51
1	A1	0.24	3.45	A1	0.11	3.47	A1	0.22	3.73
2	A1	0.62	4.07	A1	0.21	3.68	A1	0.43	4.16
3	A1	0.96	5.03	A1	0.79	4.46	A1	0.22	4.38
4	A1	2.06	7.09	A1	1.73	6.20	A1	1.73	6.11
5	A1	0.86	7.96	A1	0.47	6.67	A1	0.43	6.54
6	A1	0.48	8.44	A1	0.47	7.14	A1	0.59	7.14
7	A1	1.58	10.02	A1	1.94	9.09	A1	1.30	8.43
8	A1	0.29	10.31	A1	0.42	9.51	A1	0.27	8.70
9	A1	0.19	10.50	A1	0.68	10.19	A1	0.22	8.92
10	A1	0.58	11.07	A1	0.21	10.40	A1	0.32	9.24
11	A1	0.58	11.65	A1	0.68	11.08	A1	0.49	9.73
12	A1	0.81	12.46	A1	0.32	11.40	A1	0.76	10.49
13	A1	0.34	12.80	A1	0.53	11.92	A1	0.16	10.65
14	A2	0.48	13.28	A2	0.47	12.39	A2	0.32	10.97
15	A2	0.43	13.71	A2	0.37	12.76	A2	0.38	11.35
16	A2	0.86	14.57	A2	0.26	13.03	A2	0.43	11.78
17	A2	0.91	15.48	A2	0.74	13.76	A2	0.38	12.16
18	A2	0.58	16.06	A2	0.21	13.97	A2	0.65	12.81
19	A2	1.25	17.31	A2	0.68	14.65	A2	0.65	13.46
20	A2	0.72	18.02	A2	1.00	15.65	A2	0.92	14.38
21	A2	1.15	19.18	A2	1.05	16.70	A2	0.92	15.30
22	A2	1.05	20.23	A2	0.68	17.38	A2	0.59	15.89
23	A2	1.58	21.81	A2	1.10	18.49	A2	0.76	16.65
24	A2	1.73	23.54	A2	1.21	19.70	A2	1.14	17.78
25	A3	1.05	24.59	A3	1.10	20.80	A3	1.03	18.81
26	A3	1.77	26.37	A3	1.26	22.06	A3	1.41	20.22
27	A3	1.73	28.09	A3	2.15	24.21	A3	1.68	21.89
28	A3	1.97	30.06	A3	2.73	26.94	A3	1.62	23.51
29	A3	2.59	32.65	A3	2.36	29.31	A3	2.00	25.51
30	A3	3.26	35.91	A3	2.89	32.20	A3	3.03	28.54
31	P1	2.64	38.54	P1	3.05	35.24	P1	3.14	31.68
32	P1	3.40	41.95	P1	3.99	39.23	P1	3.95	35.62
33	P1	6.33	48.27	P1	5.93	45.17	P1	5.51	41.14
34	P2	7.86	56.14	P2	6.62	51.79	P2	7.19	48.32
35	P2	11.60	67.74	P2	12.18	63.97	P2	13.46	61.78
36	P2	32.26	100.00	P2	36.03	100.00	P2	38.22	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

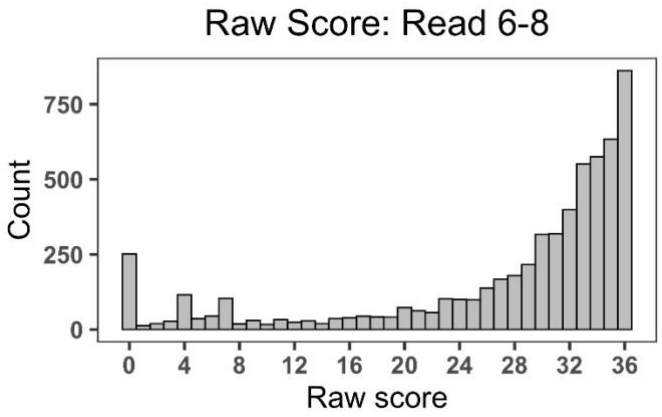


Table 6.3.2A
Raw Score Descriptive Statistics: Read 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,088	0	36	26.91	10.34
7	1,903	0	36	27.61	10.22
8	1,850	0	36	28.21	10.09
Total	5,841	0	36	27.55	10.24

Figure 6.3.2B

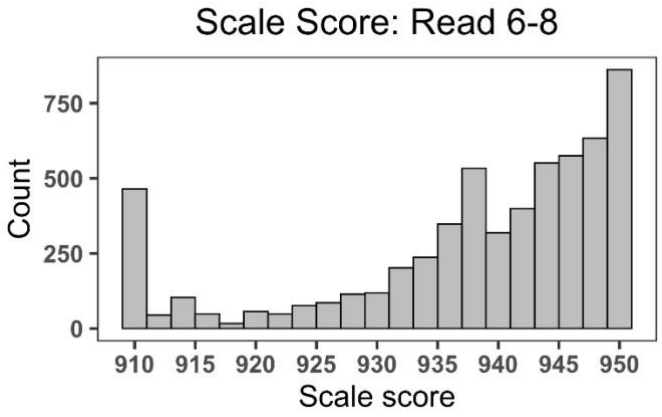


Table 6.3.2B
Scale Score Descriptive Statistics: Read 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,088	910	950	937.31	11.86
7	1,903	910	950	938.23	11.83
8	1,850	910	950	939.03	11.71
Total	5,841	910	950	938.15	11.83

Figure 6.3.2C

Table 6.3.2C

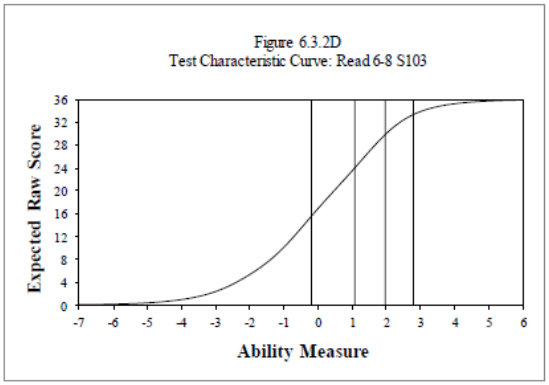
Proficiency Level Distribution: Read 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	296	14.18	260	13.66	230	12.43	786	13.46
A2	176	8.43	124	6.52	95	5.14	395	6.76
A3	236	11.30	193	10.14	178	9.62	607	10.39
P1	378	18.10	333	17.50	321	17.35	1,032	17.67
P2	1,002	47.99	993	52.18	1,026	55.46	3,021	51.72
Total	2,088	100.00	1,903	100.00	1,850	100.00	5,841	100.00

Table 6.3.2D

Equating Summary: Read 6-8

No equating summary is presented because the Alternate ACCESS Series 501 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).



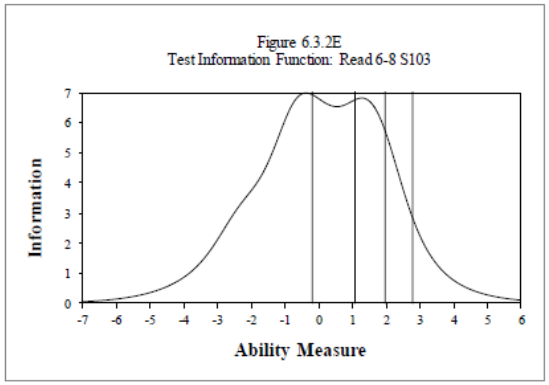


Table 6.3.2E

Reliability: Read 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
5,841	9	0.946	2.742

Table 6.3.2F

Item Analysis Summary: Read 6-8

Table 6.3.2G

Complete Item Analysis: Read 6-8

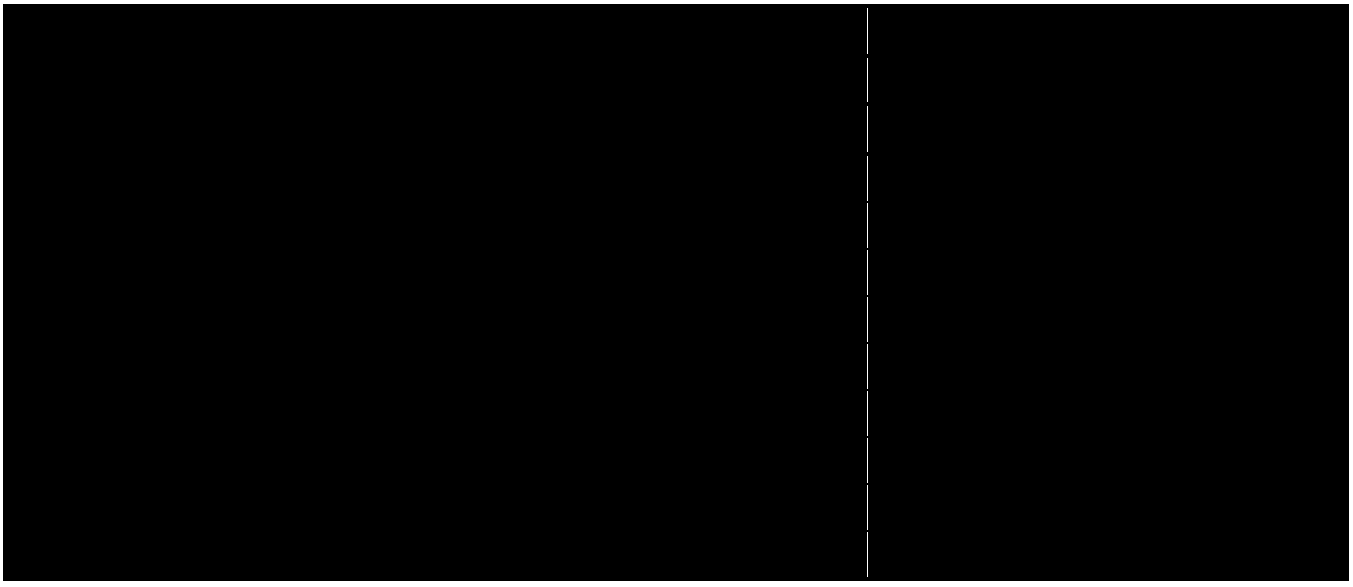


Table 6.3.2H

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	4.50	4.50	A1	4.05	4.05	A1	4.38	4.38
1	A1	0.14	4.65	A1	0.32	4.36	A1	0.22	4.59
2	A1	0.24	4.89	A1	0.47	4.83	A1	0.32	4.92
3	A1	0.67	5.56	A1	0.32	5.15	A1	0.43	5.35
4	A1	1.92	7.47	A1	2.15	7.30	A1	1.89	7.24
5	A1	0.91	8.38	A1	0.63	7.93	A1	0.27	7.51
6	A1	0.77	9.15	A1	0.68	8.62	A1	0.86	8.38
7	A1	1.82	10.97	A1	1.79	10.40	A1	1.73	10.11
8	A1	0.48	11.45	A1	0.26	10.67	A1	0.22	10.32
9	A1	0.48	11.93	A1	0.79	11.46	A1	0.27	10.59
10	A1	0.24	12.16	A1	0.16	11.61	A1	0.49	11.08
11	A1	0.77	12.93	A1	0.53	12.14	A1	0.38	11.46
12	A1	0.57	13.51	A1	0.47	12.61	A1	0.16	11.62
13	A1	0.48	13.98	A1	0.58	13.19	A1	0.43	12.05
14	A1	0.19	14.18	A1	0.47	13.66	A1	0.38	12.43
15	A2	1.01	15.18	A2	0.42	14.08	A2	0.43	12.86
16	A2	0.86	16.04	A2	0.63	14.71	A2	0.49	13.35
17	A2	1.01	17.05	A2	0.53	15.24	A2	0.70	14.05
18	A2	0.72	17.77	A2	0.84	16.08	A2	0.59	14.65
19	A2	0.96	18.73	A2	0.63	16.71	A2	0.49	15.14
20	A2	1.44	20.16	A2	1.42	18.13	A2	0.86	16.00
21	A2	1.15	21.31	A2	1.26	19.39	A2	0.81	16.81
22	A2	1.29	22.61	A2	0.79	20.18	A2	0.76	17.57
23	A3	1.92	24.52	A3	1.58	21.76	A3	1.73	19.30
24	A3	2.11	26.63	A3	1.58	23.33	A3	1.41	20.70
25	A3	1.58	28.21	A3	1.58	24.91	A3	1.95	22.65
26	A3	2.44	30.65	A3	2.52	27.43	A3	2.11	24.76
27	A3	3.26	33.91	A3	2.89	30.32	A3	2.43	27.19
28	P1	3.26	37.16	P1	3.10	33.42	P1	2.86	30.05
29	P1	3.50	40.66	P1	3.73	37.15	P1	3.89	33.95
30	P1	5.41	46.07	P1	5.10	42.25	P1	5.78	39.73
31	P1	5.94	52.01	P1	5.57	47.82	P1	4.81	44.54
32	P2	6.94	58.96	P2	6.73	54.55	P2	6.81	51.35
33	P2	8.96	67.91	P2	10.09	64.63	P2	9.30	60.65
34	P2	10.58	78.50	P2	9.41	74.04	P2	9.46	70.11
35	P2	10.73	89.22	P2	10.40	84.45	P2	11.46	81.57
36	P2	10.78	100.00	P2	15.55	100.00	P2	18.43	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

