Annual Technical Report for

**ALTERNATE ACCESS for ELLs® English Language Proficiency
Test, Series 503, 2021-2022 Administration**

Annual Technical Report No. 10

Prepared by:

Psychometrics team

WIDA

January, 2023

Table of Contents

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 content area of **Language Arts**.

---

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

Standard 3—Language of Mathematics: ELLs communicate information, ideas, and concepts necessary for academic success in the content area of **Mathematics**.

Standard 4—Language of Science: ELLs communicate information, ideas, and concepts necessary for academic success in the content area of **Science.**

For practical purposes, the four Standards are abbreviated as follows in this report:
- Social and Instructional language: SI
- Language of English Language Arts: LA
- Language of Mathematics: MA
- Language of Science: SC

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

1.2.1. Grade-level Clusters

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

1.2.2. Language Domains

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

---

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

1.2.3. Language Proficiency Levels

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

ACCESS. In this figure, PL1 has been stretched for illustrative purposes to display levels A1 – A3.
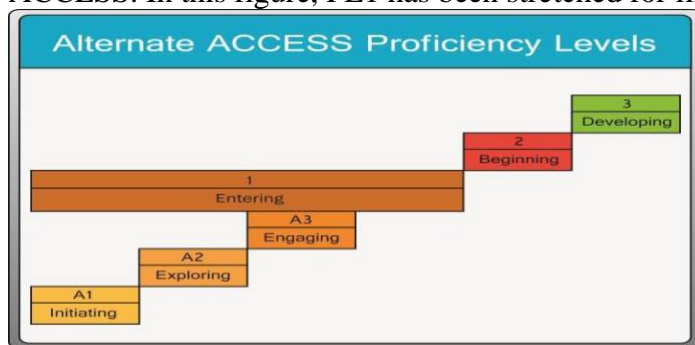
Figure 1.2.4A. Alternate ACCESS Proficiency Levels

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

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

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

Each selected-response item or performance-based task on Alternate ACCESS is carefully developed, reviewed, piloted, and field tested 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 by errors when text becomes more complex. Text is related to the task.
2-Emerging	One or more simple phrases. Text is original or adapted from model or source text. Comprehensible when text is adapted from model or source text. Comprehensibility may be impeded by errors in original text. Text is related to the task.
1-Entering	One or more general content words. Text is original or adapted from the model or source text. Generally comprehensible when text is adapted from model or source text. Comprehensibility may be significantly impeded in original text. Text is related to the task.
A3-Engaging	Single words and numbers. All or part of text is copied. If original text is present, it is not related to the task. Comprehensibility of the text may be significantly impeded by imprecise letter, symbol, or number formation. Text may or may not be related to the task.
A2-Exploring	Common single-digit numbers, letters, symbols, or syllables. All or part of text is copied. Comprehensibly 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-Initating	Pictorial representations and imprecise, but intentional markings such as drawing and scribbles. Representations may or may not be related to the task.

1.6.3. Speaking

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

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

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

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

Table 1.6.3A

Alternate ACCESS Speaking Rubric

Level	Text Features
2-Emerging	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-Initating	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.
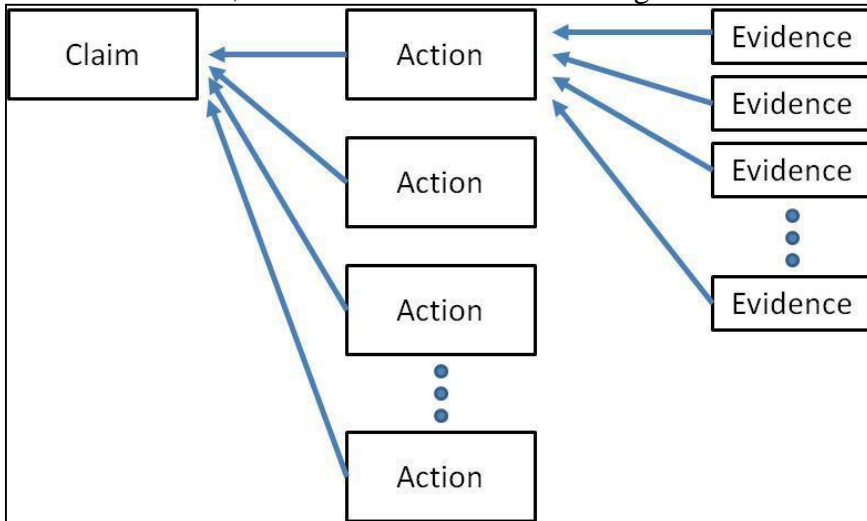
Figure 2A: General Argument Structure for Assessment Validation

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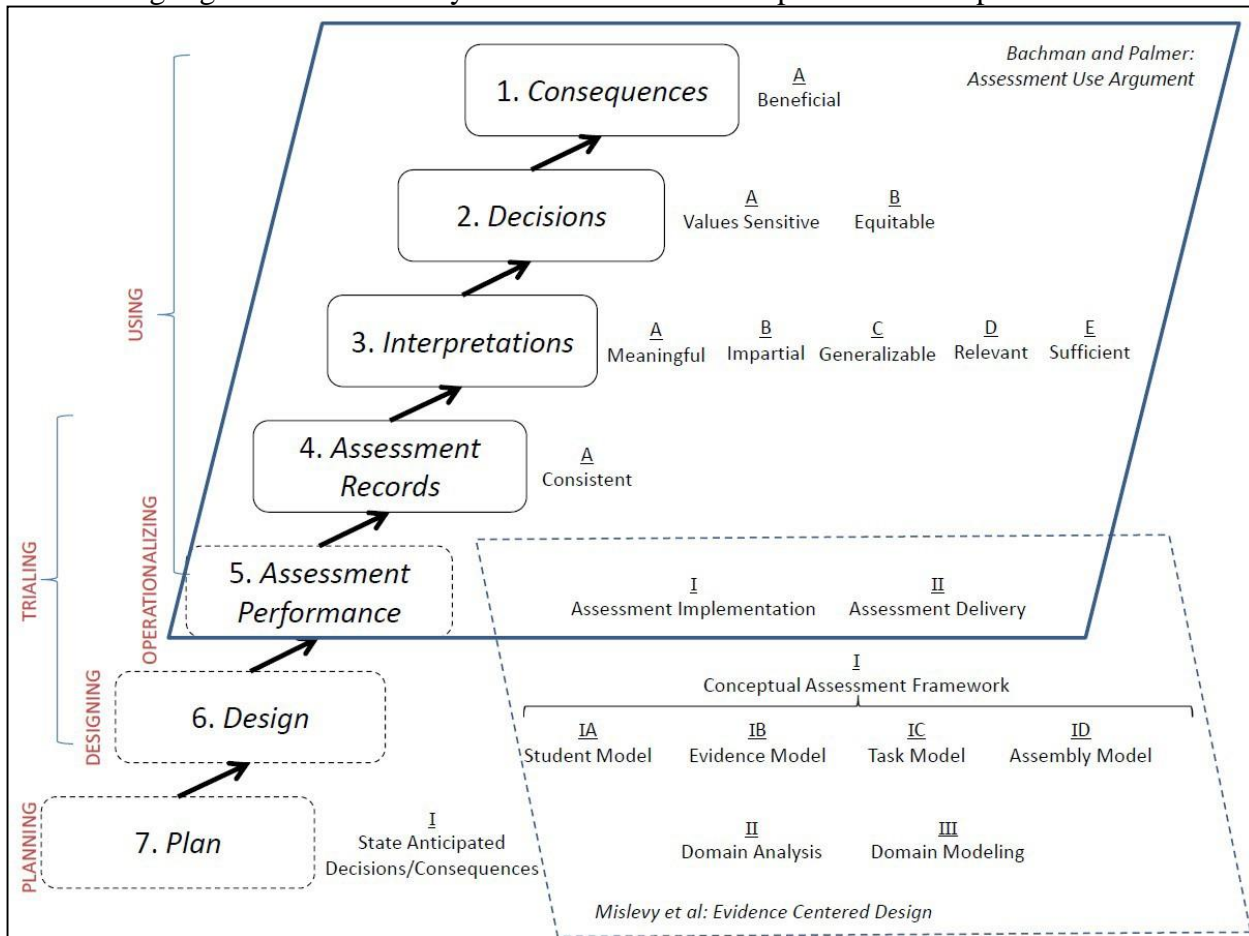
Figure 2.1A: CAL's Validation Framework (based on Bachman & Palmer, 2010; Mislevy, Almond, & Lukas, 2004)

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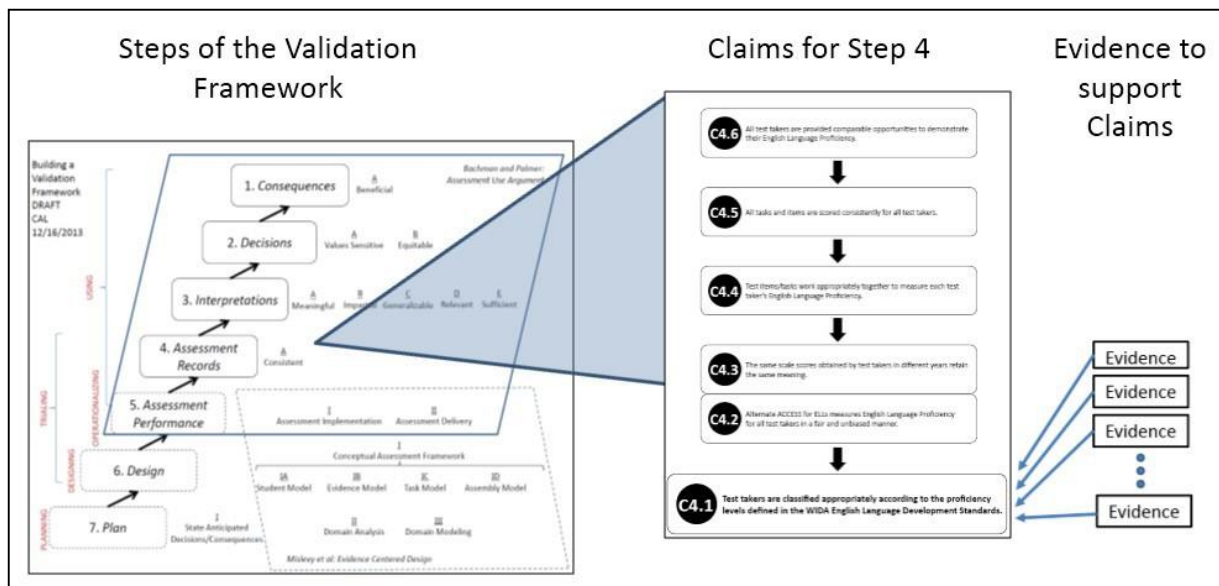
Figure 2.2A: Structure of the Argument-Based Approach Supporting Step 4 Contained in this Chapter

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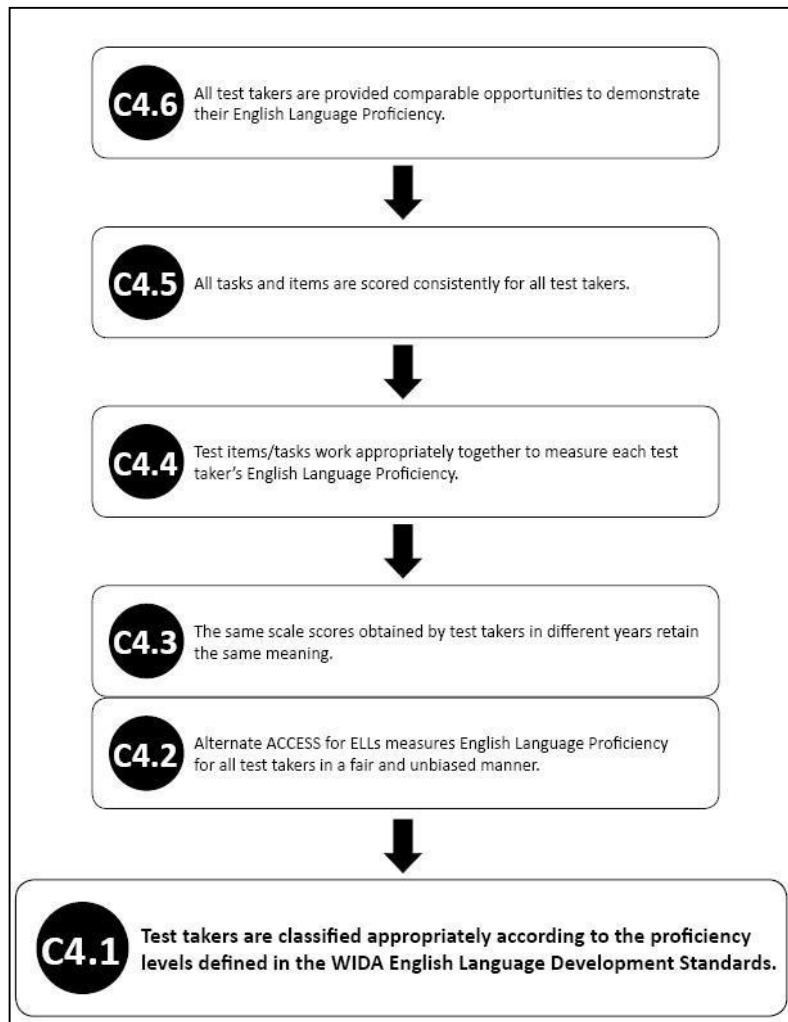
Figure 2.2.1A: Progression of Claims for Step 4: Assessment Records

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, S502).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