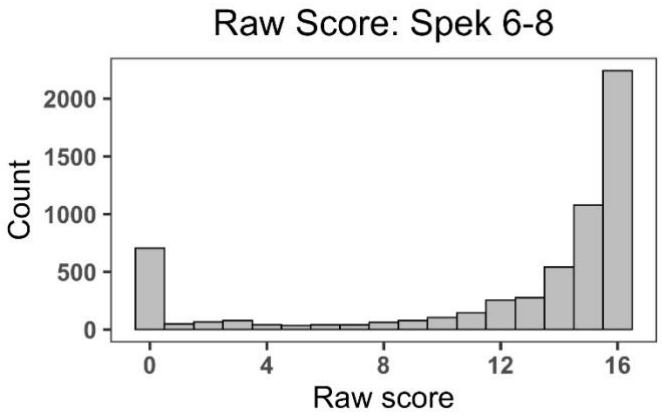


Table 6.3.3A
Raw Score Descriptive Statistics: Spek 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,089	0	16	12.05	5.55
7	1,902	0	16	12.28	5.49
8	1,849	0	16	12.38	5.47
Total	5,840	0	16	12.23	5.51

Figure 6.3.3B

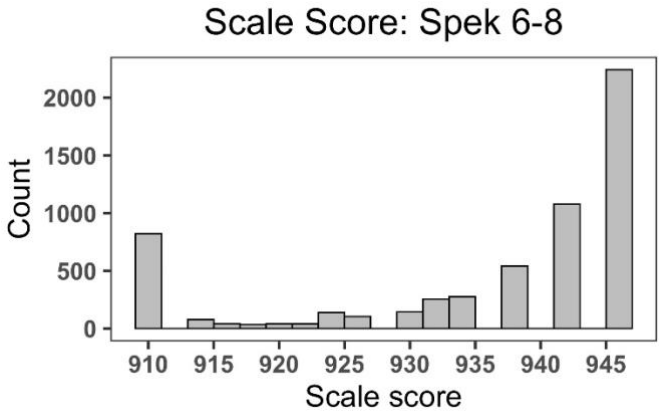


Table 6.3.3B
Scale Score Descriptive Statistics: Spek 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,089	910	947	936.08	13.16
7	1,902	910	947	936.74	13.06
8	1,849	910	947	937.03	12.98
Total	5,840	910	947	936.60	13.08

Figure 6.3.3C

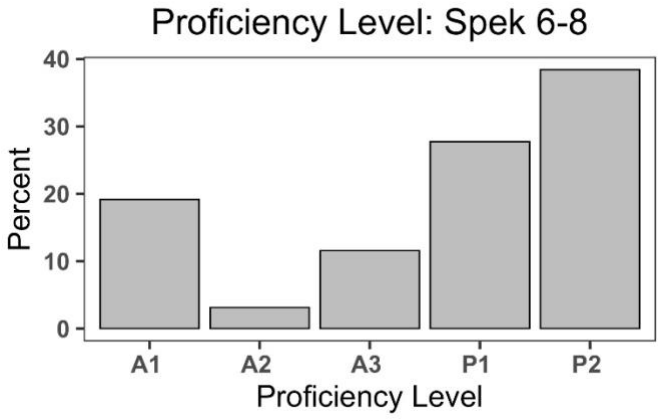


Table 6.3.3C

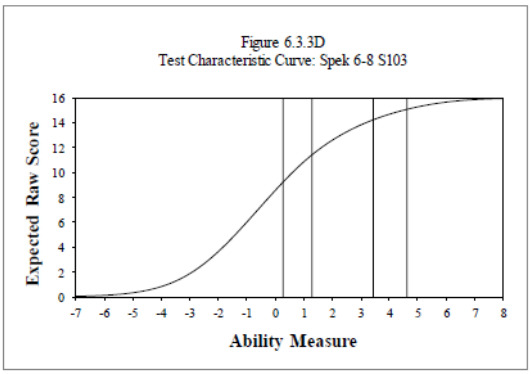
Proficiency Level Distribution: Spek 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	422	20.20	365	19.19	332	17.96	1,119	19.16
A2	70	3.35	57	3.00	55	2.97	182	3.12
A3	265	12.69	213	11.20	197	10.65	675	11.56
P1	576	27.57	517	27.18	527	28.50	1,620	27.74
P2	756	36.19	750	39.43	738	39.91	2,244	38.42
Total	2,089	100.00	1,902	100.00	1,849	100.00	5,840	100.00

Table 6.3.3D

Equating Summary: Spek 6-8

No equating summary is presented because the Alternate ACCESS Series 501 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).



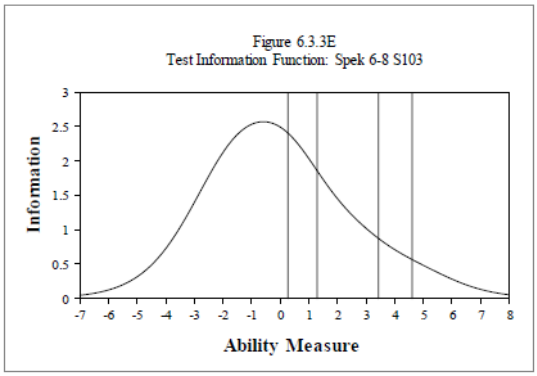


Table 6.3.3E

Reliability: Spek 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
5,840	8	0.965	2.436

Table 6.3.3F

Item Analysis Summary: Spek 6-8

Table 6.3.3G

Complete Item Analysis: Spek 6-8

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	12.35	12.35	A1	11.83	11.83	A1	12.01	12.01
1	A1	0.91	13.26	A1	0.84	12.67	A1	0.76	12.76
2	A1	1.10	14.36	A1	1.10	13.77	A1	1.24	14.01
3	A1	1.48	15.84	A1	1.31	15.09	A1	1.14	15.14
4	A1	0.86	16.71	A1	0.74	15.83	A1	0.54	15.68
5	A1	0.72	17.42	A1	0.68	16.51	A1	0.43	16.12
6	A1	0.72	18.14	A1	0.79	17.30	A1	0.59	16.71
7	A1	0.91	19.05	A1	0.58	17.88	A1	0.54	17.25
8	A1	1.15	20.20	A1	1.31	19.19	A1	0.70	17.96
9	A2	1.24	21.45	A2	1.31	20.50	A2	1.41	19.36
10	A2	2.11	23.55	A2	1.68	22.19	A2	1.57	20.93
11	A3	2.54	26.09	A3	2.47	24.66	A3	2.38	23.31
12	A3	5.03	31.12	A3	4.15	28.81	A3	3.89	27.20
13	A3	5.12	36.24	A3	4.57	33.39	A3	4.38	31.58
14	P1	9.29	45.52	P1	8.68	42.06	P1	9.90	41.48
15	P1	18.29	63.81	P1	18.51	60.57	P1	18.60	60.09
16	P2	36.19	100.00	P2	39.43	100.00	P2	39.91	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

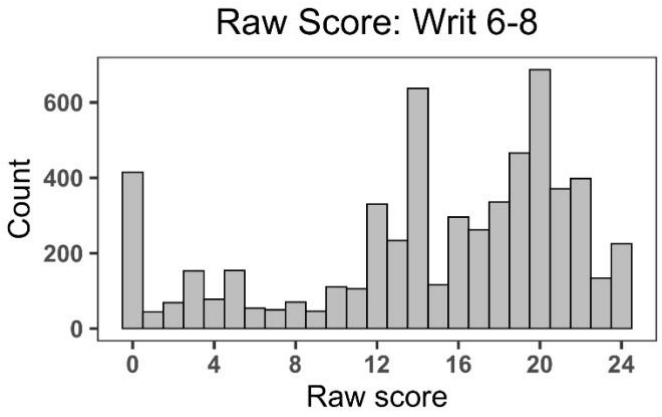


Table 6.3.4A
Raw Score Descriptive Statistics: Writ 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,089	0	24	14.50	6.71
7	1,902	0	24	14.68	6.79
8	1,851	0	24	15.10	6.91
Total	5,842	0	24	14.75	6.80

Figure 6.3.4B

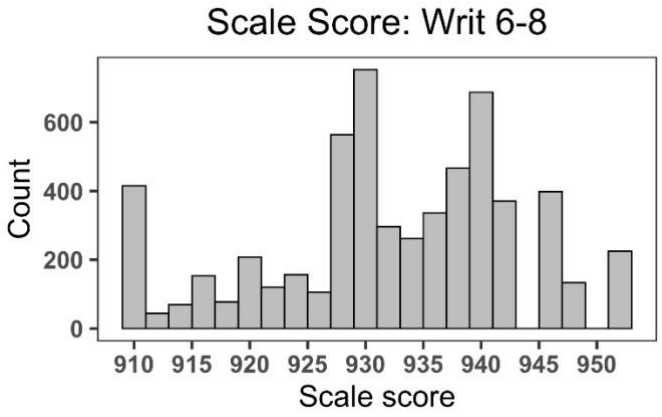


Table 6.3.4B
Scale Score Descriptive Statistics: Writ 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,089	910	953	932.30	10.59
7	1,902	910	953	932.65	10.80
8	1,851	910	953	933.41	11.13
Total	5,842	910	953	932.76	10.84

Figure 6.3.4C

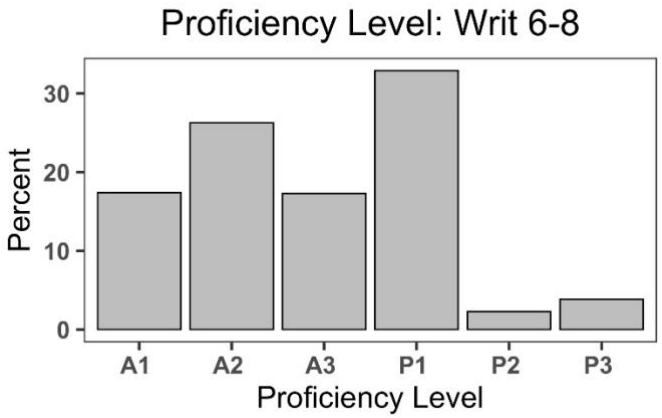


Table 6.3.4C

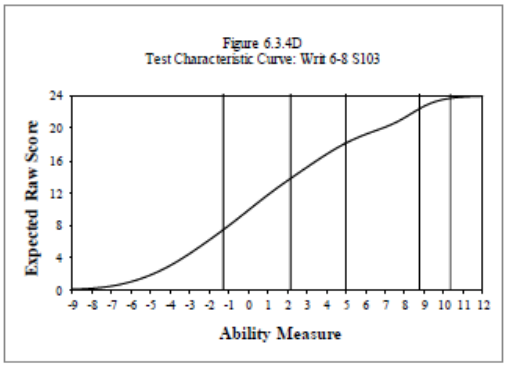
Proficiency Level Distribution: Writ 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	364	17.42	340	17.88	313	16.91	1,017	17.41
A2	597	28.58	510	26.81	427	23.07	1,534	26.26
A3	369	17.66	315	16.56	326	17.61	1,010	17.29
P1	653	31.26	618	32.49	651	35.17	1,922	32.90
P2	42	2.01	50	2.63	42	2.27	134	2.29
P3	64	3.06	69	3.63	92	4.97	225	3.85
Total	2,089	100.00	1,902	100.00	1,851	100.00	5,842	100.00

Table 6.3.4D

Equating Summary: Writ 6-8

No equating summary is presented because the Alternate ACCESS Series 501 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).



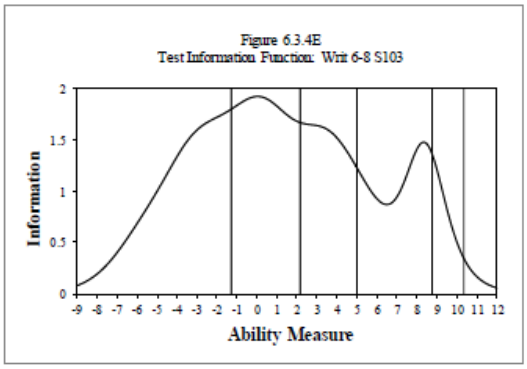


Table 6.3.4E

Reliability: Writ 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
5,842	10	0.922	3.019

Table 6.3.4F

Item Analysis Summary: Writ 6-8

Table 6.3.4G

Complete Item Analysis: Writ 6-8

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	6.94	6.94	A1	7.10	7.10	A1	7.29	7.29
1	A1	0.67	7.61	A1	0.63	7.73	A1	0.97	8.27
2	A1	1.05	8.66	A1	1.31	9.04	A1	1.19	9.45
3	A1	3.21	11.87	A1	2.31	11.36	A1	2.27	11.72
4	A1	1.34	13.21	A1	1.58	12.93	A1	1.08	12.80
5	A1	2.63	15.84	A1	2.68	15.62	A1	2.59	15.40
6	A1	0.67	16.52	A1	1.31	16.93	A1	0.81	16.21
7	A1	0.91	17.42	A1	0.95	17.88	A1	0.70	16.91
8	A2	1.34	18.76	A2	1.00	18.87	A2	1.24	18.15
9	A2	1.01	19.77	A2	0.74	19.61	A2	0.59	18.75
10	A2	2.11	21.88	A2	1.68	21.29	A2	1.89	20.64
11	A2	2.25	24.13	A2	1.63	22.92	A2	1.51	22.15
12	A2	5.98	30.11	A2	6.36	29.28	A2	4.54	26.69
13	A2	4.21	34.32	A2	4.15	33.44	A2	3.62	30.31
14	A2	11.68	46.00	A2	11.25	44.69	A2	9.67	39.98
15	A3	1.91	47.92	A3	1.79	46.48	A3	2.27	42.25
16	A3	4.93	52.85	A3	5.15	51.63	A3	5.13	47.38
17	A3	5.07	57.92	A3	4.05	55.68	A3	4.27	51.65
18	A3	5.74	63.67	A3	5.57	61.25	A3	5.94	57.59
19	P1	7.80	71.47	P1	8.36	69.61	P1	7.78	65.37
20	P1	11.20	82.67	P1	11.78	81.39	P1	12.37	77.74
21	P1	5.65	88.32	P1	5.89	87.28	P1	7.62	85.36
22	P1	6.61	94.93	P1	6.47	93.74	P1	7.40	92.76
23	P2	2.01	96.94	P2	2.63	96.37	P2	2.27	95.03
24	P3	3.06	100.00	P3	3.63	100.00	P3	4.97	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

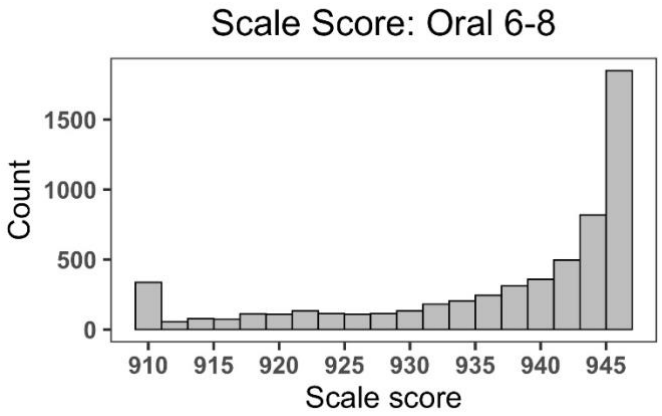


Table 6.3.5A

Scale Score Descriptive Statistics: Oral 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,085	910	947	936.96	11.45
7	1,902	910	947	937.72	11.22
8	1,848	910	947	938.18	11.01
Total	5,835	910	947	937.59	11.25

Figure 6.3.5B

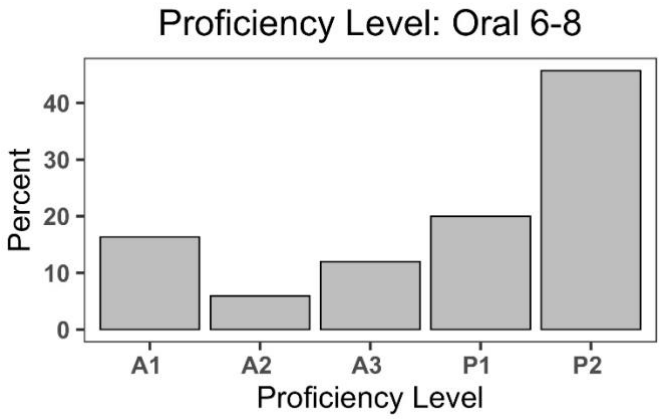


Table 6.3.5C

Proficiency Level Distribution: Oral 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	370	17.75	304	15.98	280	15.15	954	16.35
A2	128	6.14	120	6.31	99	5.36	347	5.95
A3	270	12.95	221	11.62	209	11.31	700	12.00
P1	417	20.00	377	19.82	373	20.18	1,167	20.00
P2	900	43.17	880	46.27	887	48.00	2,667	45.71
Total	2,085	100.00	1,902	100.00	1,848	100.00	5,835	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	108.735	0.951
Speaking	0.5	171.148	0.965
Oral		126.477	0.978

*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

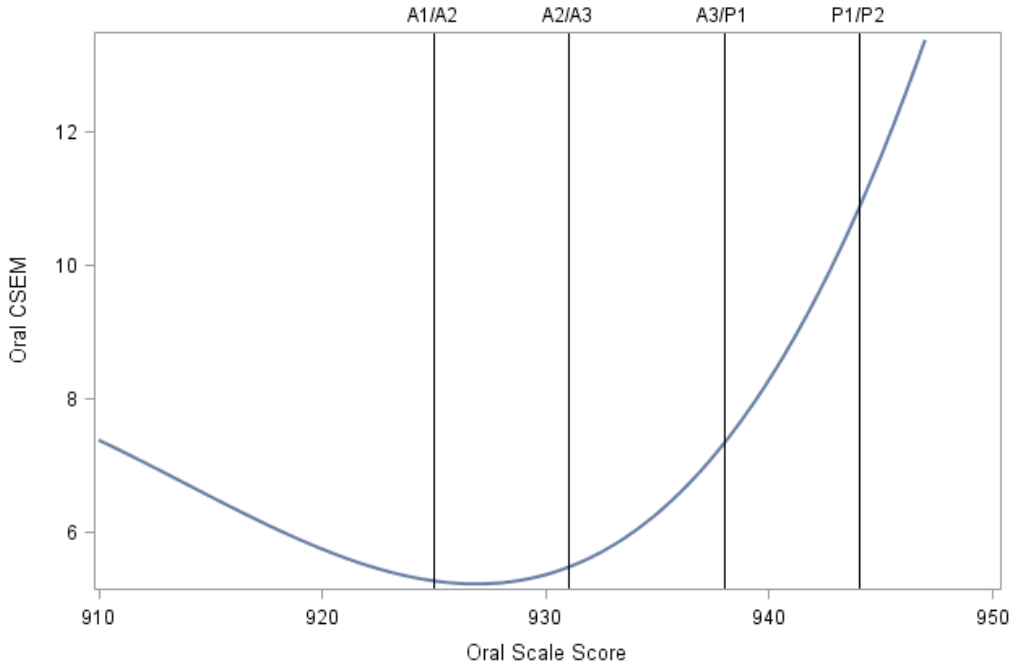
Table 6.3.5J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

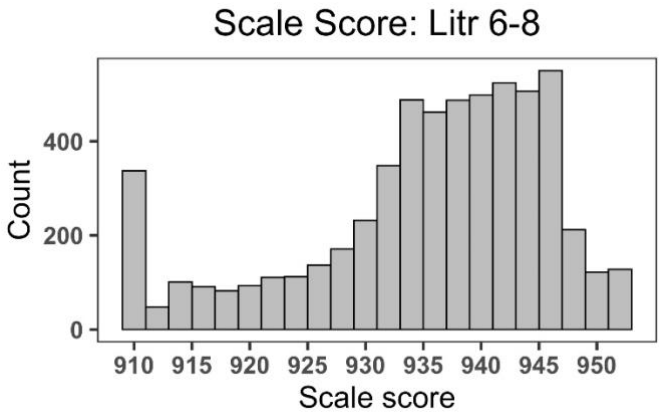


Table 6.3.6A

Scale Score Descriptive Statistics: Litr 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,088	910	952	935.04	10.59
7	1,902	910	952	935.68	10.73
8	1,850	910	952	936.47	10.83
Total	5,840	910	952	935.70	10.73

Figure 6.3.6B

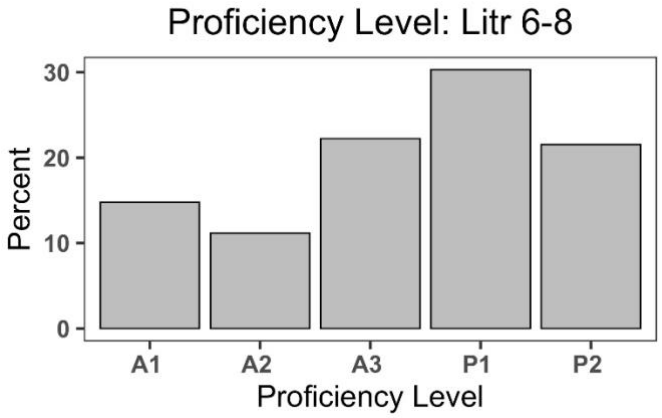


Table 6.3.6C

Proficiency Level Distribution: Litr 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	327	15.66	285	14.98	251	13.57	863	14.78
A2	257	12.31	203	10.67	192	10.38	652	11.16
A3	477	22.84	432	22.71	389	21.03	1,298	22.23
P1	646	30.94	584	30.70	540	29.19	1,770	30.31
P2	381	18.25	398	20.93	478	25.84	1,257	21.52
Total	2,088	100.00	1,902	100.00	1,850	100.00	5,840	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	139.787	0.946
Writing	0.5	117.436	0.922
Literacy		115.105	0.964

*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

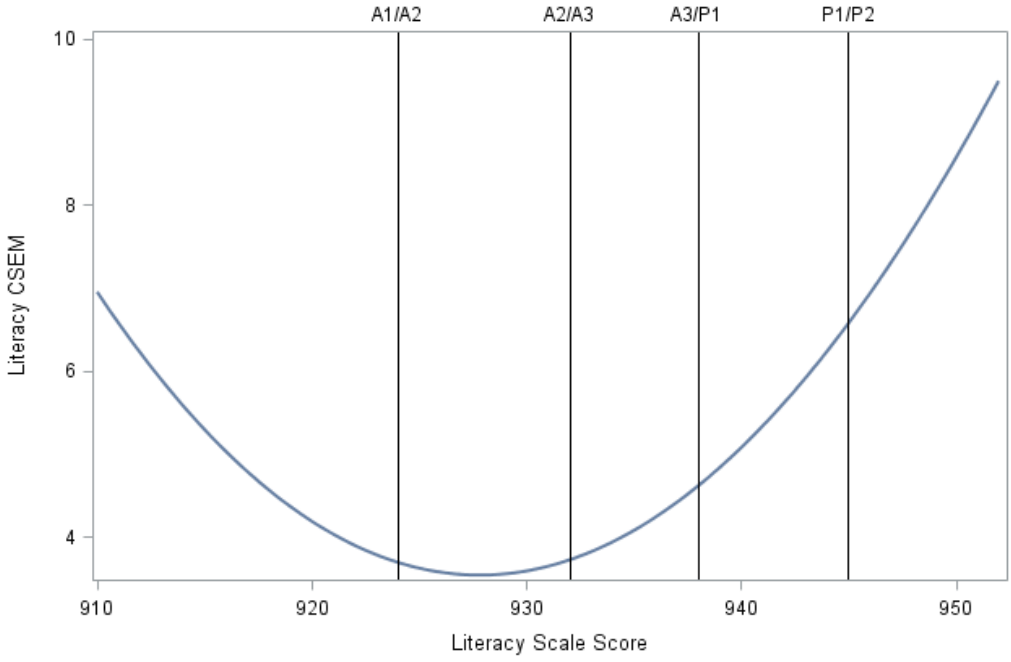
Table 6.3.6I
n/a

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

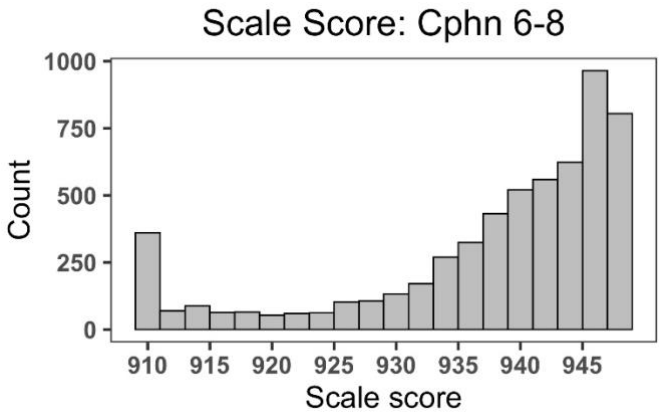


Table 6.3.7A

Scale Score Descriptive Statistics: Cphn 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,086	910	949	937.27	11.19
7	1,903	910	949	938.19	11.11
8	1,849	910	949	938.93	10.97
Total	5,838	910	949	938.09	11.12

Figure 6.3.7B

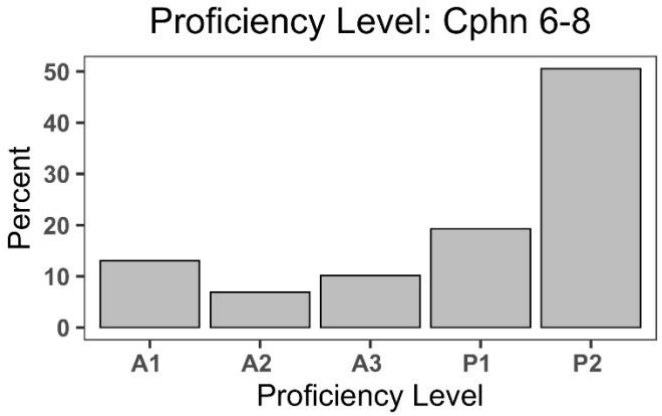


Table 6.3.7C

Proficiency Level Distribution: Cphn 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	294	14.09	251	13.19	217	11.74	762	13.05
A2	172	8.25	123	6.46	110	5.95	405	6.94
A3	227	10.88	197	10.35	169	9.14	593	10.16
P1	419	20.09	366	19.23	341	18.44	1,126	19.29
P2	974	46.69	966	50.76	1,012	54.73	2,952	50.57
Total	2,086	100.00	1,903	100.00	1,849	100.00	5,838	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	108.735	0.951
Reading	0.7	139.787	0.946
Comprehension		123.641	0.966