The Complete Items Analysis table includes more detailed information on the DIF analyses, showing the degree of measurement bias for each item and which group is favored (Table 6G). Each item is categorized into three levels of DIF: A, B, or C (Zieky, 1993). An item exhibiting A level DIF shows little or no evidence of bias toward a particular group, an item exhibiting B level DIF 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 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

4.1	Students Excluded from Analysis
4.2	Participation by Grade-Level Cluster
4.3	Participation by Grade
4.4	Participation by Domain
4.5	Scale Scores by Domain
4.6	Scale Scores by Grade-Level Cluster
4.7	Scale Scores by Grade
4.8	Correlations among Scale Scores by Grade-Level Cluster
4.9	Proficiency Level Results
4.10	Participation by Disability

Test Form Characteristics

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

Figure 2.5.1 Chapter 4 Visual Guide to Tables and Figures

2.5.2 Chapter 6 Visual Guide to Tables and Figures

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

		Grade-Level Cluster			
		1-2	3-5	6-8	9-12
Domain Composites	Overall	6.1.8	6.2.8	6.3.8	6.4.8
	Oral Language	6.1.5	6.2.5	6.3.5	6.4.5
	Literacy	6.1.6	6.2.6	6.3.6	6.4.6
	Comprehension	6.1.7	6.2.7	6.3.7	6.4.7
Domains	Listening	6.1.1	6.2.1	6.3.1	6.4.1
	Reading	6.1.2	6.2.2	6.3.2	6.4.2
	Writing	6.1.4	6.2.4	6.3.4	6.4.4
	Speaking	6.1.3	6.2.3	6.3.3	6.4.3
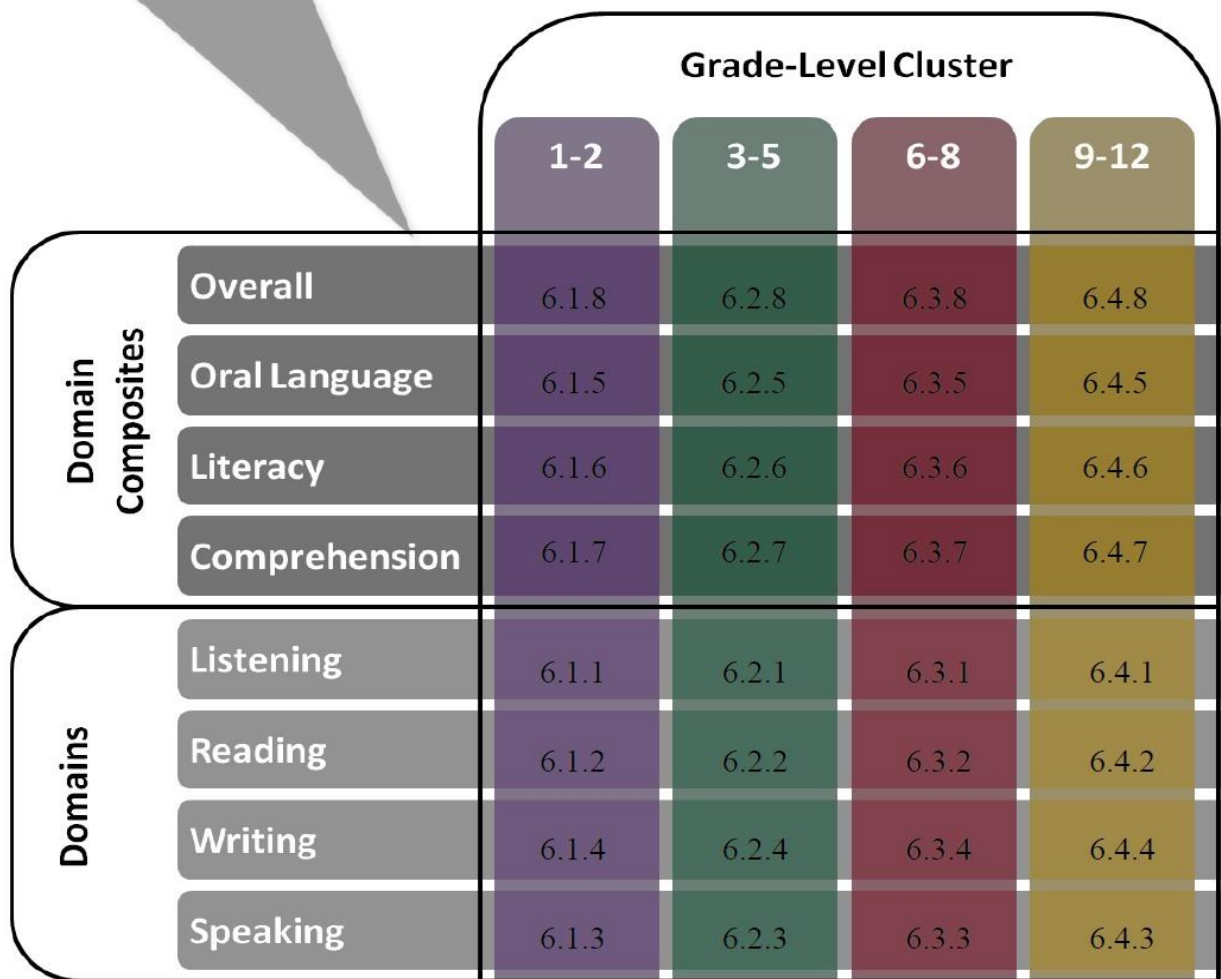
Figure 2.5.2 Chapter 6 Visual Guide to Tables and Figures

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 41 WIDA Consortium states/territories participated in the 2021-2022 Alternate ACCESS operational administration. The data used for analyses of the 503 Alternate ACCESS technical report is 27,048, which was drawn in September 2022. The final number of students that have taken the 503 Alternate ACCESS tests is 27,529 as of January 2023.

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, 30 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 30 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 41 WIDA states who participated in the 2021-2022 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
  C. Speaking

---

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

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

Grade	Cluster 1-2	Cluster 3-5	Cluster 6-8	Cluster 9-12	Total
1		1	0	0	0
2		5	0	0	0
3	4		0	0	0
4	2		0	0	0
5	1		2	0	0
6	0	5		2	0
7	0	0		2	0
8	0	0		2	0
9	0	0	3		0
10	0	0	1		0
11	0	0	0		0
12	0	0	0		0
Total	7	11	6	6	30

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				
	1-2	3-5	6-8	9-12	Total
AK	6	25	31	37	99
AL	62	89	61	62	274
BI	2	9	14	8	33
CO	135	271	214	199	819
DC	12	35	23	30	100
DD	2	9	1	4	16
DE	6	6	9	8	29
FL	304	376	139	138	957
GA	265	424	344	315	1348
HI	41	73	59	65	238
ID	25	48	51	35	159
IL	951	1220	1032	1386	4589
IN	202	277	298	457	1234
KY	82	86	83	98	349
MA	371	466	392	396	1625
MD	125	224	217	184	750
ME	17	17	12	25	71
MI	152	225	156	198	731
MN	273	352	253	291	1169
MO	55	47	58	45	205
MP	0	4	0	1	5
MT	2	13	6	0	21
NC	228	472	486	533	1719
ND	3	5	6	8	22
NH	4	8	8	7	27
NJ	204	213	99	58	574
NM	82	148	135	135	500
NV	127	227	250	332	936
OK	151	245	210	137	743
PA	294	366	325	337	1322
RI	29	51	46	55	181
SC	97	130	70	113	410
SD	5	17	8	13	43
TN	93	137	103	85	418
UT	95	169	144	157	565
VA	510	539	461	694	2204
VI	1	0	0	0	1
VT	9	11	0	8	28
WA	466	556	410	569	2001
WI	59	134	132	183	508
WY	4	12	2	7	25
Total	5551	7736	6348	7413	27048

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	1451	26.14%	3479	62.67%	621	11.19%	5551
3-5	2238	28.98%	4752	61.43%	746	9.64%	7736
6-8	2019	31.81%	3745	58.99%	584	9.20%	6348
9-12	2391	32.25%	4148	55.96%	874	11.79%	7413
Total	8099	29.94%	16124	59.61%	2825	10.44%	27048

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	2993	56.58%	2008	36.17%	402	7.24%	5551
3-5	4521	60.65%	2531	32.72%	513	6.63%	7736
6-8	3909	64.35%	1847	29.10%	416	6.55%	6348
9-12	4534	63.93%	2211	29.83%	463	6.25%	7413
Total	16657	61.58%	8597	31.78%	1794	6.63%	27048

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	3	3	7	11	7	7	11	13	9	15	6	7	99
AL	24	38	32	35	22	24	19	18	25	15	10	12	274
BI	0	2	2	4	3	3	5	6	2	4	2	0	33
CO	61	74	103	92	76	84	69	61	46	52	42	59	819
DC	8	4	11	7	17	7	10	6	7	6	7	10	100
DD	1	1	3	4	2	1	0	0	1	1	2	0	16
DE	3	3	3	0	3	3	3	3	0	3	2	3	29
FL	140	164	144	120	112	57	38	44	32	40	31	35	957
GA	122	143	153	147	124	111	131	102	102	72	52	89	1348
HI	23	18	21	25	27	26	14	19	18	14	14	19	238
ID	17	8	18	10	20	20	13	18	9	6	8	12	159
IL	464	487	449	402	369	355	339	338	291	280	254	561	4589
IN	109	93	103	85	89	77	112	109	103	107	94	153	1234
KY	38	44	32	25	29	26	30	27	23	28	26	21	349
MA	202	169	190	145	131	138	127	127	110	80	95	111	1625
MD	61	64	76	67	81	83	62	72	54	44	44	42	750
ME	10	7	7	8	2	5	2	5	6	5	7	7	71
MI	81	71	81	78	66	65	50	41	42	64	48	44	731
MN	112	161	131	114	107	86	100	67	75	59	73	84	1169
MO	26	29	19	19	9	22	20	16	9	13	9	14	205
MP	0	0	1	1	2	0	0	0	1	0	0	0	5
MT	0	2	5	2	6	1	4	1	0	0	0	0	21
NC	101	127	142	149	181	183	152	151	131	97	109	196	1719
ND	2	1	2	2	1	1	2	3	1	2	2	3	22
NH	0	4	3	4	1	1	4	3	2	2	1	2	27
NJ	112	92	82	81	50	35	34	30	17	16	18	7	574
NM	40	42	52	45	51	43	39	53	38	27	36	34	500
NV	49	78	72	77	78	82	78	90	87	84	77	84	936
OK	63	88	82	82	81	75	72	63	31	29	31	46	743
PA	142	152	124	133	109	121	109	95	80	65	80	112	1322
RI	19	10	19	16	16	15	14	17	13	16	7	19	181
SC	50	47	47	50	33	25	18	27	29	25	23	36	410
SD	2	3	7	6	4	4	3	1	5	0	6	2	43
TN	40	53	50	39	48	35	27	41	30	21	24	10	418
UT	37	58	57	55	57	49	47	48	39	47	39	32	565
VA	247	263	194	185	160	143	159	159	151	131	146	266	2204
VI	0	1	0	0	0	0	0	0	0	0	0	0	1
VT	4	5	3	6	2	0	0	0	2	2	1	3	28
WA	252	214	194	192	170	122	144	144	130	108	106	225	2001
WI	29	30	36	46	52	42	49	41	44	30	28	81	508
WY	1	3	3	5	4	0	2	0	1	3	1	2	25

State	Grade													Total
	1	2	3	4	5	6	7	8	9	10	11	12		
Total	2695	2856	2760	2574	2402	2177	2112	2059	1796	1613	1561	2443		27048

4.3.2 Participation by Grade by Gender

Table 4.3.2

Participation by Grade by Gender

	Gender						Total
	Female		Male		Missing		
Grade	Count	% within Grade	Count	% within Grade	Count	% within Grade	Total
1	670	24.86%	1689	62.67%	336	12.47%	2695
2	781	27.35%	1790	62.68%	285	9.98%	2856
3	731	26.49%	1756	63.62%	273	9.89%	2760
4	759	29.49%	1572	61.07%	243	9.44%	2574
5	748	31.14%	1424	59.28%	230	9.58%	2402
6	680	31.24%	1310	60.17%	187	8.59%	2177
7	666	31.53%	1247	59.04%	199	9.42%	2112
8	673	32.69%	1188	57.70%	198	9.62%	2059
9	581	32.35%	1008	56.12%	207	11.53%	1796
10	513	31.80%	940	58.28%	160	9.92%	1613
11	507	32.48%	895	57.34%	159	10.19%	1561
12	790	32.34%	1305	53.42%	348	14.24%	2443
Total	8099	29.94%	16124	59.61%	2825	10.44%	27048

4.3.3 Participation by Grade by Ethnicity

Table 4.3.3

Participation by Grade by Ethnicity

	Hispanic/Non-Hispanic						Total
	Hispanic		Non-Hispanic		Missing		
Grade	Count	% within Grade	Count	% within Grade	Count	% within Grade	Total
1	1524	56.55%	970	35.99%	201	7.46%	2695
2	1617	56.62%	1038	36.34%	201	7.04%	2856
3	1621	58.73%	958	34.71%	181	6.56%	2760
4	1561	60.64%	832	32.32%	181	7.03%	2574
5	1510	62.86%	741	30.85%	151	6.29%	2402
6	1354	62.20%	652	29.95%	171	7.85%	2177

| Grade | Hispanic/Non-Hispanic | | | | | | Total |
| | Hispanic | | Non-Hispanic | | Missing | | |
	Count	% within Grade	Count	% within Grade	Count	% within Grade	
7	1360	64.39%	614	29.07%	138	6.53%	2112
8	1371	66.59%	581	28.22%	107	5.20%	2059
9	1167	64.98%	510	28.40%	119	6.63%	1796
10	1040	64.48%	483	29.94%	90	5.58%	1613
11	985	63.10%	490	31.39%	86	5.51%	1485
12	1547	63.32%	728	29.80%	168	6.88%	2333
Total	16657	61.58%	8597	31.78%	1794	6.63%	27048

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	5499	5459	5420	5390
3-5	7673	7640	7581	7570
6-8	6300	6288	6229	6157
9-12	7334	7328	7270	7181
Total	26806	26715	26500	26298

4.4.2   Participation by Grade by Domain

Table 4.4.2

Participation by Grade by Domain

| Grade | Domain | | | |
	Listening	Reading	Speaking	Writing
1	2668	2649	2626	2616
2	2831	2810	2794	2774
3	2728	2713	2693	2705
4	2554	2547	2528	2522
5	2391	2380	2360	2343
6	2164	2159	2141	2126
7	2093	2086	2065	2032
8	2043	2043	2023	1999

Grade	Domain			
	Listening	Reading	Speaking	Writing
9	1780	1774	1757	1739
10	1590	1596	1587	1563
11	1541	1537	1525	1511
12	2423	2421	2401	2368
Total	26806	26715	26500	26298

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	929.76	11.56	2668	932.38	11.04	2831	931.11	11.37	5499
Reading	930.09	13.02	2649	933.10	13.07	2810	931.64	13.13	5459
Speaking	930.03	14.95	2626	933.22	14.77	2794	931.67	14.94	5420
Writing	925.81	11.02	2616	928.41	11.44	2774	927.15	11.31	5390
Oral	930.24	12.44	2610	933.12	12.12	2784	931.73	12.36	5394
Literacy	928.30	11.19	2593	931.10	11.45	2745	929.74	11.41	5338
Comprehension	930.12	12.21	2640	932.97	12.09	2807	931.59	12.23	5447
Overall	928.71	11.13	2565	931.54	11.17	2725	930.17	11.24	5290

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	934.39	11.07	2728	936.58	10.33	2554	937.14	10.44	2391	935.97	10.70	7673
Reading	933.28	11.37	2713	935.85	10.44	2547	936.47	10.74	2380	935.13	10.96	7640
Speaking	934.04	13.91	2693	936.21	13.12	2528	936.40	13.34	2360	935.50	13.52	7581
Writing	929.51	11.21	2705	931.75	11.38	2522	932.37	11.74	2343	931.14	11.50	7570
Oral	934.44	11.57	2682	936.59	10.86	2516	936.99	10.95	2353	935.95	11.20	7551
Literacy	931.73	10.60	2673	934.14	10.21	2506	934.78	10.51	2333	933.48	10.53	7512
Comprehension	933.67	10.91	2703	936.12	10.06	2538	936.74	10.30	2375	935.44	10.53	7616
Overall	932.37	10.51	2645	934.71	10.01	2484	935.31	10.24	2314	934.07	10.34	7443

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	936.40	10.61	2164	937.42	10.53	2093	937.69	10.34	2043	937.16	10.51	6300
Reading	936.74	11.86	2159	937.74	11.77	2086	937.97	11.98	2043	937.47	11.88	6288
Speaking	935.96	13.09	2141	936.76	12.92	2065	936.27	13.38	2023	936.32	13.13	6229
Writing	931.31	10.25	2126	932.25	10.54	2032	932.39	10.57	1999	931.97	10.46	6157
Oral	936.60	11.22	2136	937.59	11.08	2053	937.44	11.27	2010	937.20	11.20	6199
Literacy	934.31	10.26	2117	935.37	10.33	2021	935.51	10.62	1990	935.05	10.42	6128
Comprehension	936.70	11.13	2152	937.74	11.03	2077	937.96	11.15	2034	937.46	11.12	6263
Overall	934.79	10.18	2101	935.87	10.18	2002	935.95	10.41	1969	935.52	10.27	6072

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	937.23	10.41	1780	937.88	10.52	1590	938.56	10.03	1541	938.28	10.29	2423	938.00	10.32	7334
Reading	937.30	10.88	1774	938.12	10.82	1596	938.74	10.50	1537	938.17	11.12	2421	938.07	10.88	7328
Speaking	935.71	12.50	1757	936.17	12.17	1587	936.61	11.87	1525	935.77	12.74	2401	936.02	12.38	7270
Writing	932.92	10.72	1739	933.95	10.74	1563	934.18	10.86	1511	933.78	10.94	2368	933.69	10.83	7181
Oral	936.71	10.67	1747	937.27	10.60	1573	937.87	10.17	1515	937.27	10.69	2390	937.26	10.56	7225
Literacy	935.42	10.07	1728	936.41	10.05	1554	936.76	10.03	1504	936.28	10.40	2356	936.20	10.18	7142
Comprehension	937.42	10.40	1766	938.16	10.45	1582	938.82	10.07	1524	938.33	10.63	2409	938.17	10.43	7281
Overall	935.70	9.93	1713	936.58	9.88	1538	936.97	9.77	1485	936.49	10.16	2333	936.42	9.97	7069

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	931.03	11.25	1434	931.21	11.46	3448	930.79	11.13	617
Reading	931.28	12.77	1419	931.84	13.32	3423	931.35	12.93	617
Speaking	930.63	15.01	1408	932.14	14.89	3401	931.49	14.94	611
Writing	926.52	10.87	1401	927.53	11.47	3389	926.44	11.32	600
Oral	931.19	12.22	1402	931.99	12.43	3384	931.48	12.23	608
Literacy	929.29	10.97	1384	930.01	11.59	3356	929.25	11.38	598
Comprehension	931.32	11.99	1419	931.77	12.38	3414	931.23	11.97	614
Overall	929.69	10.89	1376	930.43	11.39	3326	929.80	11.12	588

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	936.24	10.48	2220	936.06	10.77	4714	934.65	10.82	739
Reading	935.03	10.73	2208	935.42	11.07	4696	933.54	10.84	736
Speaking	935.38	13.55	2194	935.78	13.39	4656	934.09	14.14	731
Writing	930.64	11.22	2186	931.64	11.61	4659	929.44	11.40	725
Oral	936.03	11.10	2184	936.13	11.18	4641	934.60	11.51	726
Literacy	933.19	10.34	2168	933.88	10.61	4623	931.83	10.39	721
Comprehension	935.46	10.31	2200	935.66	10.63	4682	933.97	10.45	734
Overall	933.89	10.18	2148	934.39	10.40	4581	932.52	10.29	714

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	937.65	10.35	2002	937.29	10.39	3718	934.62	11.46	580
Reading	937.66	11.42	1999	937.72	11.98	3720	935.14	12.57	569
Speaking	936.65	12.86	1983	936.50	13.09	3678	934.05	14.15	568
Writing	932.11	10.43	1962	932.23	10.40	3647	929.76	10.74	548
Oral	937.61	10.95	1971	937.32	11.18	3664	935.00	11.94	564
Literacy	935.24	10.14	1951	935.27	10.47	3634	932.88	10.80	543
Comprehension	937.75	10.75	1991	937.66	11.16	3703	935.14	11.83	569
Overall	935.80	9.97	1933	935.71	10.30	3598	933.31	10.81	541

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	937.93	10.47	2359	938.22	10.22	4109	937.10	10.36	866
Reading	937.98	10.83	2359	938.31	10.88	4107	937.13	10.92	862
Speaking	936.08	12.25	2339	936.26	12.37	4077	934.73	12.76	854
Writing	933.60	10.92	2308	933.86	10.77	4045	933.14	10.88	828
Oral	937.26	10.57	2323	937.49	10.53	4055	936.18	10.65	847
Literacy	936.14	10.22	2293	936.38	10.13	4024	935.52	10.25	825
Comprehension	938.11	10.45	2340	938.39	10.42	4084	937.31	10.38	857
Overall	936.37	10.00	2265	936.60	9.93	3988	935.66	10.04	816

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	930.74	931.97	931.48	928.15	931.43	930.45	931.70	930.61
	Std. Dev.	11.16	13.48	15.03	11.79	12.16	11.66	12.41	11.35
	N	874	872	863	857	857	851	870	840
Non-Hispanic Pacific Islander	Mean	928.85	927.82	930.72	927.65	930.25	928.47	928.15	928.90
	Std. Dev.	13.22	13.89	14.70	11.12	13.20	11.85	13.38	11.95
	N	34	33	32	31	32	30	33	30
Non-Hispanic Black	Mean	929.86	930.78	931.11	927.55	930.83	929.51	930.61	929.65
	Std. Dev.	11.33	13.78	15.10	11.98	12.46	12.03	12.63	11.72
	N	505	498	494	485	493	482	498	481
Hispanic (Of Any Race)	Mean	931.46	931.84	931.86	927.03	931.99	929.77	931.83	930.28
	Std. Dev.	11.44	13.04	14.89	10.98	12.44	11.28	12.20	11.17
	N	3107	3087	3060	3055	3046	3023	3080	2993
Non-Hispanic American Indian	Mean	932.77	931.76	934.12	927.88	933.42	929.59	932.06	930.59
	Std. Dev.	11.60	11.35	14.40	9.59	12.15	9.61	11.06	10.02
	N	35	34	33	33	33	32	34	32
Non-Hispanic Multi-racial	Mean	930.47	928.50	923.50	923.06	927.25	926.06	929.25	926.16
	Std. Dev.	10.35	11.74	15.07	9.38	11.05	9.60	11.15	9.45
	N	32	32	32	32	32	32	32	32
Non-Hispanic White	Mean	930.30	930.48	930.66	925.24	930.79	928.11	930.53	928.71
	Std. Dev.	11.31	12.71	15.03	11.53	12.31	11.13	11.92	11.05
	N	514	508	512	505	509	501	508	499
Missing	Mean	931.99	932.54	933.23	928.09	933.04	930.77	932.55	931.30
	Std. Dev.	11.13	12.84	14.61	11.61	11.96	11.48	11.92	11.11
	N	398	395	394	392	392	387	392	383

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	934.69	934.74	934.76	932.24	934.93	933.83	934.75	933.91
	Std. Dev.	10.48	11.32	13.71	11.97	11.14	10.95	10.73	10.64
	N	1070	1068	1062	1057	1059	1055	1065	1049
Non-Hispanic Pacific Islander	Mean	934.64	933.09	936.36	929.28	935.67	931.40	933.65	932.47
	Std. Dev.	11.27	11.21	12.37	12.08	11.16	10.63	10.98	10.33
	N	58	57	58	57	57	57	57	57
Non-Hispanic Black	Mean	934.33	933.57	935.55	930.72	935.16	932.43	933.89	933.16
	Std. Dev.	11.09	11.61	13.59	11.85	11.48	10.95	11.05	10.69
	N	566	562	557	558	554	549	560	542
Hispanic (Of Any Race)	Mean	936.45	935.46	935.53	931.03	936.21	933.60	935.83	934.22
	Std. Dev.	10.61	10.73	13.49	11.32	11.14	10.34	10.34	10.20
	N	4654	4634	4596	4597	4577	4562	4619	4521
Non-Hispanic American Indian	Mean	939.85	938.98	938.03	933.65	939.10	936.93	939.25	937.71
	Std. Dev.	8.91	9.34	12.45	11.50	10.24	9.00	8.92	8.83
	N	61	61	60	62	59	60	61	58
Non-Hispanic Multi-racial	Mean	934.27	931.48	933.60	928.81	934.33	930.36	932.21	931.35
	Std. Dev.	11.43	12.63	13.68	11.58	11.63	11.53	11.94	11.19
	N	49	48	48	48	48	47	48	46
Non-Hispanic White	Mean	934.83	933.75	934.22	929.78	934.70	932.13	934.15	932.82
	Std. Dev.	11.14	11.31	14.25	11.81	11.84	10.92	10.89	10.78
	N	707	704	697	690	696	686	703	679
Missing	Mean	937.54	936.64	938.22	932.33	938.15	934.85	936.97	935.69
	Std. Dev.	10.33	10.51	11.99	10.92	10.16	9.97	10.12	9.64
	N	508	506	503	501	501	496	503	491

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	935.75	936.37	934.61	931.48	935.67	934.26	936.28	934.50
	Std. Dev.	10.75	12.42	13.74	11.46	11.46	11.10	11.55	10.84
	N	737	742	732	729	726	725	735	713
Non-Hispanic Pacific Islander	Mean	935.26	936.10	935.29	930.81	935.91	933.71	935.87	933.98
	Std. Dev.	11.40	11.45	13.78	9.17	11.60	9.60	11.16	10.02
	N	46	48	45	48	45	48	45	45
Non-Hispanic Black	Mean	936.05	936.76	935.78	931.81	936.34	934.61	936.61	935.01
	Std. Dev.	11.01	12.57	13.35	10.94	11.66	11.26	11.77	11.06
	N	432	430	427	421	424	418	430	415
Hispanic (Of Any Race)	Mean	937.32	937.63	936.36	931.90	937.30	935.08	937.61	935.57
	Std. Dev.	10.51	11.79	13.12	10.27	11.18	10.28	11.06	10.16
	N	4059	4045	4004	3963	3986	3944	4033	3909
Non-Hispanic American Indian	Mean	940.88	941.12	939.90	934.92	940.67	938.33	941.10	938.69
	Std. Dev.	7.32	11.05	12.54	7.82	9.77	8.43	9.62	8.54
	N	49	50	51	50	49	49	49	48
Non-Hispanic Multi-racial	Mean	937.87	938.41	938.33	931.93	938.41	935.37	938.26	936.07
	Std. Dev.	9.26	9.86	10.80	9.13	9.60	9.07	9.46	8.96
	N	46	46	46	46	46	46	46	46
Non-Hispanic White	Mean	937.01	936.63	936.00	931.06	936.91	934.28	936.83	934.93
	Std. Dev.	10.38	12.07	13.16	10.51	11.23	10.43	11.20	10.27
	N	516	513	513	499	512	497	512	495
Missing	Mean	939.13	939.30	939.44	934.65	939.78	937.33	939.34	937.83
	Std. Dev.	9.58	10.79	11.43	9.95	10.03	9.59	10.07	9.31
	N	415	414	411	401	411	401	413	401