*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

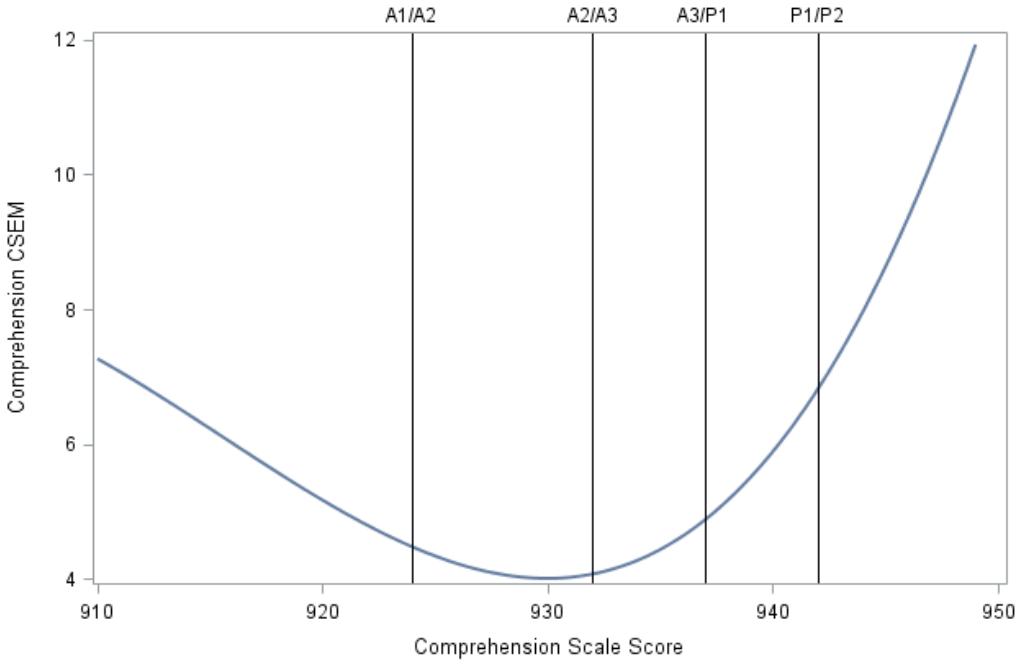
Table 6.3.7J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

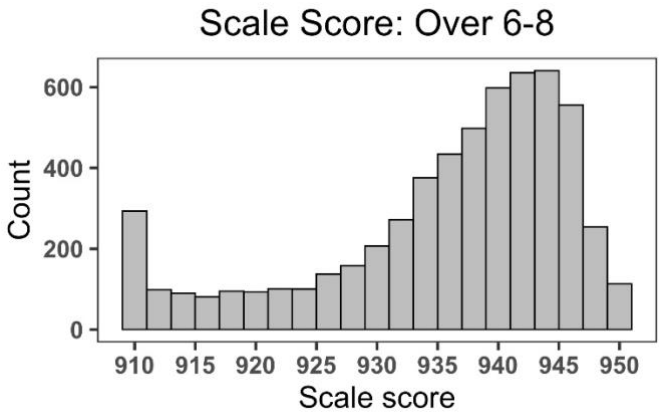


Table 6.3.8A

Scale Score Descriptive Statistics: Over 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,085	910	950	935.38	10.49
7	1,900	910	950	936.06	10.56
8	1,847	910	950	936.73	10.55
Total	5,832	910	950	936.03	10.55

Figure 6.3.8B

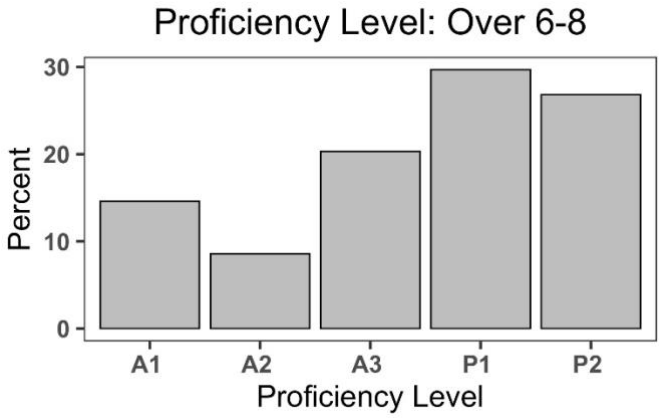


Table 6.3.8C

Proficiency Level Distribution: Over 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	319	15.30	280	14.74	253	13.70	852	14.61
A2	210	10.07	163	8.58	127	6.88	500	8.57
A3	442	21.20	374	19.68	368	19.92	1,184	20.30
P1	633	30.36	567	29.84	532	28.80	1,732	29.70
P2	481	23.07	516	27.16	567	30.70	1,564	26.82
Total	2,085	100.00	1,900	100.00	1,847	100.00	5,832	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	108.735	0.951
Reading	0.35	139.787	0.946
Speaking	0.15	171.148	0.965
Writing	0.35	117.436	0.922
Overall Composite		111.259	0.979

*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

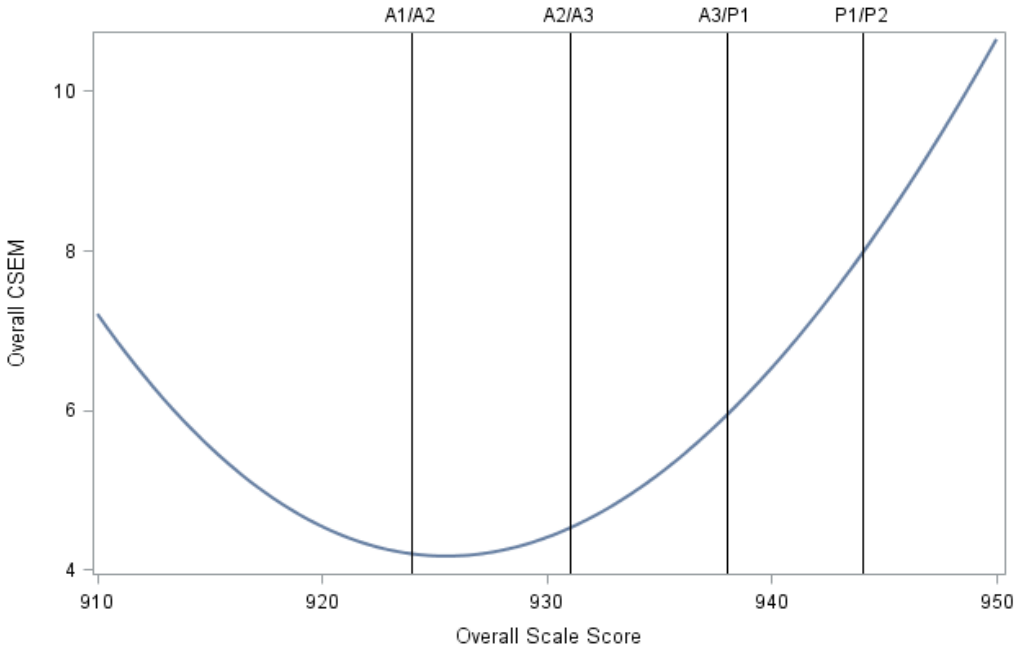
Table 6.3.8J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

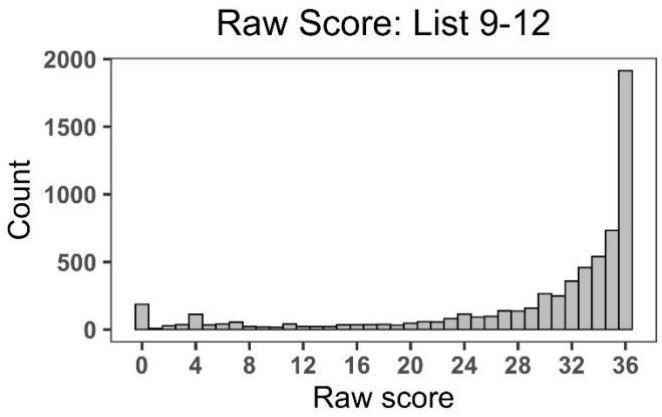


Table 6.4.1A
Raw Score Descriptive Statistics: List 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,686	0	36	29.01	9.73
10	1,418	0	36	29.80	9.26
11	1,379	0	36	29.80	9.35
12	1,862	0	36	29.14	10.04
Total	6,345	0	36	29.39	9.65

Figure 6.4.1B

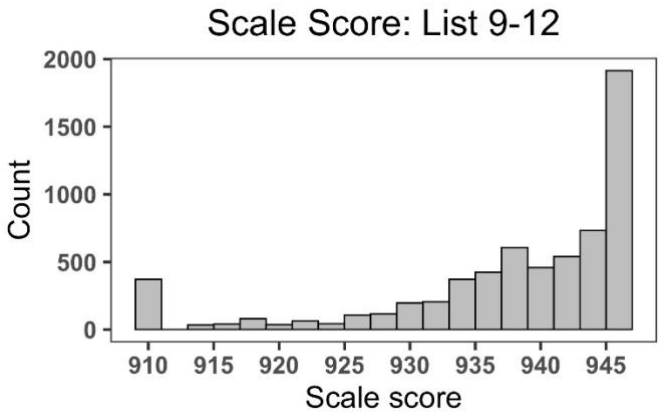


Table 6.4.1B
Scale Score Descriptive Statistics: List 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,686	910	947	938.19	10.16
10	1,418	910	947	939.08	9.80
11	1,379	910	947	939.08	9.78
12	1,862	910	947	938.40	10.49
Total	6,345	910	947	938.64	10.11

Figure 6.4.1B

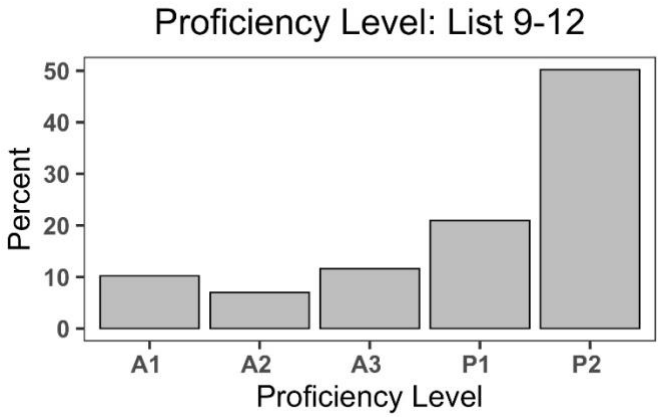


Table 6.4.1C

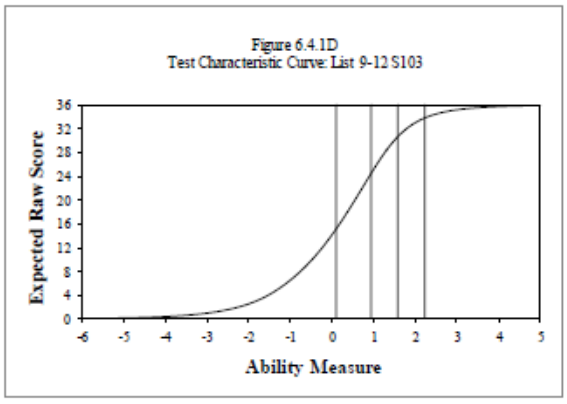
Proficiency Level Distribution: List 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	179	10.62	134	9.45	126	9.14	208	11.17	647	10.20
A2	135	8.01	91	6.42	93	6.74	124	6.66	443	6.98
A3	214	12.69	162	11.42	144	10.44	217	11.65	737	11.62
P1	354	21.00	282	19.89	306	22.19	389	20.89	1,331	20.98
P2	804	47.69	749	52.82	710	51.49	924	49.62	3,187	50.23
Total	1,686	100.00	1,418	100.00	1,379	100.00	1,862	100.00	6,345	100.00

Table 6.4.1D

Equating Summary: List 9-12

No equating summary is presented because the Alternate ACCESS Series 501 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.