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	936.83	937.19	935.03	933.59	936.20	935.67	937.21	935.74
	Std. Dev.	10.38	11.27	12.68	11.30	10.61	10.57	10.67	10.23
	N	951	953	942	928	937	926	947	913
Non-Hispanic Pacific Islander	Mean	936.35	936.56	935.20	932.67	936.08	934.87	936.47	934.80
	Std. Dev.	11.27	12.10	13.47	10.17	11.52	10.73	11.66	10.78
	N	51	52	51	52	50	52	51	50
Non-Hispanic Black	Mean	937.04	936.75	935.71	932.94	936.62	935.10	936.87	935.39
	Std. Dev.	11.13	11.54	12.72	11.23	11.24	10.74	11.18	10.58
	N	528	529	521	516	518	515	523	506
Hispanic (Of Any Race)	Mean	938.29	938.36	936.07	933.72	937.42	936.37	938.48	936.59
	Std. Dev.	10.21	10.78	12.41	10.71	10.53	10.09	10.33	9.91
	N	4685	4682	4654	4609	4623	4578	4651	4534
Non-Hispanic American Indian	Mean	940.11	941.05	939.04	934.89	939.62	938.26	940.80	938.65
	Std. Dev.	9.37	9.17	10.36	9.89	9.37	8.67	9.05	8.44
	N	55	55	57	56	55	54	55	54
Non-Hispanic Multi-racial	Mean	937.24	937.68	935.89	934.41	937.53	936.94	937.68	936.97
	Std. Dev.	10.41	11.77	11.68	11.21	9.49	9.94	11.11	9.48
	N	37	37	37	37	36	36	37	36
Non-Hispanic White	Mean	937.76	937.91	936.23	933.83	937.24	936.30	938.00	936.49
	Std. Dev.	10.44	10.65	12.19	11.24	10.57	10.11	10.32	9.88
	N	566	564	555	539	554	538	562	536
Missing	Mean	938.82	938.39	937.39	934.30	938.42	936.60	938.68	937.04
	Std. Dev.	9.95	10.25	11.30	10.32	9.92	9.66	9.88	9.42
	N	461	456	453	444	452	443	455	440

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	929.48	11.45	660	929.96	11.65	1673	929.34	11.30	335	929.76	11.56	2668
Reading	929.48	12.77	651	930.32	13.18	1663	930.16	12.73	335	930.09	13.02	2649
Speaking	929.01	14.74	647	930.51	15.07	1646	929.68	14.67	333	930.03	14.95	2626
Writing	925.15	10.74	641	926.21	11.13	1647	925.13	10.92	328	925.81	11.02	2616
Oral	929.63	12.15	643	930.57	12.59	1635	929.83	12.19	332	930.24	12.44	2610
Literacy	927.70	10.76	631	928.58	11.39	1634	927.99	11.01	328	928.30	11.19	2593
Comprehension	929.60	12.02	651	930.34	12.34	1655	930.05	11.93	334	930.12	12.21	2640
Overall	928.08	10.71	628	929.02	11.32	1612	928.43	10.91	325	928.71	11.13	2565

Table 4.7.1 B

Mean Scale Scores by Gender: Grade 2

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	932.35	10.91	774	932.38	11.16	1775	932.51	10.69	282	932.38	11.04	2831
Reading	932.81	12.59	768	933.28	13.29	1760	932.77	13.04	282	933.10	13.07	2810
Speaking	932.02	15.11	761	933.67	14.56	1755	933.65	15.00	278	933.22	14.77	2794
Writing	927.68	10.85	760	928.78	11.66	1742	928.02	11.61	272	928.41	11.44	2774
Oral	932.52	12.14	759	933.33	12.13	1749	933.48	11.99	276	933.12	12.12	2784
Literacy	930.62	10.97	753	931.36	11.62	1722	930.78	11.66	270	931.10	11.45	2745
Comprehension	932.78	11.77	768	933.11	12.27	1759	932.64	11.89	280	932.97	12.09	2807
Overall	931.04	10.85	748	931.76	11.30	1714	931.50	11.17	263	931.54	11.17	2725

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	934.11	11.18	725	934.71	10.96	1735	933.05	11.38	268	934.39	11.07	2728
Reading	932.68	11.22	718	933.74	11.40	1729	931.84	11.47	266	933.28	11.37	2713
Speaking	933.57	14.06	714	934.37	13.77	1714	933.16	14.38	265	934.04	13.91	2693
Writing	928.54	11.13	717	930.11	11.18	1723	928.27	11.37	265	929.51	11.21	2705
Oral	934.12	11.73	712	934.73	11.44	1709	933.38	11.87	261	934.44	11.57	2682
Literacy	930.94	10.56	709	932.27	10.57	1703	930.38	10.71	261	931.73	10.60	2673
Comprehension	933.20	10.87	716	934.08	10.89	1722	932.36	11.08	265	933.67	10.91	2703
Overall	931.73	10.57	702	932.82	10.46	1685	931.22	10.57	258	932.37	10.51	2645

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	936.88	10.24	754	936.58	10.39	1559	935.65	10.28	241	936.58	10.33	2554
Reading	935.77	10.49	751	936.12	10.46	1555	934.33	10.03	241	935.85	10.44	2547
Speaking	936.43	13.05	745	936.35	13.03	1545	934.65	13.87	238	936.21	13.12	2528
Writing	931.62	11.13	742	932.11	11.50	1546	929.80	11.20	234	931.75	11.38	2522
Oral	936.88	10.74	741	936.65	10.84	1538	935.32	11.30	237	936.59	10.86	2516
Literacy	934.09	10.21	737	934.44	10.24	1535	932.39	9.93	234	934.14	10.21	2506
Comprehension	936.16	10.07	748	936.30	10.11	1550	934.76	9.69	240	936.12	10.06	2538
Overall	934.79	9.96	730	934.92	10.03	1523	933.04	9.91	231	934.71	10.01	2484

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	937.65	9.66	741	937.14	10.79	1420	935.47	10.52	230	937.14	10.44	2391
Reading	936.56	10.08	739	936.71	11.05	1412	934.69	10.70	229	936.47	10.74	2380
Speaking	936.08	13.39	735	936.87	13.17	1397	934.58	14.13	228	936.40	13.34	2360
Writing	931.72	11.12	727	933.03	12.04	1390	930.44	11.55	226	932.37	11.74	2343
Oral	937.03	10.61	731	937.26	11.06	1394	935.25	11.23	228	936.99	10.95	2353
Literacy	934.49	9.88	722	935.24	10.82	1385	932.92	10.33	226	934.78	10.51	2333
Comprehension	936.96	9.61	736	936.90	10.63	1410	935.01	10.28	229	936.74	10.30	2375
Overall	935.09	9.69	716	935.73	10.49	1373	933.48	10.24	225	935.31	10.24	2314

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	936.96	10.47	679	936.37	10.61	1300	934.59	10.98	185	936.40	10.61	2164
Reading	937.13	11.42	673	936.85	11.97	1303	934.51	12.40	183	936.74	11.86	2159
Speaking	936.19	12.87	670	935.96	13.18	1287	935.10	13.31	184	935.96	13.09	2141
Writing	931.59	10.55	665	931.28	10.07	1284	930.49	10.35	177	931.31	10.25	2126
Oral	937.04	10.94	669	936.54	11.34	1285	935.43	11.37	182	936.60	11.22	2136
Literacy	934.67	10.22	662	934.32	10.25	1279	932.87	10.50	176	934.31	10.26	2117
Comprehension	937.16	10.77	673	936.74	11.23	1296	934.70	11.60	183	936.70	11.13	2152
Overall	935.17	10.05	661	934.78	10.21	1264	933.41	10.37	176	934.79	10.18	2101

Table 4.7.1 G

Mean Scale Scores by Gender: Grade 7

	Female			Male			Missing			Total		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Listening	938.15	10.23	657	937.57	10.32	1238	934.04	12.16	198	937.42	10.53	2093
Reading	937.93	11.19	659	938.09	11.85	1235	934.79	12.74	192	937.74	11.77	2086
Speaking	937.46	12.34	649	936.92	12.91	1224	933.36	14.41	192	936.76	12.92	2065
Writing	932.49	10.04	640	932.57	10.61	1208	929.27	11.28	184	932.25	10.54	2032
Oral	938.30	10.64	643	937.70	11.02	1219	934.43	12.37	191	937.59	11.08	2053
Literacy	935.65	9.73	636	935.66	10.44	1203	932.43	11.26	182	935.37	10.33	2021
Comprehension	938.08	10.58	655	938.03	11.02	1230	934.76	12.14	192	937.74	11.03	2077
Overall	936.32	9.62	627	936.10	10.21	1194	932.82	11.32	181	935.87	10.18	2002

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	937.84	10.32	666	938.02	10.16	1180	935.23	11.20	197	937.69	10.34	2043
Reading	937.94	11.63	667	938.30	12.06	1182	936.08	12.56	194	937.97	11.98	2043
Speaking	936.31	13.32	664	936.66	13.16	1167	933.73	14.68	192	936.27	13.38	2023
Writing	932.27	10.66	657	932.92	10.46	1155	929.56	10.56	187	932.39	10.57	1999
Oral	937.51	11.22	659	937.78	11.13	1160	935.16	12.08	191	937.44	11.27	2010
Literacy	935.43	10.45	653	935.90	10.68	1152	933.35	10.65	185	935.51	10.62	1990
Comprehension	938.01	10.89	663	938.28	11.17	1177	935.92	11.76	194	937.96	11.15	2034
Overall	935.95	10.21	645	936.31	10.43	1140	933.71	10.73	184	935.95	10.41	1969

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	936.30	10.98	576	937.98	9.93	998	936.20	10.77	206	937.23	10.41	1780
Reading	936.25	11.17	572	938.19	10.62	997	935.84	10.89	205	937.30	10.88	1774
Speaking	934.59	12.97	573	936.57	12.22	981	934.67	12.19	203	935.71	12.50	1757
Writing	932.13	10.95	559	933.49	10.51	984	932.32	11.00	196	932.92	10.72	1739
Oral	935.66	11.18	569	937.54	10.27	976	935.63	10.77	202	936.71	10.67	1747
Literacy	934.49	10.40	556	936.13	9.82	977	934.54	10.12	195	935.42	10.07	1728
Comprehension	936.43	10.75	570	938.26	10.11	991	936.07	10.45	205	937.42	10.40	1766
Overall	934.74	10.31	552	936.46	9.64	966	934.72	9.99	195	935.70	9.93	1713

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	938.26	10.24	503	937.99	10.59	931	935.99	10.84	156	937.88	10.52	1590
Reading	938.80	10.26	504	937.96	11.05	934	936.92	11.10	158	938.12	10.82	1596
Speaking	936.66	11.86	503	936.24	12.23	928	934.16	12.72	156	936.17	12.17	1587
Writing	934.19	10.59	496	933.96	10.82	915	933.13	10.75	152	933.95	10.74	1563
Oral	937.74	10.23	497	937.35	10.73	923	935.29	10.87	153	937.27	10.60	1573
Literacy	936.90	9.74	490	936.32	10.15	912	935.33	10.34	152	936.41	10.05	1554
Comprehension	938.80	9.95	500	938.04	10.66	928	936.84	10.66	154	938.16	10.45	1582
Overall	937.05	9.56	486	936.55	10.00	904	935.24	10.16	148	936.58	9.88	1538

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	938.26	10.45	497	938.85	9.71	885	937.87	10.43	159	938.56	10.03	1541
Reading	938.24	10.72	498	939.10	10.34	883	938.26	10.62	156	938.74	10.50	1537
Speaking	936.16	11.87	493	937.01	11.78	878	935.71	12.34	154	936.61	11.87	1525
Writing	933.61	11.22	488	934.61	10.79	871	933.63	10.04	152	934.18	10.86	1511
Oral	937.50	10.35	488	938.21	10.00	873	937.14	10.49	154	937.87	10.17	1515
Literacy	936.24	10.35	486	937.14	9.88	866	936.22	9.80	152	936.76	10.03	1504
Comprehension	938.39	10.34	491	939.15	9.88	877	938.32	10.21	156	938.82	10.07	1524
Overall	936.48	10.07	477	937.33	9.60	858	936.44	9.75	150	936.97	9.77	1485

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	938.72	10.13	783	938.14	10.50	1295	937.78	9.79	345	938.28	10.29	2423
Reading	938.55	10.89	785	938.11	11.30	1293	937.49	10.95	343	938.17	11.12	2421
Speaking	936.75	12.11	770	935.50	12.93	1290	934.60	13.32	341	935.77	12.74	2401
Writing	934.29	10.82	765	933.56	10.92	1275	933.41	11.26	328	933.78	10.94	2368
Oral	937.98	10.37	769	937.05	10.90	1283	936.48	10.54	338	937.27	10.69	2390
Literacy	936.79	10.18	761	936.08	10.50	1269	935.86	10.48	326	936.28	10.40	2356
Comprehension	938.71	10.49	779	938.24	10.82	1288	937.80	10.23	342	938.33	10.63	2409
Overall	937.07	9.89	750	936.26	10.31	1260	936.07	10.15	323	936.49	10.16	2333

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	929.26	930.58	930.21	926.76	930.06	929.03	930.32	929.18
	Std. Dev.	11.57	13.83	15.15	11.43	12.40	11.65	12.78	11.43
	N	441	442	435	435	431	433	440	425
Non-Hispanic Pacific Islander	Mean	930.00	925.55	929.36	927.42	929.82	926.55	926.82	927.36
	Std. Dev.	13.86	10.74	14.90	10.57	13.72	9.50	11.55	10.22
	N	11	11	11	12	11	11	11	11
Non-Hispanic Black	Mean	928.44	929.51	928.28	926.14	928.69	928.23	929.29	928.09
	Std. Dev.	11.54	13.67	15.11	11.51	12.45	11.67	12.58	11.47
	N	240	236	234	228	233	227	236	227
Hispanic (Of Any Race)	Mean	930.20	930.33	930.38	925.80	930.64	928.38	930.41	928.91
	Std. Dev.	11.61	12.97	14.87	10.79	12.49	11.14	12.21	11.12
	N	1509	1499	1482	1484	1474	1471	1495	1454
Non-Hispanic American Indian	Mean	929.2	930.07	931.00	926.57	929.93	928.21	929.93	928.64
	Std. Dev.	13.27	11.38	14.16	8.16	13.19	9.26	11.83	10.23
	N	15	15	14	14	14	14	15	14
Non-Hispanic Multi-racial	Mean	929.75	926.06	923.06	920.88	926.69	923.81	927.31	924.38
	Std. Dev.	8.80	10.14	15.15	9.15	10.52	8.71	9.53	8.82
	N	16	16	16	16	16	16	16	16
Non-Hispanic White	Mean	928.59	928.53	928.50	923.40	928.89	926.22	928.66	926.79
	Std. Dev.	11.49	12.13	15.16	11.21	12.45	10.79	11.62	10.86
	N	239	235	239	232	238	230	235	230
Missing	Mean	930.58	930.33	931.52	926.53	931.56	929.04	930.67	929.69
	Std. Dev.	11.06	12.12	14.45	10.96	11.89	10.60	11.32	10.46
	N	197	195	195	195	193	191	192	188

Table 4.7.2 B

Mean Scale Scores by Ethnicity: Grade 2

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	932.25	933.39	932.77	929.58	932.83	931.92	933.12	932.07
	Std. Dev.	10.54	12.97	14.80	11.99	11.77	11.49	11.88	11.08
	N	433	430	428	422	426	418	430	415
Non-Hispanic Pacific Islander	Mean	928.30	928.95	931.43	927.79	930.48	929.58	928.82	929.79
	Std. Dev.	13.18	15.33	14.90	11.73	13.26	13.14	14.42	13.02
	N	23	22	21	19	21	19	22	19
Non-Hispanic Black	Mean	931.14	931.93	933.65	928.81	932.74	930.65	931.81	931.04
	Std. Dev.	11.00	13.81	14.66	12.27	12.18	12.25	12.58	11.79
	N	265	262	260	257	260	255	262	254
Hispanic (Of Any Race)	Mean	932.64	933.26	933.24	928.18	933.26	931.08	933.16	931.57
	Std. Dev.	11.15	12.95	14.78	11.04	12.25	11.26	12.04	11.07
	N	1598	1588	1578	1571	1572	1552	1585	1539
Non-Hispanic American Indian	Mean	935.45	933.11	936.42	928.84	936	930.67	933.74	932.11
	Std. Dev.	9.67	11.45	14.51	10.63	10.97	10.01	10.41	9.87
	N	20	19	19	19	19	18	19	18
Non-Hispanic Multi-racial	Mean	931.19	930.94	923.94	925.25	927.81	928.31	931.19	927.94
	Std. Dev.	11.95	13.00	15.47	9.38	11.88	10.20	12.57	10.00
	N	16	16	16	16	16	16	16	16
Non-Hispanic White	Mean	931.78	932.16	932.56	926.79	932.46	929.72	932.15	930.35
	Std. Dev.	10.96	12.98	14.68	11.59	11.96	11.18	11.96	10.97
	N	275	273	273	273	271	271	273	269
Missing	Mean	933.36	934.69	934.91	929.64	934.47	932.45	934.36	932.85
	Std. Dev.	11.06	13.18	14.61	12.05	11.89	12.06	12.23	11.51
	N	201	200	199	197	199	196	200	195

Table 4.7.2 C

Mean Scale Scores by Ethnicity: Grade 3

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	933.72	933.35	933.87	931.13	934.01	932.59	933.50	932.75
	Std. Dev.	10.72	11.74	14.00	11.80	11.43	11.19	11.10	10.88
	N	432	430	430	427	428	425	429	423
Non-Hispanic Pacific Islander	Mean	933.41	931	934.59	924.23	934.23	927.77	931.77	929.45
	Std. Dev.	10.64	9.93	12.83	10.43	10.69	9.03	9.92	9.08
	N	22	22	22	22	22	22	22	22
Non-Hispanic Black	Mean	933.06	932	933.87	929.79	933.66	931.21	932.39	931.82
	Std. Dev.	11.15	12.07	14.04	11.53	11.79	11.20	11.38	10.97
	N	218	217	212	211	212	208	215	203
Hispanic (Of Any Race)	Mean	934.60	933.43	933.78	929.10	934.42	931.60	933.85	932.28
	Std. Dev.	11.13	11.14	14.05	10.98	11.63	10.34	10.77	10.36
	N	1602	1593	1581	1596	1575	1575	1588	1561
Non-Hispanic American Indian	Mean	939.32	937.14	936.81	930.36	938.19	934	937.86	935.10
	Std. Dev.	8.93	10.51	13.46	11.99	10.65	9.55	9.77	9.63
	N	22	22	21	22	21	22	22	21
Non-Hispanic Multi-racial	Mean	932.61	927.18	929.55	924.13	931.73	925.82	928.55	927.48
	Std. Dev.	12.85	14.05	14.22	11.79	12.33	12.30	13.22	11.91
	N	23	22	22	23	22	22	22	21
Non-Hispanic White	Mean	933.25	932.16	933.62	928.87	933.62	930.91	932.62	931.67
	Std. Dev.	11.67	11.88	14.10	11.46	12.07	11.08	11.40	10.97
	N	231	229	227	227	226	224	228	220
Missing	Mean	936.90	935.29	937.62	931.15	937.51	933.47	935.82	934.45
	Std. Dev.	9.90	10.48	11.68	10.32	9.98	9.82	9.97	9.56
	N	178	178	178	177	176	175	177	174

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	935.07	935.71	935.55	933.15	935.49	934.77	935.50	934.73
	Std. Dev.	10.04	10.72	13.25	11.52	10.77	10.37	10.19	10.11
	N	332	332	330	328	329	328	331	326
Non-Hispanic Pacific Islander	Mean	934.22	931.83	935.33	927.89	934.94	930.11	932.67	931.33
	Std. Dev.	13.02	11.62	12.59	11.08	12.23	9.46	11.76	9.77
	N	18	18	18	18	18	18	18	18
Non-Hispanic Black	Mean	935.39	934.93	936.72	930.96	936.34	933.21	935.17	934.12
	Std. Dev.	11.16	10.88	13.23	11.94	11.25	10.68	10.54	10.40
	N	180	178	177	179	175	175	178	173
Hispanic (Of Any Race)	Mean	937.12	936.15	936.30	931.62	936.92	934.24	936.49	934.87
	Std. Dev.	10.08	10.18	13.04	11.15	10.70	10.00	9.82	9.82
	N	1547	1544	1529	1525	1520	1517	1537	1501
Non-Hispanic American Indian	Mean	940.47	940.79	938.32	934.95	939.53	938.32	940.68	938.37
	Std. Dev.	7.55	6.21	10.67	11.95	8.36	8.27	6.20	7.94
	N	19	19	19	20	19	19	19	19
Non-Hispanic Multi-racial	Mean	937.38	935.81	939.25	935.81	938.50	936.13	936.31	936.50
	Std. Dev.	8.27	8.03	9.77	8.32	8.11	7.48	7.85	7.54
	N	16	16	16	16	16	16	16	16
Non-Hispanic White	Mean	935.45	934.31	934.20	930.48	934.95	932.70	934.71	933.22
	Std. Dev.	11.04	11.07	14.29	11.98	11.76	10.90	10.66	10.73
	N	262	261	261	255	261	254	261	254
Missing	Mean	937.27	936.60	938.71	932.59	938.19	934.97	936.87	935.81
	Std. Dev.	10.81	10.87	11.90	11.53	10.55	10.45	10.55	10.19
	N	180	179	178	181	178	179	178	177

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	935.64	935.65	935.18	932.81	935.62	934.57	935.68	934.64
	Std. Dev.	10.53	11.17	13.77	12.58	11.06	11.11	10.64	10.74
	N	306	306	302	302	302	302	305	300
Non-Hispanic Pacific Islander	Mean	936.56	937.12	939.56	937.29	938.29	937.47	937.12	937.59
	Std. Dev.	10.53	11.92	11.63	11.48	10.76	11.60	11.26	11.05
	N	18	17	18	17	17	17	17	17
Non-Hispanic Black	Mean	934.85	934.18	936.43	931.65	935.83	933.13	934.45	933.79
	Std. Dev.	10.83	11.57	13.25	12.12	11.19	10.86	11.00	10.56
	N	168	167	168	168	167	166	167	166
Hispanic (Of Any Race)	Mean	937.74	936.93	936.61	932.52	937.39	935.08	937.25	935.64
	Std. Dev.	10.29	10.53	13.17	11.56	10.82	10.37	10.09	10.11
	N	1505	1497	1486	1476	1482	1470	1494	1459
Non-Hispanic American Indian	Mean	939.85	939.3	939.05	935.95	939.68	938.95	939.4	940.06
	Std. Dev.	10.39	10.46	13.41	10.15	11.85	8.61	10.26	8.41
	N	20	20	20	20	19	19	20	18
Non-Hispanic Multi-racial	Mean	933.1	934	933.5	928.33	933.4	931.22	933.7	931.22
	Std. Dev.	12.40	13.42	15.81	10.70	13.91	11.73	12.96	12.28
	N	10	10	10	9	10	9	10	9
Non-Hispanic White	Mean	935.79	934.76	934.89	929.93	935.56	932.75	935.10	933.55
	Std. Dev.	10.53	10.83	14.39	11.97	11.64	10.71	10.49	10.59
	N	214	214	209	208	209	208	214	205
Missing	Mean	938.63	938.28	938.34	933.46	938.88	936.41	938.48	937.08
	Std. Dev.	10.23	9.93	12.52	10.79	9.90	9.36	9.64	8.88
	N	150	149	147	143	147	142	148	140

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	935.22	935.69	934.13	931.59	935.05	933.86	935.58	933.92
	Std. Dev.	10.51	12.39	13.46	11.14	11.19	10.94	11.46	10.72
	N	267	268	267	266	266	264	267	262
Non-Hispanic Pacific Islander	Mean	932.06	932.28	933.63	927.72	933.19	930.28	931.31	930.25
	Std. Dev.	13.02	11.72	14.23	8.77	13.67	9.83	11.90	10.76
	N	16	18	16	18	16	18	16	16
Non-Hispanic Black	Mean	935.41	935.64	935.14	930.91	935.71	933.59	935.63	933.99
	Std. Dev.	11.14	12.62	13.23	10.87	11.71	11.24	11.90	11.09
	N	148	148	147	146	145	144	148	144
Hispanic (Of Any Race)	Mean	936.64	937.07	936.16	931.16	936.82	934.38	937.01	934.92
	Std. Dev.	10.55	11.64	13.05	9.97	11.15	9.99	10.96	9.96
	N	1348	1343	1329	1325	1327	1320	1339	1308
Non-Hispanic American Indian	Mean	939	940	938.45	932.09	939.09	936.27	939.73	936.91
	Std. Dev.	9.56	11.94	12.16	7.25	10.83	8.91	11.05	9.49
	N	11	11	11	11	11	11	11	11
Non-Hispanic Multi-racial	Mean	938.90	940	940.05	933.52	939.81	937	939.76	937.62
	Std. Dev.	8.58	8.83	10.95	7.87	9.28	8.04	8.50	8.08
	N	21	21	21	21	21	21	21	21
Non-Hispanic White	Mean	936.02	935.54	935.58	930.66	936.26	933.56	935.77	934.27
	Std. Dev.	10.77	11.99	13.11	10.66	11.26	10.44	11.15	10.16
	N	182	180	180	174	180	174	180	174
Missing	Mean	937.61	937.89	937.92	933.18	938.19	935.83	937.84	936.32
	Std. Dev.	10.43	11.98	12.66	10.38	11.07	10.48	11.22	10.22
	N	171	170	170	165	170	165	170	165