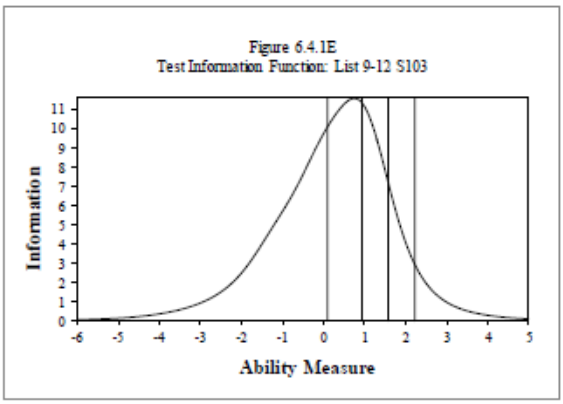


Table 6.4.1E

Reliability: List 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
6,345	9	0.943	2.412

Table 6.4.1F

Item Analysis Summary: List 9-12

Table 6.4.1G

Complete Item Analysis: List 9-12

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 6.4.1I

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	2.97	2.97	A1	1.83	1.83	A1	2.83	2.83	A1	3.87	3.87
1	A1	0.18	3.14	A1	0.07	1.90	A1	0.00	2.83	A1	0.21	4.08
2	A1	0.30	3.44	A1	0.35	2.26	A1	0.51	3.34	A1	0.54	4.62
3	A1	0.71	4.15	A1	0.63	2.89	A1	0.65	3.99	A1	0.38	4.99
4	A1	1.90	6.05	A1	1.97	4.87	A1	1.23	5.22	A1	1.93	6.93
5	A1	0.36	6.41	A1	0.71	5.57	A1	0.51	5.73	A1	0.54	7.47
6	A1	0.77	7.18	A1	0.71	6.28	A1	0.51	6.24	A1	0.54	8.00
7	A1	0.83	8.01	A1	1.06	7.33	A1	0.73	6.96	A1	0.91	8.92
8	A1	0.42	8.42	A1	0.35	7.69	A1	0.29	7.25	A1	0.38	9.29
9	A1	0.12	8.54	A1	0.21	7.90	A1	0.51	7.76	A1	0.38	9.67
10	A1	0.36	8.90	A1	0.21	8.11	A1	0.36	8.12	A1	0.21	9.88
11	A1	0.53	9.43	A1	0.71	8.82	A1	0.58	8.70	A1	0.75	10.63
12	A1	0.59	10.02	A1	0.35	9.17	A1	0.36	9.06	A1	0.16	10.79
13	A1	0.59	10.62	A1	0.28	9.45	A1	0.07	9.14	A1	0.38	11.17
14	A2	0.65	11.27	A2	0.28	9.73	A2	0.29	9.43	A2	0.16	11.33
15	A2	0.53	11.80	A2	0.42	10.16	A2	0.80	10.22	A2	0.48	11.82
16	A2	0.65	12.46	A2	0.56	10.72	A2	0.51	10.73	A2	0.54	12.35
17	A2	0.47	12.93	A2	0.71	11.42	A2	0.51	11.24	A2	0.64	13.00
18	A2	0.59	13.52	A2	0.85	12.27	A2	0.65	11.89	A2	0.38	13.37
19	A2	0.65	14.18	A2	0.14	12.41	A2	0.87	12.76	A2	0.38	13.75
20	A2	0.89	15.07	A2	0.63	13.05	A2	0.73	13.49	A2	0.64	14.39
21	A2	1.25	16.31	A2	0.92	13.96	A2	0.73	14.21	A2	0.75	15.15
22	A2	1.01	17.32	A2	0.78	14.74	A2	0.65	14.87	A2	1.07	16.22
23	A2	1.30	18.62	A2	1.13	15.87	A2	1.02	15.88	A2	1.61	17.83
24	A3	2.25	20.88	A3	1.69	17.56	A3	1.52	17.40	A3	1.72	19.55
25	A3	1.25	22.12	A3	1.55	19.11	A3	1.31	18.71	A3	1.61	21.16
26	A3	1.54	23.67	A3	1.41	20.52	A3	1.52	20.23	A3	1.61	22.77
27	A3	2.31	25.98	A3	1.76	22.28	A3	2.10	22.34	A3	2.47	25.24
28	A3	2.31	28.29	A3	2.05	24.33	A3	2.18	24.51	A3	2.04	27.28
29	A3	3.02	31.32	A3	2.96	27.29	A3	1.81	26.32	A3	2.20	29.48
30	P1	4.63	35.94	P1	4.51	31.81	P1	3.84	30.17	P1	3.76	33.24
31	P1	3.80	39.74	P1	4.23	36.04	P1	3.41	33.58	P1	4.14	37.38
32	P1	5.34	45.08	P1	5.50	41.54	P1	6.02	39.59	P1	5.80	43.18
33	P1	7.24	52.31	P1	5.64	47.18	P1	8.92	48.51	P1	7.20	50.38
34	P2	8.84	61.15	P2	8.25	55.43	P2	9.43	57.94	P2	7.73	58.11
35	P2	11.09	72.24	P2	13.40	68.83	P2	11.31	69.25	P2	10.69	68.80
36	P2	27.76	100.00	P2	31.17	100.00	P2	30.75	100.00	P2	31.20	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

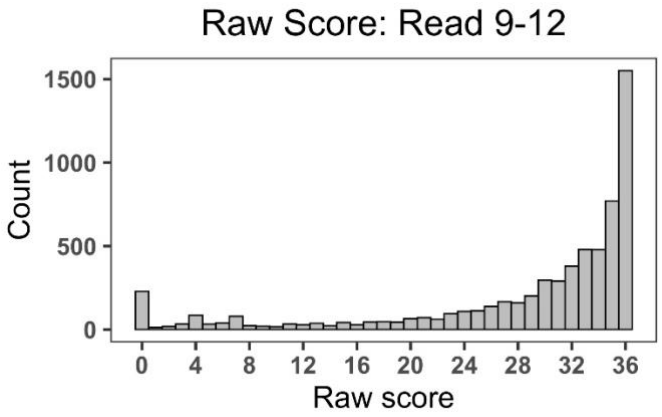


Table 6.4.2A
Raw Score Descriptive Statistics: Read 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,687	0	36	28.45	9.88
10	1,418	0	36	29.03	9.46
11	1,380	0	36	29.16	9.46
12	1,861	0	36	28.35	10.15
Total	6,346	0	36	28.71	9.79

Figure 6.4.2B

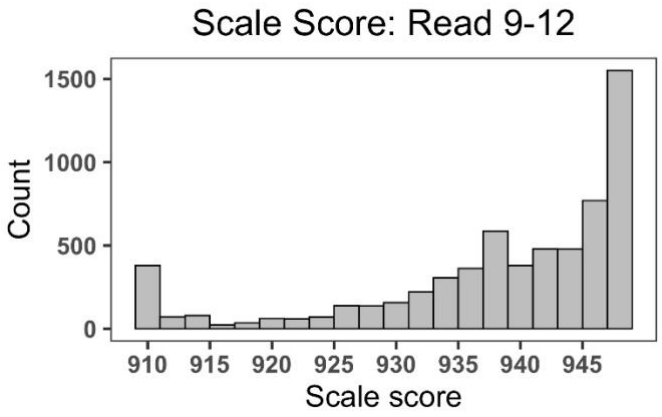


Table 6.4.2B
Scale Score Descriptive Statistics: Read 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,687	910	948	938.02	10.82
10	1,418	910	948	938.71	10.49
11	1,380	910	948	938.90	10.46
12	1,861	910	948	937.99	11.13
Total	6,346	910	948	938.36	10.77

Figure 6.4.2C

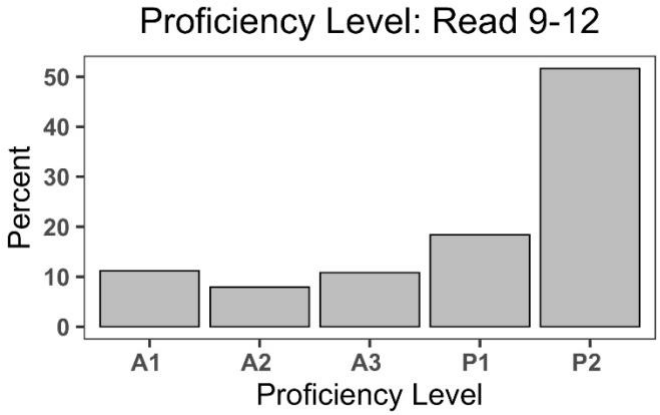


Table 6.4.2C

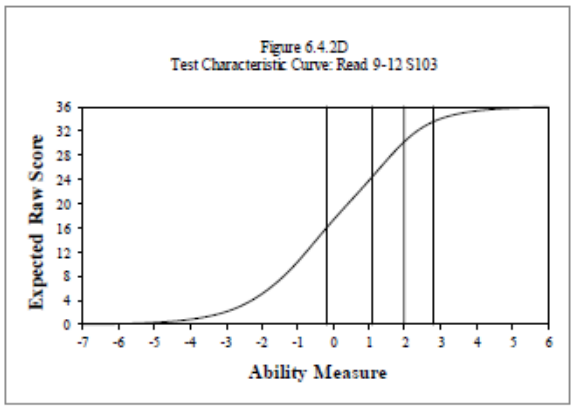
Proficiency Level Distribution: Read 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	191	11.32	152	10.72	143	10.36	224	12.04	710	11.19
A2	138	8.18	111	7.83	106	7.68	148	7.95	503	7.93
A3	192	11.38	151	10.65	129	9.35	215	11.55	687	10.83
P1	327	19.38	247	17.42	257	18.62	336	18.05	1,167	18.39
P2	839	49.73	757	53.39	745	53.99	938	50.40	3,279	51.67
Total	1,687	100.00	1,418	100.00	1,380	100.00	1,861	100.00	6,346	100.00

Table 6.4.2D

Equating Summary: Read 9-12

No equating summary is presented because the Alternate ACCESS Series 501 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).



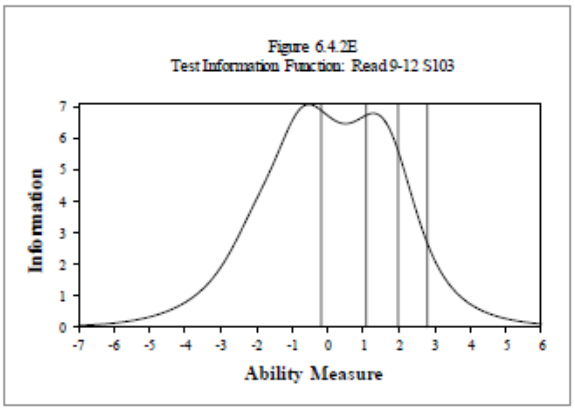


Table 6.4.2E

Reliability: Read 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
6,346	9	0.944	2.554

Table 6.4.2F

Item Analysis Summary: Read 9-12

Table 6.4.2G

Complete Item Analysis: Read 9-12

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	3.73	3.73	A1	2.89	2.89	A1	3.26	3.26	A1	4.30	4.30
1	A1	0.36	4.09	A1	0.14	3.03	A1	0.07	3.33	A1	0.21	4.51
2	A1	0.18	4.27	A1	0.14	3.17	A1	0.43	3.77	A1	0.43	4.94
3	A1	0.71	4.98	A1	0.21	3.39	A1	0.65	4.42	A1	0.48	5.43
4	A1	1.24	6.22	A1	1.48	4.87	A1	1.01	5.43	A1	1.56	6.99
5	A1	0.24	6.46	A1	0.42	5.29	A1	0.65	6.09	A1	0.70	7.68
6	A1	0.77	7.23	A1	0.99	6.28	A1	0.22	6.30	A1	0.54	8.22
7	A1	1.72	8.95	A1	1.13	7.40	A1	0.94	7.25	A1	1.13	9.35
8	A1	0.30	9.25	A1	0.49	7.90	A1	0.14	7.39	A1	0.48	9.83
9	A1	0.30	9.54	A1	0.42	8.32	A1	0.29	7.68	A1	0.27	10.10
10	A1	0.36	9.90	A1	0.28	8.60	A1	0.14	7.83	A1	0.21	10.32
11	A1	0.59	10.49	A1	0.56	9.17	A1	0.72	8.55	A1	0.27	10.59
12	A1	0.36	10.85	A1	0.42	9.59	A1	0.36	8.91	A1	0.64	11.23
13	A1	0.36	11.20	A1	0.63	10.23	A1	0.80	9.71	A1	0.59	11.82
14	A1	0.12	11.32	A1	0.49	10.72	A1	0.65	10.36	A1	0.21	12.04
15	A2	0.83	12.15	A2	0.42	11.14	A2	0.72	11.09	A2	0.64	12.68
16	A2	0.47	12.63	A2	0.56	11.71	A2	0.43	11.52	A2	0.38	13.06
17	A2	1.07	13.69	A2	0.56	12.27	A2	0.43	11.96	A2	0.75	13.81
18	A2	0.71	14.40	A2	0.99	13.26	A2	0.58	12.54	A2	0.70	14.51
19	A2	0.71	15.12	A2	0.28	13.54	A2	0.80	13.33	A2	0.97	15.48
20	A2	0.83	15.95	A2	1.34	14.88	A2	0.87	14.20	A2	1.07	16.55
21	A2	1.24	17.19	A2	1.13	16.01	A2	1.09	15.29	A2	1.07	17.62
22	A2	0.65	17.84	A2	0.99	17.00	A2	1.09	16.38	A2	1.18	18.81
23	A2	1.66	19.50	A2	1.55	18.55	A2	1.67	18.04	A2	1.18	19.99
24	A3	1.60	21.10	A3	1.20	19.75	A3	2.03	20.07	A3	1.99	21.98
25	A3	1.84	22.94	A3	1.83	21.58	A3	1.45	21.52	A3	1.88	23.86
26	A3	2.61	25.55	A3	2.12	23.70	A3	1.67	23.19	A3	2.20	26.06
27	A3	2.73	28.28	A3	2.82	26.52	A3	1.74	24.93	A3	3.06	29.12
28	A3	2.61	30.88	A3	2.68	29.20	A3	2.46	27.39	A3	2.42	31.54
29	P1	3.20	34.08	P1	2.68	31.88	P1	3.48	30.87	P1	3.28	34.82
30	P1	4.86	38.94	P1	4.37	36.25	P1	4.78	35.65	P1	4.62	39.44
31	P1	5.04	43.98	P1	4.23	40.48	P1	4.93	40.58	P1	4.14	43.58
32	P1	6.28	50.27	P1	6.14	46.61	P1	5.43	46.01	P1	6.02	49.60
33	P2	8.42	58.68	P2	7.62	54.23	P2	7.39	53.41	P2	6.88	56.48
34	P2	7.11	65.80	P2	8.32	62.55	P2	7.46	60.87	P2	7.36	63.84
35	P2	11.44	77.24	P2	12.76	75.32	P2	13.70	74.57	P2	11.12	74.96
36	P2	22.76	100.00	P2	24.68	100.00	P2	25.43	100.00	P2	25.04	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

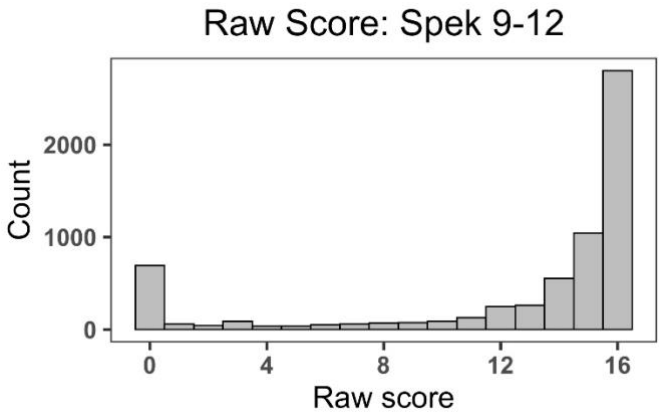


Table 6.4.3A
Raw Score Descriptive Statistics: Spek 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,687	0	16	12.45	5.41
10	1,419	0	16	12.67	5.32
11	1,381	0	16	12.76	5.27
12	1,859	0	16	12.41	5.44
Total	6,346	0	16	12.55	5.37

Figure 6.4.3B

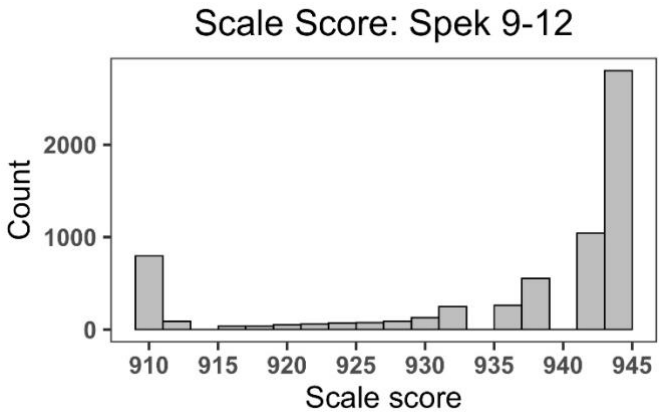


Table 6.4.3B
Scale Score Descriptive Statistics: Spek 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,687	910	945	936.28	12.23
10	1,419	910	945	936.82	12.04
11	1,381	910	945	937.04	11.93
12	1,859	910	945	936.21	12.26
Total	6,346	910	945	936.55	12.14

Figure 6.4.3C

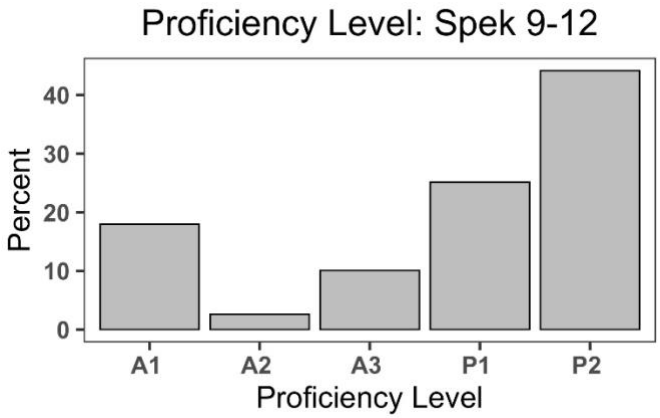


Table 6.4.3C

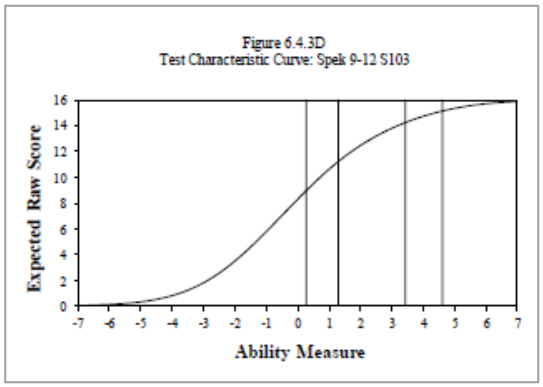
Proficiency Level Distribution: Spek 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	306	18.14	251	17.69	233	16.87	351	18.88	1,141	17.98
A2	48	2.85	31	2.18	33	2.39	54	2.90	166	2.62
A3	181	10.73	143	10.08	126	9.12	191	10.27	641	10.10
P1	440	26.08	342	24.10	352	25.49	462	24.85	1,596	25.15
P2	712	42.21	652	45.95	637	46.13	801	43.09	2,802	44.15
Total	1,687	100.00	1,419	100.00	1,381	100.00	1,859	100.00	6,346	100.00

Table 6.4.3D

Equating Summary: Spek 9-12

No equating summary is presented because the Alternate ACCESS Series 501 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).



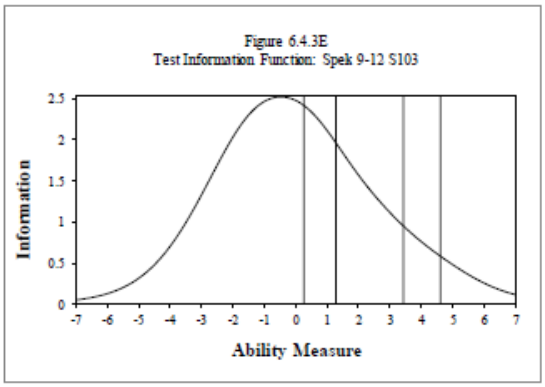


Table 6.4.3E

Reliability: Spek 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
6,346	8	0.967	2.192

Table 6.4.3F

Item Analysis Summary: Spek 9-12

[redacted]

Table 6.4.3G

Complete Item Analysis: Spek 9-12

[redacted]

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	11.08	11.08	A1	10.50	10.50	A1	10.28	10.28	A1	11.51	11.51
1	A1	1.13	12.21	A1	0.63	11.13	A1	1.01	11.30	A1	0.91	12.43
2	A1	0.95	13.16	A1	0.78	11.91	A1	0.80	12.09	A1	0.38	12.80
3	A1	1.42	14.58	A1	1.69	13.60	A1	1.01	13.11	A1	1.45	14.25
4	A1	0.36	14.94	A1	0.56	14.16	A1	0.36	13.47	A1	0.97	15.22
5	A1	0.59	15.53	A1	0.63	14.80	A1	0.51	13.98	A1	0.59	15.81
6	A1	0.95	16.48	A1	0.70	15.50	A1	0.94	14.92	A1	0.70	16.51
7	A1	0.77	17.25	A1	1.13	16.63	A1	0.72	15.64	A1	1.08	17.59
8	A1	0.89	18.14	A1	1.06	17.69	A1	1.23	16.87	A1	1.29	18.88
9	A2	1.30	19.44	A2	0.92	18.60	A2	1.01	17.89	A2	1.40	20.28
10	A2	1.54	20.98	A2	1.27	19.87	A2	1.38	19.26	A2	1.51	21.79
11	A3	2.37	23.36	A3	1.69	21.56	A3	2.32	21.58	A3	1.83	23.61
12	A3	4.21	27.56	A3	3.59	25.16	A3	3.77	25.34	A3	3.98	27.60
13	A3	4.15	31.71	A3	4.79	29.95	A3	3.04	28.39	A3	4.46	32.06
14	P1	8.83	40.55	P1	8.39	38.34	P1	7.82	36.21	P1	9.52	41.58
15	P1	17.25	57.79	P1	15.72	54.05	P1	17.67	53.87	P1	15.33	56.91
16	P2	42.21	100.00	P2	45.95	100.00	P2	46.13	100.00	P2	43.09	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

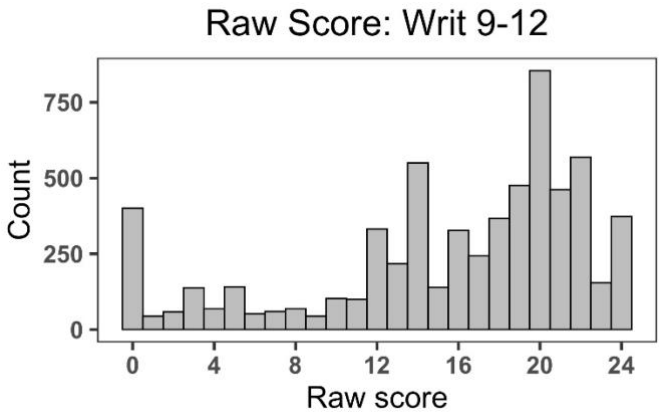


Table 6.4.4A
Raw Score Descriptive Statistics: Writ 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,689	0	24	15.49	6.72
10	1,419	0	24	15.53	6.69
11	1,380	0	24	15.91	6.68
12	1,856	0	24	15.44	6.85
Total	6,344	0	24	15.57	6.75

Figure 6.4.4B

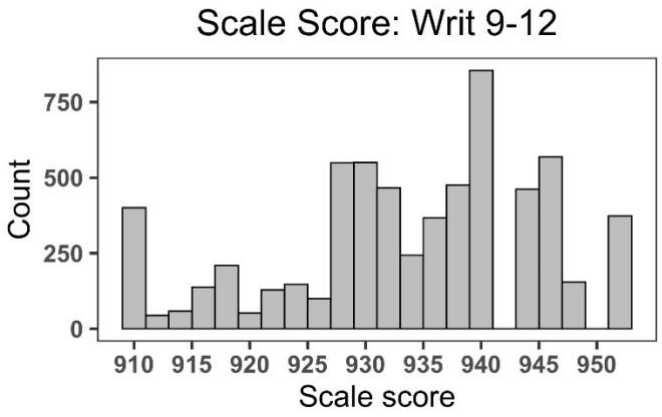


Table 6.4.4B
Scale Score Descriptive Statistics: Writ 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,689	910	953	934.10	11.04
10	1,419	910	953	934.26	11.00
11	1,380	910	953	934.80	11.07
12	1,856	910	953	934.05	11.28
Total	6,344	910	953	934.27	11.11

Figure 6.4.4C

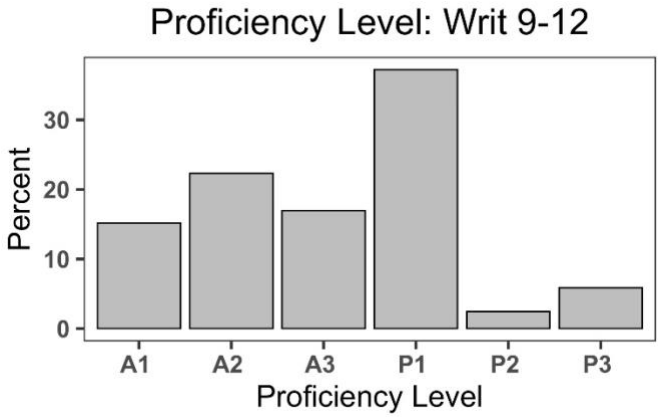


Table 6.4.4C

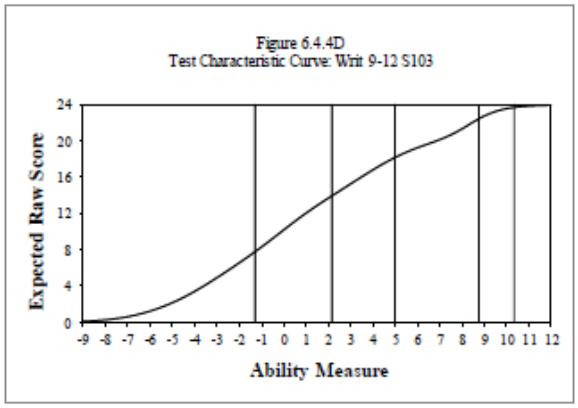
Proficiency Level Distribution: Writ 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	258	15.28	222	15.64	193	13.99	290	15.63	963	15.18
A2	386	22.85	324	22.83	295	21.38	410	22.09	1,415	22.30
A3	282	16.70	233	16.42	230	16.67	331	17.83	1,076	16.96
P1	637	37.71	521	36.72	538	38.99	666	35.88	2,362	37.23
P2	35	2.07	31	2.18	42	3.04	47	2.53	155	2.44
P3	91	5.39	88	6.20	82	5.94	112	6.03	373	5.88
Total	1,689	100.00	1,419	100.00	1,380	100.00	1,856	100.00	6,344	100.00

Table 6.4.4D

Equating Summary: Writ 9-12

No equating summary is presented because the Alternate ACCESS Series 501 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).