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	935.64	936.79	935.72	931.35	936.19	934.59	936.58	934.90
	Std. Dev.	10.96	11.97	13.43	11.18	11.58	10.56	11.31	10.51
	N	247	248	242	239	240	238	245	234
Non-Hispanic Pacific Islander	Mean	935.46	935.54	934.25	930.85	936.25	933.46	936.25	934
	Std. Dev.	10.94	11.03	13.53	11.97	9.58	10.47	10.00	10.24
	N	13	13	12	13	12	13	12	12
Non-Hispanic Black	Mean	935.87	937.19	934.99	932.02	935.79	934.99	936.90	935.13
	Std. Dev.	11.34	13.17	13.61	11.47	11.90	11.79	12.24	11.49
	N	146	145	146	143	145	142	145	141
Hispanic (Of Any Race)	Mean	937.5	937.69	936.69	932.13	937.60	935.25	937.72	935.81
	Std. Dev.	10.61	11.78	12.97	10.38	11.12	10.29	11.07	10.14
	N	1350	1344	1330	1311	1322	1304	1340	1291
Non-Hispanic American Indian	Mean	940.82	939.88	939.28	932.88	940.29	936.56	940.18	937.25
	Std. Dev.	7.44	11.56	13.67	5.83	10.32	7.63	9.69	8.14
	N	17	17	18	17	17	16	17	16
Non-Hispanic Multi-racial	Mean	941.33	941.56	942.22	932.67	942.11	937.33	941.33	938.56
	Std. Dev.	4.87	5.00	4.52	7.58	4.14	5.61	4.66	5.03
	N	9	9	9	9	9	9	9	9
Non-Hispanic White	Mean	937.92	937.25	936.43	931.57	937.58	934.81	937.50	935.44
	Std. Dev.	9.68	11.57	12.68	10.78	10.65	10.32	10.76	10.11
	N	174	173	173	169	173	168	173	168
Missing	Mean	940.32	940.75	941.18	936.17	941.31	938.92	940.84	939.43
	Std. Dev.	8.84	9.83	10.19	9.35	8.92	8.56	9.00	8.29
	N	137	137	135	131	135	131	136	131

Table 4.7.2 H

Mean Scale Scores by Ethnicity: Grade 8

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	936.49	936.69	933.96	931.48	935.85	934.40	936.78	934.78
	Std. Dev.	10.79	12.95	14.38	12.15	11.68	11.88	11.92	11.35
	N	223	226	223	224	220	223	223	217
Non-Hispanic Pacific Islander	Mean	938.12	940.59	937.59	934.06	938.24	937.53	939.88	937.47
	Std. Dev.	9.91	10.49	14.05	5.99	10.89	7.56	10.14	8.29
	N	17	17	17	17	17	17	17	17
Non-Hispanic Black	Mean	936.93	937.50	937.33	932.57	937.63	935.31	937.36	936
	Std. Dev.	10.53	11.86	13.16	10.42	11.30	10.69	11.12	10.52
	N	138	137	134	132	134	132	137	130
Hispanic (Of Any Race)	Mean	937.81	938.12	936.23	932.41	937.47	935.59	938.10	936.00
	Std. Dev.	10.34	11.92	13.35	10.43	11.28	10.53	11.12	10.35
	N	1361	1358	1345	1327	1337	1320	1354	1310
Non-Hispanic American Indian	Mean	941.90	942.64	941.14	937.91	941.81	940.64	942.57	940.71
	Std. Dev.	5.96	10.53	12.23	8.66	9.08	8.53	9.06	8.31
	N	21	22	22	22	21	22	21	21
Non-Hispanic Multi-racial	Mean	934.56	934.56	933.88	929.44	934.5	932.13	934.56	932.63
	Std. Dev.	11.22	12.22	12.03	11.25	11.18	11.24	11.70	10.98
	N	16	16	16	16	16	16	16	16
Non-Hispanic White	Mean	937.16	937.19	936.03	930.96	936.92	934.52	937.29	935.10
	Std. Dev.	10.62	12.66	13.77	10.07	11.82	10.55	11.71	10.61
	N	160	160	160	156	159	155	159	153
Missing	Mean	940.01	939.68	939.68	935.08	940.37	937.69	939.83	938.2
	Std. Dev.	8.80	9.73	10.59	9.75	9.29	9.05	9.13	8.73
	N	107	107	106	105	106	105	107	105

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	935.82	936.30	934.48	932.66	935.37	934.72	936.36	934.81
	Std. Dev.	10.14	11.27	12.59	10.88	10.43	10.08	10.43	9.86
	N	219	218	220	219	218	218	217	217
Non-Hispanic Pacific Islander	Mean	938.36	938.64	936.75	933.91	939	936.45	938.64	936.91
	Std. Dev.	6.71	8.25	13.40	4.64	7.75	5.92	7.59	5.96
	N	11	11	12	11	11	11	11	11
Non-Hispanic Black	Mean	937.15	935.98	936.26	933.60	936.98	935.18	936.45	935.73
	Std. Dev.	11.41	12.15	12.50	11.05	11.29	10.99	11.72	10.72
	N	134	133	130	129	129	128	133	127
Hispanic (Of Any Race)	Mean	937.41	937.49	935.73	932.73	936.80	935.45	937.60	935.75
	Std. Dev.	10.28	10.74	12.51	10.58	10.63	9.94	10.26	9.84
	N	1155	1153	1138	1128	1132	1119	1146	1106
Non-Hispanic American Indian	Mean	937.75	940.33	939.58	936.45	938.75	939.45	939.58	939.91
	Std. Dev.	10.82	7.94	10.19	9.22	10.33	6.62	8.64	5.91
	N	12	12	12	11	12	11	12	11
Non-Hispanic Multi-racial	Mean	933.56	934.11	928.11	931.33	931	932.89	934	932.22
	Std. Dev.	13.87	14.57	15.34	13.59	13.34	13.49	14.31	13.18
	N	9	9	9	9	9	9	9	9
Non-Hispanic White	Mean	937.86	938.47	936.51	934.64	937.29	936.81	938.36	936.83
	Std. Dev.	10.45	10.46	12.17	11.43	10.61	10.45	10.29	10.20
	N	121	121	119	116	119	116	121	116
Missing	Mean	937.64	937.36	936.44	932.41	937.41	935.05	937.64	935.53
	Std. Dev.	10.81	10.49	12.32	10.99	10.83	10.16	10.26	10.09
	N	119	117	117	116	117	116	117	116

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	936.98	937.41	935.46	933.67	936.45	935.81	937.40	935.99
	Std. Dev.	9.85	10.67	11.94	10.64	9.97	9.88	10.10	9.51
	N	204	205	203	200	202	199	204	197
Non-Hispanic Pacific Islander	Mean	934.13	934.63	934.38	931.31	934.38	933.19	934.63	933.25
	Std. Dev.	14.70	15.41	14.81	13.04	14.71	13.99	15.16	13.97
	N	16	16	16	16	16	16	16	16
Non-Hispanic Black	Mean	935.56	935.92	934.74	932.19	935.25	934.47	935.73	934.52
	Std. Dev.	11.96	12.18	13.40	11.60	11.98	11.08	11.96	11.01
	N	113	115	114	112	113	112	113	110
Hispanic (Of Any Race)	Mean	938.47	938.68	936.29	934.18	937.64	936.79	938.75	936.94
	Std. Dev.	10.12	10.42	12.15	10.59	10.41	9.83	10.04	9.69
	N	1023	1026	1023	1011	1012	1004	1016	993
Non-Hispanic American Indian	Mean	941.5	940	939.08	933.62	940.42	937	940.58	937.83
	Std. Dev.	8.08	10.83	10.37	9.00	9.11	9.22	10.05	9.23
	N	12	12	13	13	12	12	12	12
Non-Hispanic Multi-racial	Mean	942.67	945.33	944	943	943.33	944.33	944.67	944
	Std. Dev.	4.51	2.31	1.73	11.79	3.06	6.43	3.06	5.20
	N	3	3	3	3	3	3	3	3
Non-Hispanic White	Mean	935.67	936.35	935.48	933.07	935.81	935.39	936.22	935.55
	Std. Dev.	12.39	12.55	12.79	11.91	11.91	11.08	12.23	10.98
	N	129	130	127	122	127	122	129	121
Missing	Mean	939.28	938.89	938.97	935.65	939.57	937.52	939.20	938.01
	Std. Dev.	10.23	9.71	9.81	9.32	9.08	9.03	9.46	8.65
	N	90	89	88	86	88	86	89	86

Table 4.7.2 K

Mean Scale Scores by Ethnicity: Grade 11

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Non-Hispanic Asian	Mean	937.76	938.49	935.92	934.27	937.24	936.61	938.37	936.62
	Std. Dev.	10.51	10.82	12.16	10.95	10.36	10.28	10.39	10.01
	N	204	204	199	198	198	198	202	194
Non-Hispanic Pacific Islander	Mean	941.9	941.3	939.44	935.2	940.44	938.5	941.6	938.89
	Std. Dev.	4.98	6.46	11.44	7.22	7.67	6.88	5.83	6.99
	N	10	10	9	10	9	10	10	9
Non-Hispanic Black	Mean	938.02	937.87	936.95	933.96	937.75	936	937.84	936.12
	Std. Dev.	10.51	10.87	11.76	10.51	10.42	10.05	10.48	9.96
	N	111	114	111	112	110	112	110	108
Hispanic (Of Any Race)	Mean	938.49	938.71	936.38	934.00	937.71	936.70	938.80	936.94
	Std. Dev.	10.17	10.70	12.10	10.85	10.35	10.12	10.26	9.87
	N	972	969	967	955	960	949	962	940
Non-Hispanic American Indian	Mean	939.29	939.65	937.71	933.29	938.53	936.71	939.53	937
	Std. Dev.	11.58	11.37	11.95	12.42	11.57	11.39	11.29	11.08
	N	17	17	17	17	17	17	17	17
Non-Hispanic Multi-racial	Mean	937.25	935.92	939.45	936.82	939.82	937.82	936.42	938.09
	Std. Dev.	10.22	12.38	5.89	10.20	5.49	8.80	11.42	7.40
	N	12	12	11	11	11	11	12	11
Non-Hispanic White	Mean	939.93	939.50	937.58	934.28	939.06	937.16	939.83	937.54
	Std. Dev.	8.01	8.38	10.92	11.27	8.71	9.25	7.83	8.74
	N	129	127	127	126	126	125	127	125
Missing	Mean	939.52	939.61	938.04	935.98	939	937.93	939.71	938.06
	Std. Dev.	9.41	9.79	10.72	10.97	9.76	9.90	9.52	9.65
	N	86	84	84	82	84	82	84	81

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	936.84	936.84	934.58	933.76	935.95	935.64	936.95	935.69
	Std. Dev.	10.76	11.86	13.50	12.21	11.26	11.48	11.30	11.02
	N	324	326	320	311	319	311	324	305
Non-Hispanic Pacific Islander	Mean	933.36	933.93	932.07	931.53	932.93	933.07	933.21	932.29
	Std. Dev.	12.01	13.11	13.66	11.69	11.54	11.67	12.25	11.41
	N	14	15	14	15	14	15	14	14
Non-Hispanic Black	Mean	937.3	937.16	935.11	932.23	936.52	934.86	937.35	935.22
	Std. Dev.	10.70	11.06	13.04	11.59	11.20	10.81	10.65	10.62
	N	170	167	166	163	166	163	167	161
Hispanic (Of Any Race)	Mean	938.69	938.59	935.98	933.95	937.56	936.56	938.76	936.76
	Std. Dev.	10.22	11.08	12.71	10.76	10.63	10.32	10.59	10.11
	N	1535	1534	1526	1515	1519	1506	1527	1495
Non-Hispanic American Indian	Mean	941.93	944.29	940.07	936.67	941	940.29	943.57	940.36
	Std. Dev.	5.85	4.76	9.44	8.22	5.91	5.57	4.77	5.64
	N	14	14	15	15	14	14	14	14
Non-Hispanic Multi-racial	Mean	938.54	940	936.36	932.64	938.77	937.31	939.77	937.69
	Std. Dev.	8.86	9.96	11.74	10.06	8.38	8.27	9.18	8.08
	N	13	13	14	14	13	13	13	13
Non-Hispanic White	Mean	937.65	937.55	935.61	933.49	936.93	935.97	937.76	936.18
	Std. Dev.	10.24	10.60	12.63	10.63	10.62	9.77	10.25	9.63
	N	187	186	182	175	182	175	185	174
Missing	Mean	939.05	938.25	936.89	934.07	938.24	936.55	938.61	937.10
	Std. Dev.	9.42	10.60	11.57	9.82	9.75	9.41	10.00	9.14
	N	166	166	164	160	163	159	165	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	Speaking	Writing
Listening	Pearson Correlation	1	0.846	0.743	0.688
	N	5499	5447	5394	5341
Reading	Pearson Correlation		1	0.731	0.726
	N		5459	5385	5338
Speaking	Pearson Correlation			1	0.723
	N			5420	5325
Writing	Pearson Correlation				1
	N				5390

**. 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	Speaking	Writing
Listening	Pearson Correlation	1	0.867	0.746	0.677
	N	7673	7616	7551	7510
Reading	Pearson Correlation		1	0.761	0.756
	N		7640	7543	7512
Speaking	Pearson Correlation			1	0.7028
	N			7581	7487
Writing	Pearson Correlation				1
	N				7570

**. 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	Speaking	Writing
Listening	Pearson Correlation	1	0.867	0.767	0.698
	N	6300	6263	6199	6112
Reading	Pearson Correlation		1	0.775	0.743
	N		6288	6201	6128
Speaking	Pearson Correlation			1	0.721
	N			6229	6105
Writing	Pearson Correlation				1
	N				6157

**. 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	Speaking	Writing
Listening	Pearson Correlation	1	0.879	0.760	0.738
	N	7334	7281	7225	7108
Reading	Pearson Correlation		1	0.775	0.771
	N		7328	7227	7142
Speaking	Pearson Correlation			1	0.735
	N			7270	7114
Writing	Pearson Correlation				1
	N				7181