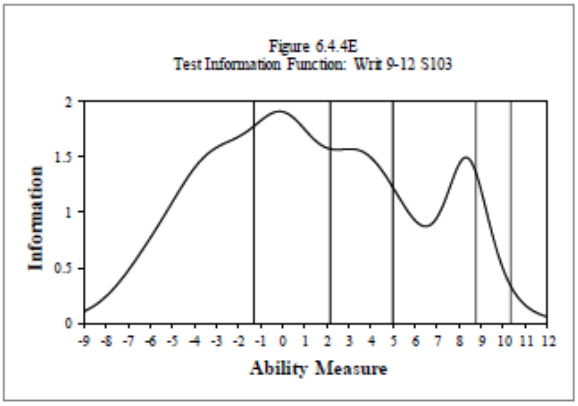


Table 6.4.4E

Reliability: Writ 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
6,344	10	0.917	3.193

Table 6.4.4F

Item Analysis Summary: Writ 9-12

Table 6.4.4G

Complete Item Analysis: Writ 9-12

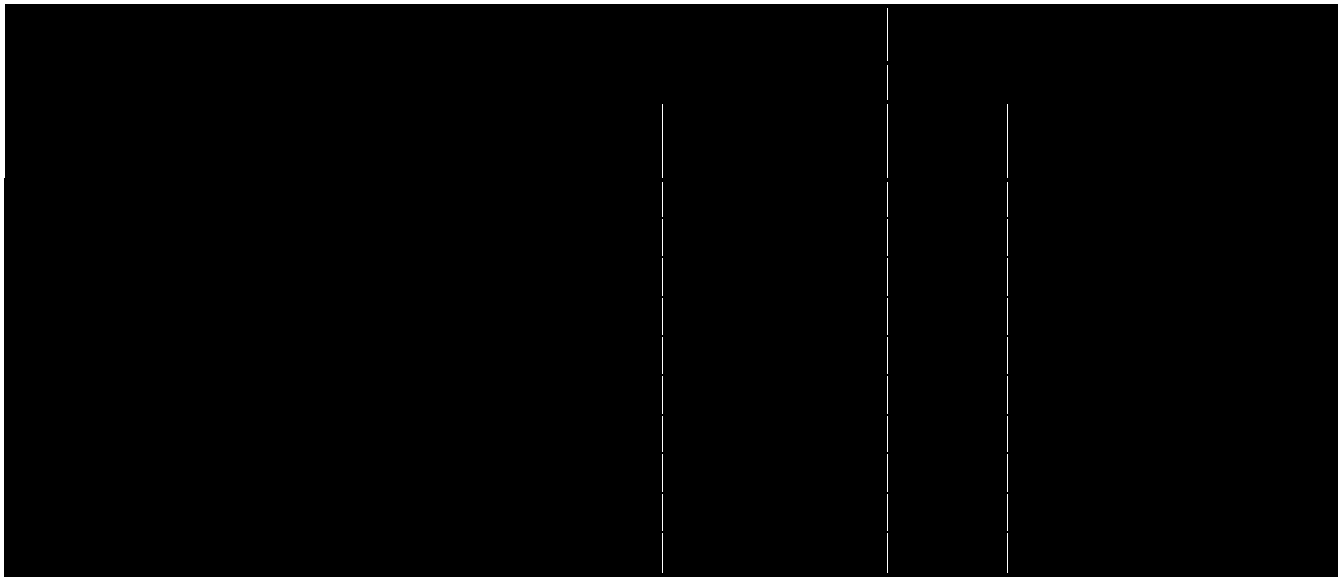


Table 6.4.4H

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	6.57	6.57	A1	5.00	5.00	A1	6.38	6.38	A1	7.06	7.06
1	A1	0.71	7.28	A1	0.49	5.50	A1	0.94	7.32	A1	0.65	7.70
2	A1	0.59	7.87	A1	1.41	6.91	A1	0.58	7.90	A1	1.08	8.78
3	A1	2.01	9.89	A1	2.68	9.58	A1	1.74	9.64	A1	2.26	11.05
4	A1	1.01	10.89	A1	1.48	11.06	A1	0.72	10.36	A1	1.13	12.18
5	A1	2.19	13.08	A1	2.75	13.81	A1	2.46	12.83	A1	1.67	13.85
6	A1	1.12	14.21	A1	0.49	14.31	A1	0.80	13.62	A1	0.81	14.66
7	A1	1.07	15.28	A1	1.34	15.64	A1	0.36	13.99	A1	0.97	15.63
8	A2	1.24	16.52	A2	0.99	16.63	A2	1.16	15.14	A2	0.97	16.59
9	A2	0.59	17.11	A2	0.70	17.34	A2	0.58	15.72	A2	0.86	17.46
10	A2	1.84	18.95	A2	1.76	19.10	A2	1.30	17.03	A2	1.56	19.02
11	A2	1.66	20.60	A2	1.41	20.51	A2	1.09	18.12	A2	1.99	21.01
12	A2	5.03	25.64	A2	6.41	26.92	A2	4.35	22.46	A2	5.17	26.19
13	A2	3.49	29.13	A2	3.45	30.37	A2	3.33	25.80	A2	3.39	29.58
14	A2	9.00	38.13	A2	8.10	38.48	A2	9.57	35.36	A2	8.14	37.72
15	A3	2.55	40.67	A3	1.97	40.45	A3	1.67	37.03	A3	2.42	40.14
16	A3	4.62	45.29	A3	5.00	45.45	A3	5.80	42.83	A3	5.28	45.42
17	A3	3.97	49.26	A3	3.45	48.91	A3	2.90	45.72	A3	4.69	50.11
18	A3	5.57	54.83	A3	5.99	54.90	A3	6.30	52.03	A3	5.44	55.55
19	P1	8.29	63.11	P1	7.47	62.37	P1	6.81	58.84	P1	7.33	62.88
20	P1	13.03	76.14	P1	13.04	75.41	P1	15.43	74.28	P1	12.77	75.65
21	P1	6.99	83.13	P1	7.40	82.80	P1	7.68	81.96	P1	7.17	82.81
22	P1	9.41	92.54	P1	8.81	91.61	P1	9.06	91.01	P1	8.62	91.43
23	P2	2.07	94.61	P2	2.18	93.80	P2	3.04	94.06	P2	2.53	93.97
24	P3	5.39	100.00	P3	6.20	100.00	P3	5.94	100.00	P3	6.03	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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

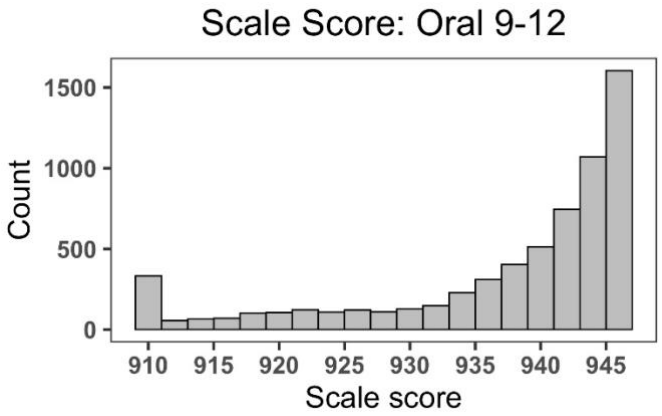


Table 6.4.5A

Scale Score Descriptive Statistics: Oral 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,685	910	946	937.39	10.52
10	1,418	910	946	938.09	10.24
11	1,379	910	946	938.21	10.21
12	1,859	910	946	937.46	10.74
Total	6,341	910	946	937.75	10.46

Figure 6.4.5B

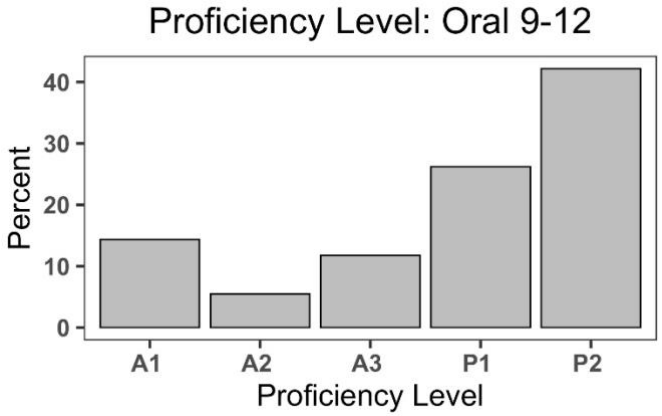


Table 6.4.5C

Proficiency Level Distribution: Oral 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	249	14.78	200	14.10	185	13.42	277	14.90	911	14.37
A2	10	5.99	72	5.08	77	5.58	97	5.22	347	5.47
A3	204	12.11	167	11.78	135	9.79	240	12.91	746	11.76
P1	462	27.42	336	23.70	381	27.63	483	25.98	1,662	26.21
P2	669	39.70	643	45.35	601	43.58	762	40.99	2,675	42.19
Total	1,685	100.00	1,418	100.00	1,379	100.00	1,859	100.00	6,341	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	102.262	0.943
Speaking	0.5	147.478	0.967
Oral		109.541	0.976

*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

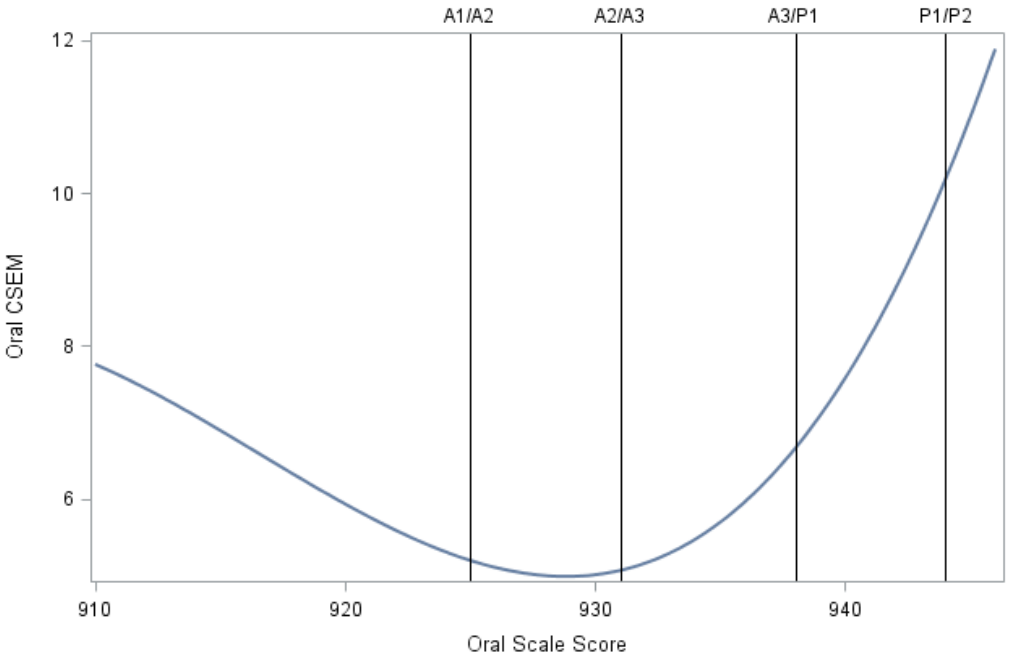
Table 6.4.5J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

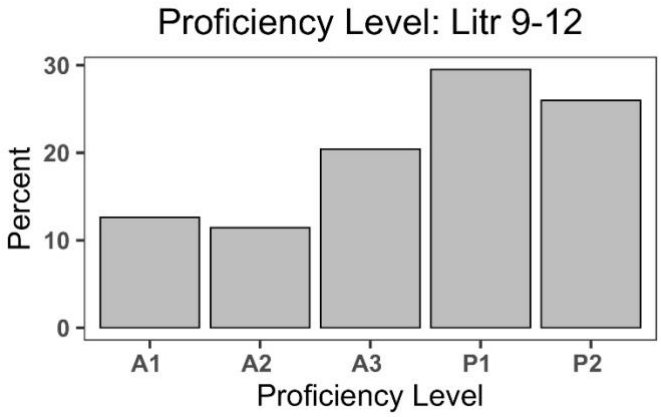


Table 6.4.6A

Scale Score Descriptive Statistics: Litr 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,687	910	951	936.27	10.31
10	1,418	910	951	936.69	10.10
11	1,380	910	951	937.06	10.18
12	1,856	910	951	936.23	10.57
Total	6,341	910	951	936.52	10.32