**. 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 | | | | | | | | | | Total |
| | A1 | | A2 | | A3 | | P1 | | P2 | | |
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1401	25.48	764	13.89	1052	19.13	1050	19.094	1232	22.40	5499
3-5	1181	15.39	811	10.57	1113	14.51	1710	22.286	2858	37.25	7673
6-8	795	12.62	632	10.03	800	12.70	891	14.143	3182	50.51	6300
9-12	841	11.47	559	7.62	970	13.23	1506	20.534	3458	47.15	7334
Total	4218	15.74	2766	10.32	3935	14.68	5157	19.238	10730	40.03	26806

Table 4.9.1 B

Proficiency Level by Cluster: Reading

| | Reading Proficiency Range | | | | | | | | | | Total |
| | A1 | | A2 | | A3 | | P1 | | P2 | | |
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1476	27.04	828	15.17	965	17.68	1076	19.71	1076	19.71	5459
3-5	1185	15.51	1110	14.53	1082	14.16	1741	22.79	1741	22.79	7640
6-8	904	14.38	533	8.48	683	10.86	1109	17.64	1109	17.64	6288
9-12	860	11.74	632	8.62	775	10.58	1334	18.20	1334	18.20	7328
Total	4425	16.56	3103	11.62	3505	13.12	5260	19.69	10422	39.01	26715

Table 4.9.1 C

Proficiency Level by Cluster: Speaking

| | Speaking Proficiency Range | | | | | | | | | | Total |
| | A1 | | A2 | | A3 | | P1 | | P2 | | |
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1820	33.58	237	4.37	670	12.36	1625	29.98	1068	19.70	5420
3-5	1693	22.33	393	5.18	648	8.55	2338	30.84	2509	33.10	7581
6-8	1254	20.13	231	3.71	686	11.01	1718	27.58	2340	37.57	6229
9-12	1391	19.13	231	3.18	782	10.76	1821	25.05	3045	41.88	7270
Total	6158	23.24	1092	4.12	2786	10.51	7502	28.31	8962	33.82	26500

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	1880	34.88	1135	21.06	1440	26.72	805	14.94	102	1.89	28	0.52	5390
3-5	1725	22.79	1699	22.44	1906	25.18	1391	18.38	691	9.13	158	2.09	7570
6-8	1152	18.71	1678	27.25	1203	19.54	1835	29.80	107	1.74	182	2.96	6157
9-12	1138	15.85	1643	22.88	1258	17.52	2676	37.27	159	2.21	307	4.28	7181
Total	5895	22.42	6155	23.40	5807	22.08	6707	25.50	1059	4.03	675	2.57	26298

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	1646	30.52	493	9.14	780	14.46	1363	25.27	1112	20.62	5394
3-5	1461	19.35	553	7.32	994	13.16	1986	26.30	2557	33.86	7551
6-8	1050	16.94	429	6.92	758	12.23	1286	20.75	2676	43.17	6199
9-12	1116	15.45	417	5.77	889	12.30	1947	26.95	2856	39.53	7225
Total	5273	20.00	1892	7.18	3421	12.97	6582	24.96	9201	34.89	26369

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	1601	29.99	1052	19.71	1341	25.12	843	15.79	501	9.39	5338
3-5	1375	18.30	1425	18.97	1661	22.11	1764	23.48	1287	17.13	7512
6-8	920	15.01	834	13.61	1403	22.89	1869	30.50	1102	17.98	6128
9-12	908	12.71	877	12.28	1450	20.30	2218	31.06	1689	23.65	7142
Total	4804	18.39	4188	16.03	5855	22.42	6694	25.63	4579	17.53	26120

able 4.9.1 G

Proficiency Level by Cluster: Comprehension

	Comprehension Proficiency Range										
Cluster	A1		A2		A3		P1		P2		Total
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1-2	1433	26.31	793	14.56	949	17.42	1280	23.50	992	18.21	5447
3-5	1146	15.05	1026	13.47	1070	14.05	1596	20.96	2778	36.48	7616
6-8	848	13.54	561	8.96	653	10.43	1269	20.26	2932	46.81	6263
9-12	806	11.07	603	8.28	802	11.01	1453	19.96	3617	49.68	7281
Total	4233	15.91	2983	11.21	3474	13.06	5598	21.04	10319	38.78	26607

Table 4.9.1 H

Proficiency Level by Cluster: Overall

	Overall Proficiency Range										
Cluster	A1		A2		A3		P1		P2		Total
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1-2	1524	28.81	800	15.12	1368	25.86	1035	19.57	563	10.64	5290
3-5	1298	17.44	981	13.18	1740	23.38	1958	26.31	1466	19.70	7443
6-8	903	14.87	611	10.06	1305	21.49	1839	30.29	1414	23.29	6072
9-12	878	12.42	664	9.39	1425	20.16	2095	29.64	2007	28.39	7069
Total	4603	17.79	3056	11.81	5838	22.56	6927	26.77	5450	21.06	25874

4.9.2 Proficiency Level by Grade

Table 4.9.2 A

Proficiency Level by Grade: Listening

| | Listening Proficiency Range | | | | | | | | | | |
| | A1 | | A2 | | A3 | | P1 | | P2 | | |
Grade	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1	794	29.76	406	15.22	510	19.12	478	17.92	480	17.99	2668
2	607	21.44	358	12.65	542	19.15	572	20.20	752	26.56	2831
3	524	19.21	332	12.17	427	15.65	597	21.88	848	31.09	2728
4	350	13.70	265	10.38	367	14.37	579	22.67	993	38.88	2554
5	307	12.84	214	8.95	319	13.34	534	22.33	1017	42.53	2391
6	298	13.77	240	11.09	298	13.77	332	15.34	996	46.03	2164
7	262	12.52	195	9.32	255	12.18	281	13.43	1100	52.56	2093
8	235	11.50	197	9.64	247	12.09	278	13.61	1086	53.16	2043
9	227	12.75	150	8.43	269	15.11	371	20.84	763	42.87	1780
10	189	11.89	118	7.42	218	13.71	314	19.75	751	47.23	1590
11	155	10.06	106	6.88	196	12.72	325	21.09	759	49.25	1541
12	270	11.14	185	7.64	287	11.84	496	20.47	1185	48.91	2423
Total	4218	15.74	2766	10.32	3935	14.68	5157	19.24	10730	40.03	26806

Table 4.9.2 B

Proficiency Level by Grade: Reading

| | Reading Proficiency Range | | | | | | | | | | |
| | A1 | | A2 | | A3 | | P1 | | P2 | | |
Grade	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1	820	30.96	436	16.46	480	18.12	482	18.20	431	16.27	2649
2	656	23.35	392	13.95	485	17.26	594	21.14	683	24.31	2810
3	538	19.83	454	16.73	383	14.12	640	23.59	698	25.73	2713
4	336	13.19	360	14.13	365	14.33	603	23.67	883	34.67	2547
5	311	13.07	296	12.44	334	14.03	498	20.92	941	39.54	2380
6	327	15.15	194	8.99	246	11.39	422	19.55	970	44.93	2159
7	288	13.81	168	8.05	243	11.65	367	17.59	1020	48.90	2086
8	289	14.15	171	8.37	194	9.50	320	15.66	1069	52.33	2043
9	223	12.57	177	9.98	203	11.44	342	19.28	829	46.73	1774
10	177	11.09	140	8.77	170	10.65	291	18.23	818	51.25	1596
11	162	10.54	112	7.29	153	9.95	296	19.26	814	52.96	1537
12	298	12.31	203	8.38	249	10.29	405	16.73	1266	52.29	2421
Total	4425	16.56	3103	11.62	3505	13.12	5260	19.69	10422	39.01	26715

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	989	37.66	124	4.72	356	13.56	739	28.14	418	15.92	2626
2	831	29.74	113	4.04	314	11.24	886	31.71	650	23.26	2794
3	698	25.92	146	5.42	264	9.80	834	30.97	751	27.89	2693
4	514	20.33	129	5.10	202	7.99	821	32.48	862	34.10	2528
5	481	20.38	118	5.00	182	7.71	683	28.94	896	37.97	2360
6	443	20.69	84	3.92	256	11.96	611	28.54	747	34.89	2141
7	380	18.40	91	4.41	225	10.90	563	27.26	806	39.03	2065
8	431	21.30	56	2.77	205	10.13	544	26.89	787	38.90	2023
9	348	19.81	49	2.79	208	11.84	446	25.38	706	40.18	1757
10	301	18.97	55	3.47	164	10.33	412	25.96	655	41.27	1587
11	248	16.26	63	4.13	167	10.95	405	26.56	642	42.10	1525
12	494	20.57	64	2.67	243	10.12	558	23.24	1042	43.40	2401
Total	6158	23.24	1092	4.12	2786	10.51	7502	28.31	8962	33.82	26500

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	1024	39.14	568	21.71	671	25.65	310	11.85	33	1.26	10	0.38	2616
2	856	30.86	567	20.44	769	27.72	495	17.84	69	2.49	18	0.65	2774
3	705	26.06	676	24.99	669	24.73	455	16.82	166	6.14	34	1.26	2705
4	528	20.94	552	21.89	669	26.53	460	18.24	254	10.07	59	2.34	2522
5	492	21.00	471	20.10	568	24.24	476	20.32	271	11.57	65	2.77	2343
6	408	19.19	646	30.39	417	19.61	568	26.72	32	1.51	55	2.59	2126
7	381	18.75	520	25.59	391	19.24	639	31.45	40	1.97	61	3.00	2032
8	363	18.16	512	25.61	395	19.76	628	31.42	35	1.75	66	3.30	1999
9	301	17.31	433	24.90	331	19.03	572	32.89	37	2.13	65	3.74	1739
10	236	15.10	342	21.88	263	16.83	632	40.44	28	1.79	62	3.97	1563
11	225	14.89	336	22.24	262	17.34	570	37.72	45	2.98	73	4.83	1511
12	376	15.88	532	22.47	402	16.98	902	38.09	49	2.07	107	4.52	2368
Total	5895	22.42	6155	23.40	5807	22.08	6707	25.50	1059	4.03	675	2.57	26298

Table 4.9.2 E

Proficiency Level by Grade: Oral

| | Oral Proficiency Range | | | | | | | | | | |
| | A1 | | A2 | | A3 | | P1 | | P2 | | |
Grade	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1	903	34.60	252	9.66	418	16.02	612	23.45	425	16.28	2610
2	743	26.69	241	8.66	362	13.00	751	26.98	687	24.68	2784
3	614	22.89	219	8.17	391	14.58	702	26.17	756	28.19	2682
4	445	17.69	173	6.88	324	12.88	676	26.87	898	35.69	2516
5	402	17.08	161	6.84	279	11.86	608	25.84	903	38.38	2353
6	386	18.07	158	7.40	278	13.01	470	22.00	844	39.51	2136
7	327	15.93	131	6.38	256	12.47	417	20.31	922	44.91	2053
8	337	16.77	140	6.97	224	11.14	399	19.85	910	45.27	2010
9	280	16.03	115	6.58	247	14.14	473	27.07	632	36.18	1747
10	242	15.38	90	5.72	196	12.46	411	26.13	634	40.31	1573
11	201	13.27	78	5.15	194	12.81	421	27.79	621	40.99	1515
12	393	16.44	134	5.61	252	10.54	642	26.86	969	40.54	2390
Total	5273	20.00	1892	7.18	3421	12.97	6582	24.96	9201	34.89	26369

Table 4.9.2 F

Proficiency Level by Grade: Literacy

| | Literacy Proficiency Range | | | | | | | | | | |
| | A1 | | A2 | | A3 | | P1 | | P2 | | |
Grade	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1	888	34.25	548	21.13	643	24.80	336	12.96	178	6.86	2593
2	713	25.97	504	18.36	698	25.43	507	18.47	323	11.77	2745
3	602	22.52	558	20.88	636	23.79	566	21.17	311	11.63	2673
4	402	16.04	472	18.83	559	22.31	617	24.62	456	18.20	2506
5	371	15.90	395	16.93	466	19.97	581	24.90	520	22.29	2333
6	342	16.15	302	14.27	535	25.27	645	30.47	293	13.84	2117
7	284	14.05	281	13.90	462	22.86	587	29.05	407	20.14	2021
8	294	14.77	251	12.61	406	20.40	637	32.01	402	20.20	1990
9	229	13.25	255	14.76	367	21.24	556	32.18	321	18.58	1728
10	190	12.23	190	12.23	306	19.69	501	32.24	367	23.62	1554
11	172	11.44	163	10.84	323	21.48	454	30.19	392	26.06	1504
12	317	13.46	269	11.42	454	19.27	707	30.01	609	25.85	2356
Total	4804	18.39	4188	16.03	5855	22.42	6694	25.63	4579	17.53	26120

Table 4.9.2 G

Proficiency Level by Grade: Comprehension

| | Comprehension Proficiency Range | | | | | | | | | | |
| | A1 | | A2 | | A3 | | P1 | | P2 | | |
Grade	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1	799	30.27	414	15.68	473	17.92	571	21.63	383	14.51	2640
2	634	22.59	379	13.50	476	16.96	709	25.26	609	21.70	2807
3	514	19.02	421	15.58	396	14.65	600	22.20	772	28.56	2703
4	330	13.00	335	13.20	362	14.26	544	21.43	967	38.10	2538
5	302	12.72	270	11.37	312	13.14	452	19.03	1039	43.75	2375
6	317	14.73	198	9.20	240	11.15	481	22.35	916	42.57	2152
7	269	12.95	180	8.67	222	10.69	416	20.03	990	47.66	2077
8	262	12.88	183	9.00	191	9.39	372	18.29	1026	50.44	2034
9	209	11.83	174	9.85	221	12.51	371	21.01	791	44.79	1766
10	174	11.00	124	7.84	190	12.01	306	19.34	788	49.81	1582
11	145	9.51	116	7.61	151	9.91	303	19.88	809	53.08	1524
12	278	11.54	189	7.85	240	9.96	473	19.63	1229	51.02	2409
Total	4233	15.91	2983	11.21	3474	13.06	5598	21.04	10319	38.78	26607