Figure 6.4.6B

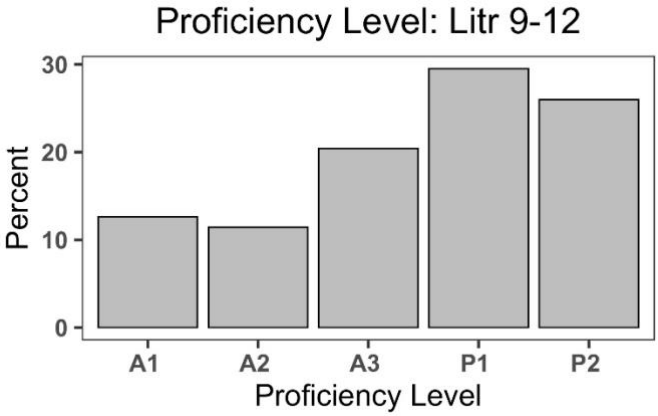


Table 6.4.6C

Proficiency Level Distribution: Litr 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	212	12.57	174	12.27	161	11.67	254	13.69	801	12.63
A2	208	12.33	157	11.07	147	10.65	214	11.53	726	11.45
A3	352	20.87	298	21.02	268	19.42	377	20.31	1,295	20.42
P1	508	30.11	418	29.48	415	30.07	531	28.61	1,872	29.52
P2	407	24.13	371	26.16	389	28.19	480	25.86	1,647	25.97
Total	1,687	100.00	1,418	100.00	1,380	100.00	1,856	100.00	6,341	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	116.061	0.944
Writing	0.5	123.508	0.917
Literacy		106.572	0.961

*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

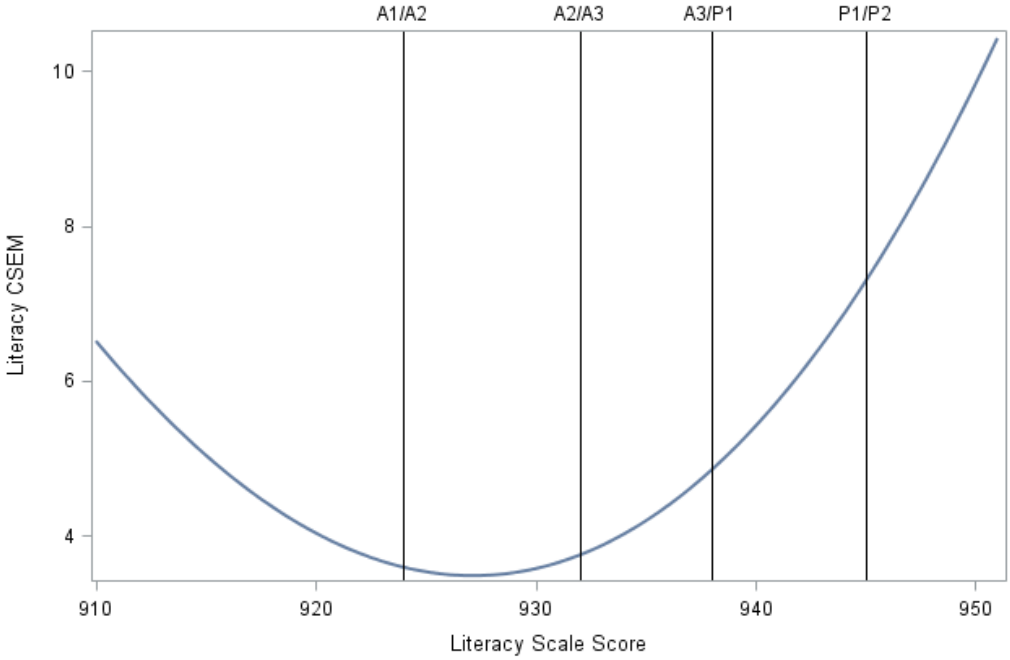
Table 6.4.6J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

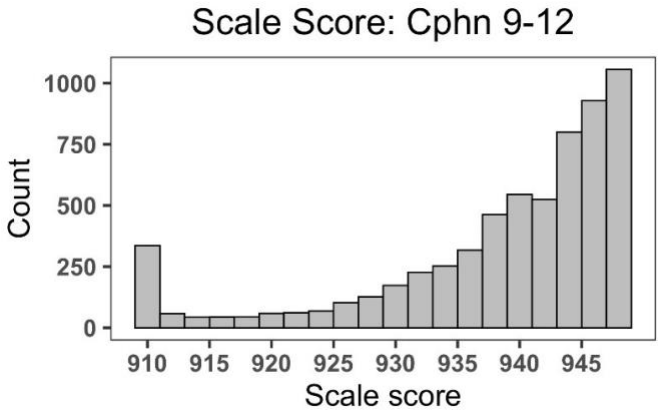


Table 6.4.7A
Scale Score Descriptive Statistics: Cphn 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,686	910	948	938.16	10.38
10	1,418	910	948	938.91	10.06
11	1,379	910	948	939.03	10.06
12	1,861	910	948	938.20	10.76
Total	6,344	910	948	938.53	10.36

Figure 6.4.7B

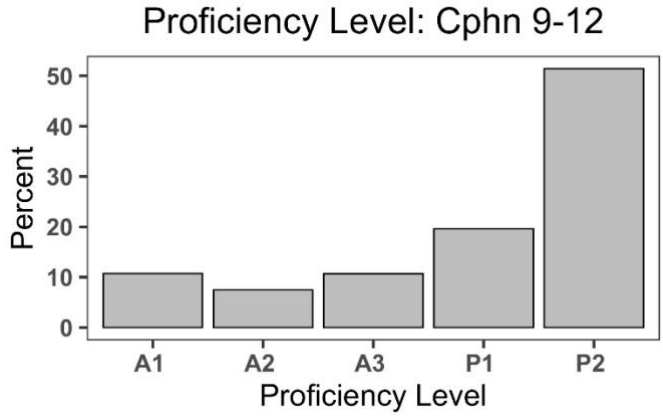


Table 6.4.7C

Proficiency Level Distribution: Cphn 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	188	11.15	145	10.23	136	9.86	213	11.45	682	10.75
A2	131	7.77	101	7.12	96	6.96	146	7.85	474	7.47
A3	186	11.03	144	10.16	135	9.79	215	11.55	680	10.72
P1	349	20.70	269	18.97	275	19.94	352	18.91	1,245	19.62
P2	832	49.35	759	53.53	737	53.44	935	50.24	3,263	51.43
Total	1,686	100.00	1,418	100.00	1,379	100.00	1,861	100.00	6,344	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	102.262	0.943
Reading	0.7	116.061	0.944
Comprehension		107.457	0.965

*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

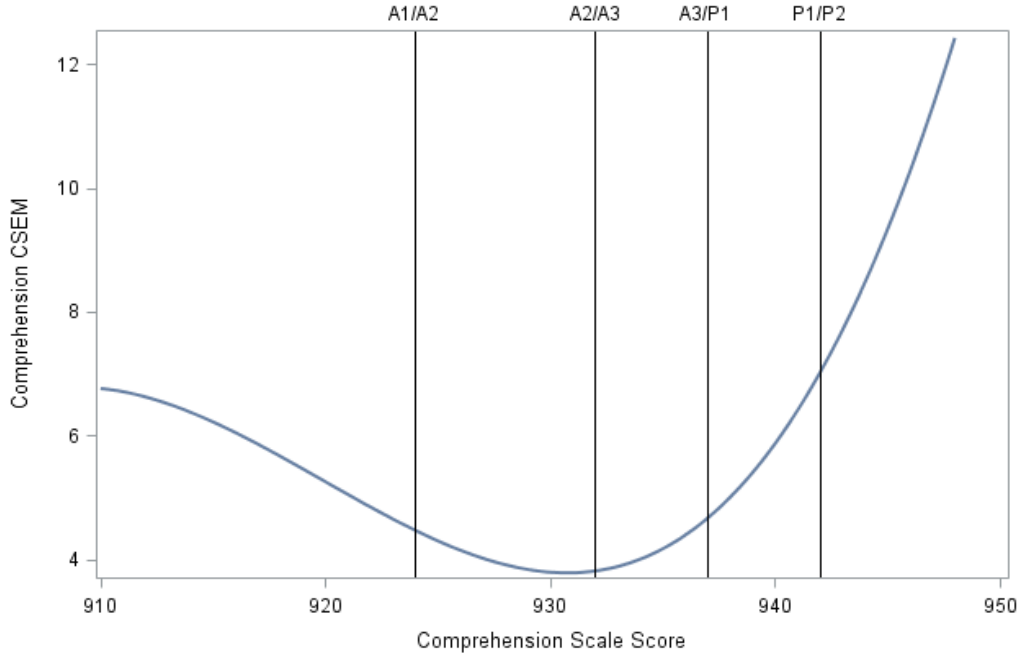
Table 6.4.7J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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

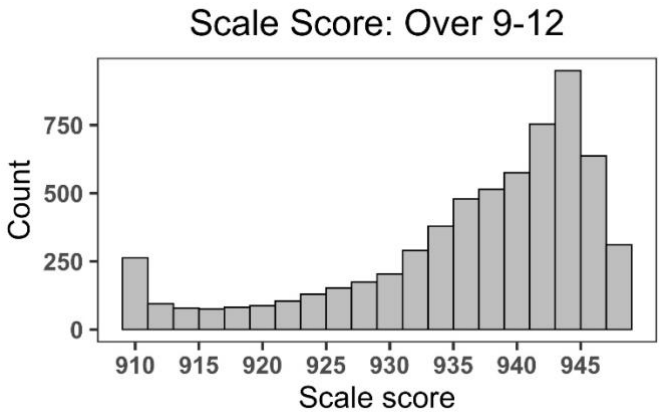


Table 6.4.8A

Scale Score Descriptive Statistics: Over 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1,685	910	949	936.44	10.12
10	1,418	910	949	936.95	9.84
11	1,379	910	949	937.22	9.93
12	1,855	910	949	936.43	10.37
Total	6,337	910	949	936.72	10.10

Figure 6.4.8B

Table 6.4.8C

Proficiency Level Distribution: Over 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	210	12.46	166	11.71	165	11.97	246	13.26	787	12.42
A2	154	9.14	132	9.31	108	7.83	166	8.95	560	8.84
A3	354	21.01	277	19.53	249	18.06	369	19.89	1,249	19.71
P1	493	29.26	410	28.91	410	29.73	530	28.57	1,843	29.08
P2	474	28.13	433	30.54	447	32.41	544	29.33	1,898	29.95
Total	1,685	100.00	1,418	100.00	1,379	100.00	1,855	100.00	6,337	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	102.262	0.943
Reading	0.35	116.061	0.944
Speaking	0.15	147.478	0.967
Writing	0.35	123.508	0.917
Overall Composite		102.005	0.978

*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

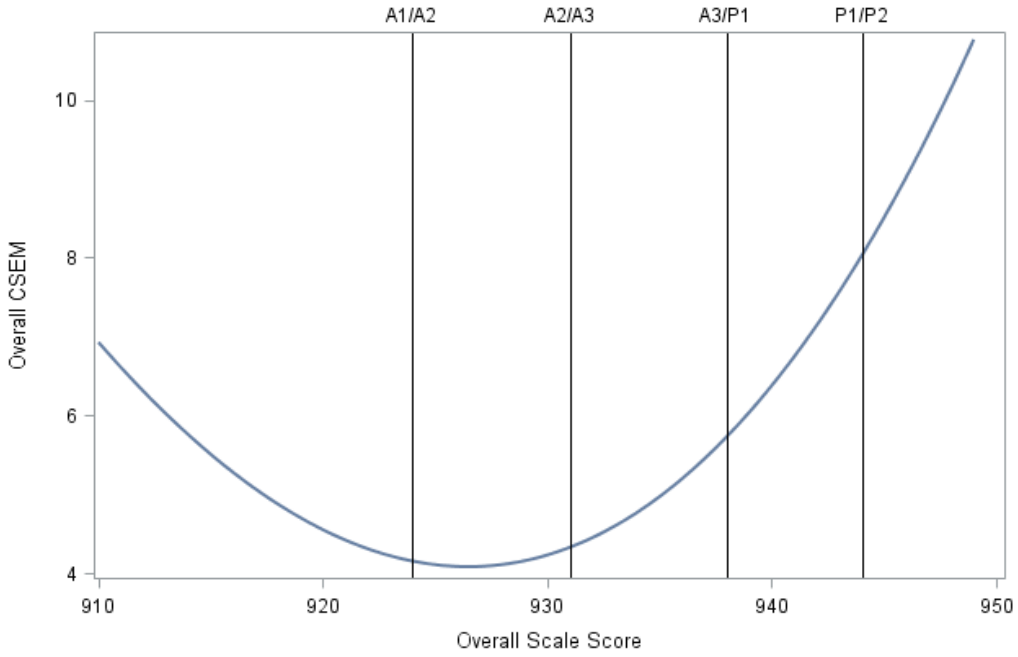
Table 6.4.8J

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		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	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.