Table 4.9.2 H

Proficiency Level by Grade: Overall

| | Overall Proficiency Range | | | | | | | | | | |
| | A1 | | A2 | | A3 | | P1 | | P2 | | |
Grade	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1	847	33.02	419	16.34	665	25.93	436	17.00	198	7.72	2565
2	677	24.84	381	13.98	703	25.80	599	21.98	365	13.39	2725
3	564	21.32	391	14.78	662	25.03	661	24.99	367	13.88	2645
4	378	15.22	320	12.88	585	23.55	680	27.38	521	20.97	2484
5	356	15.38	270	11.67	493	21.31	617	26.66	578	24.98	2314
6	330	15.71	231	10.99	497	23.66	648	30.84	395	18.80	2101
7	286	14.29	191	9.54	430	21.48	598	29.87	497	24.83	2002
8	287	14.58	189	9.60	378	19.20	593	30.12	522	26.51	1969
9	226	13.19	181	10.57	367	21.42	546	31.87	393	22.94	1713
10	185	12.03	146	9.49	297	19.31	478	31.08	432	28.09	1538
11	157	10.57	136	9.16	309	20.81	422	28.42	461	31.04	1485
12	310	13.29	201	8.62	452	19.37	649	27.82	721	30.90	2333
Total	4603	17.79	3056	11.81	5838	22.56	6927	26.77	5450	21.06	25874

4.10 Participation by Disability

4.10.1 Participation by Disability

Table 4.10.1

Participation by Disability

		Secondary Disability														Total
		AS	DB	DD	ED	HI	ID	MD	OHI	OI	SLD	SLI	TBI	VI	NSD	
Primary Disability	AS	11	5	86	4	31	606	88	107	10	59	1650		13	6225	8895
	DB			3			5	1		1		3			6	19
	DD	43	4	2	3	27	77	22	53	11	16	374		9	1064	1705
	ED	2		1		1	13	2			3	6			33	61
	HI	5		2			16	8	4	1	3	27		1	55	122
	ID	285	8	32	22	115	20	107	347	126	74	1788	5	60	6151	9140
	MD	79	11	12	4	36	165	110	76	40	13	298	4	65	1269	2182
	OHI	31	1	24		14	97	26	3	11	13	213	3	24	795	1255
	OI			3		1	15	9	7			28		4	58	125
	SLD	3		1	1	7	6	3	10	1		95		1	321	449
	SLI	6		10		1	9	1	5	1	6	2		1	196	238
	TBI	2					11	2	5	1	2	19		4	70	116
	VI	3		5	1	1	9	2	4		1	3			14	43
	NPD	11	5	86	4	31	606	88	107	10	59	1650		13	6225	2689
Total		478	29	183	35	234	1060	381	622	203	190	4516	12	182	18914	27039

Acronyms for Table 4.10.1

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

5.   Analyses of Test Forms: Overview

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

5.1   Background

5.1.1   Measurement Models Used

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

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

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

where

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

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

B_n = ability of person "n"

D_i = difficulty of task "i"

F_k = calibration of step "k" on the rating scale

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

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

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

According to Linacre (2002):

values greater than 2.0 "distort or degrade the measurement system"

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

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

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

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

5.1.2 Sampling

The results presented in most of the tables in Chapter 6 are based on the full data set of all students who were administered operational Series 503 of Alternate ACCESS for ELLs in the academic year 2021-2022. 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

$\sigma_c{}^2$ = 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)

---

[4] In Series 503, 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.

No equating summary is presented because the Alternate ACCESS Series 503 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*

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., 503).

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
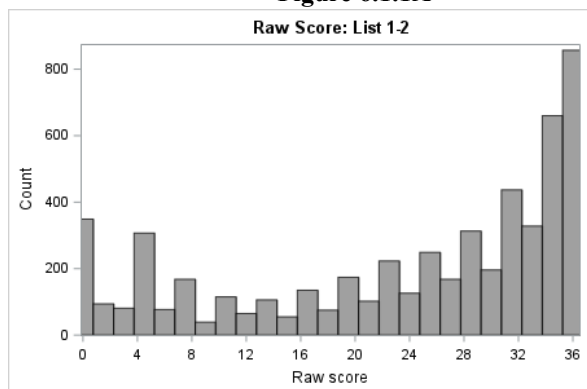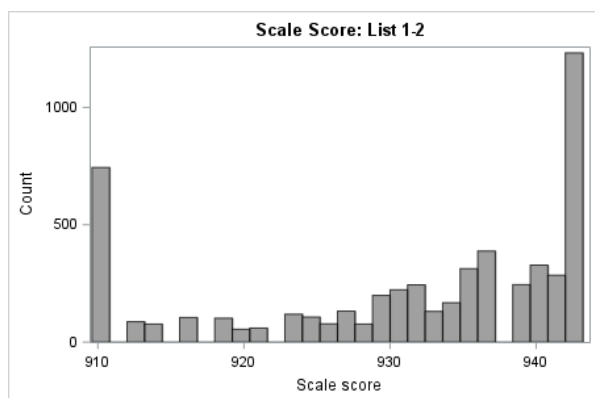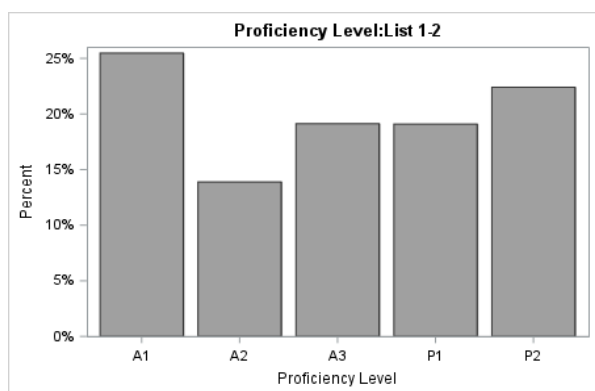
Table 6.1.1A

Raw Score Descriptive Statistics: List 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2668	0	36	22.04	12.31
2	2831	0	36	24.86	11.79
Total	5499	0	36	23.49	12.13

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	2668	910	943	929.76	11.56
2	2831	910	943	932.38	11.04
Total	5499	910	943	931.11	11.37

Figure 6.1.1C

Table 6.1.1C

Proficiency Level Distribution: List 1-2

Level	Grade 1 Count	Grade 1 Percent	Grade 2 Count	Grade 2 Percent	Total Count	Total Percent
A1	794	29.76	607	21.44	1401	25.48
A2	406	15.22	358	12.65	764	13.89
A3	510	19.12	542	19.15	1052	19.13
P1	478	17.92	572	20.20	1050	19.09
P2	480	17.99	752	26.56	1232	22.40
Total	2668	100	2831	100	5499	100

Table 6.1.1D

Equating Summary: List 1-2

No equating summary is presented because the Alternate ACCESS Series 503 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.

Figure 6.1.1D
Test Characteristic Curve: List 1-2 S402

Figure 6.1.1E
Test Information Function: List 1-2 S402
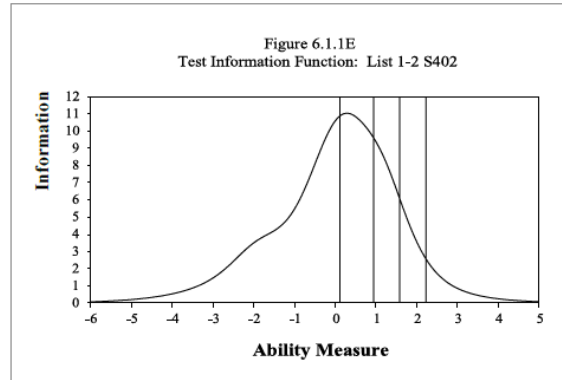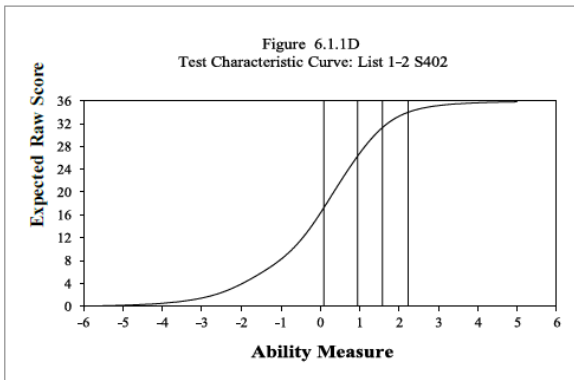
Table 6.1.1E

Reliability: List 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
5499	9	0.941	2.955

Table 6.1.1F

Item Analysis Summary: List 1-2

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	MOSR	9	0.00	71.17%	1.18	1.58
DIF Summary			Male/Female		Hispanic/Other	
	DIF Level		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA		3	6	6	3
	BB		0	0	0	0
	CC		0	0	0	0

Table 6.1.1G

Complete Item Analysis: List 1-2

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Infit Mnsq	Outfit Mnsq	Point Biserial	DIF Level (M/F)	Favored Group (M/F)	DIF Level (H/O)	Favored Group (H/O)
1. L1_A1_103	-1.59		86.00%	2.90	3.78	0.59	AA	F	AA	H
2. L2_A2_103	0.43	Yes	67.00%	1.33	1.48	0.78	AA	M	AA	O
3. L3_A2_103	0.17	Yes	73.25%	1.06	0.99	0.84	AA	M	AA	O
4. L4_A3_103	-0.56	Yes	75.75%	0.93	0.69	0.86	AA	F	AA	H
5. L5_A3_103	-0.12	Yes	73.50%	0.79	0.62	0.88	AA	F	AA	O
6. L6_P1_103	0.75	Yes	60.00%	1.10	1.08	0.79	AA	F	AA	H
7. L7_P1_103	0.98	Yes	58.50%	0.98	0.90	0.80	AA	M	AA	O
8. L8_P2_103	0.41	Yes	63.25%	0.85	0.65	0.85	AA	M	AA	O
9. L9_P2_103	0.96	Yes	55.00%	1.04	0.94	0.77	AA	F	AA	H

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

^ Truncate.   * 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	7.05	7.05	A1	5.69	5.69
1	A1	0.79	7.83	A1	0.35	6.04
2	A1	1.42	9.26	A1	0.88	6.92
3	A1	1.65	10.91	A1	1.31	8.23
4	A1	4.69	15.59	A1	3.36	11.59
5	A1	1.91	17.50	A1	1.27	12.86
6	A1	1.57	19.08	A1	1.24	14.09
7	A1	2.25	21.33	A1	1.59	15.68
8	A1	1.12	22.45	A1	1.17	16.85
9	A1	0.71	23.16	A1	0.71	17.56
10	A1	1.35	24.51	A1	0.67	18.23
11	A1	1.42	25.94	A1	0.78	19.00
12	A1	1.54	27.47	A1	0.85	19.85
13	A1	1.24	28.71	A1	0.74	20.59
14	A1	1.05	29.76	A1	0.85	21.44
15	A2	1.05	30.81	A2	0.95	22.39
16	A2	1.61	32.42	A2	1.24	23.63
17	A2	1.20	33.62	A2	0.88	24.51
18	A2	1.20	34.82	A2	1.52	26.03
19	A2	1.39	36.21	A2	1.41	27.45
20	A2	2.10	38.31	A2	1.45	28.89
21	A2	2.02	40.33	A2	1.70	30.59
22	A2	1.99	42.32	A2	1.55	32.14
23	A2	2.66	44.98	A2	1.94	34.09
24	A3	2.25	47.23	A3	2.33	36.42
25	A3	2.21	49.44	A3	2.08	38.50
26	A3	2.55	51.99	A3	2.23	40.73
27	A3	2.59	54.57	A3	3.50	44.22
28	A3	2.74	57.31	A3	2.12	46.34
29	A3	3.45	60.76	A3	3.11	49.45
30	A3	3.34	64.09	A3	3.78	53.23
31	P1	3.26	67.35	P1	3.71	56.94
32	P1	3.94	71.29	P1	4.95	61.89
33	P1	6.30	77.59	P1	5.65	67.54
34	P1	4.42	82.01	P1	5.90	73.44
35	P2	5.77	87.78	P2	7.81	81.24
36	P2	12.22	100.00	P2	18.76	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

Raw Score: Read 1-2
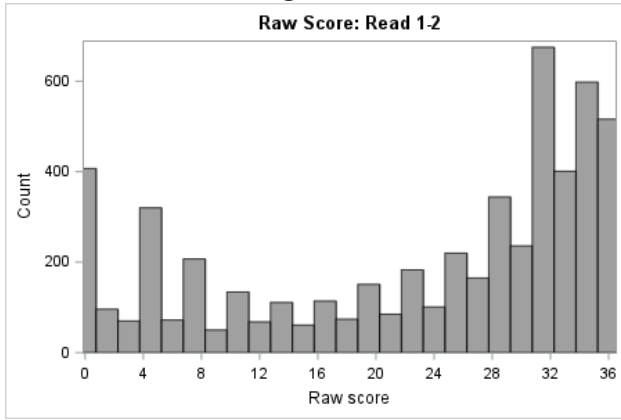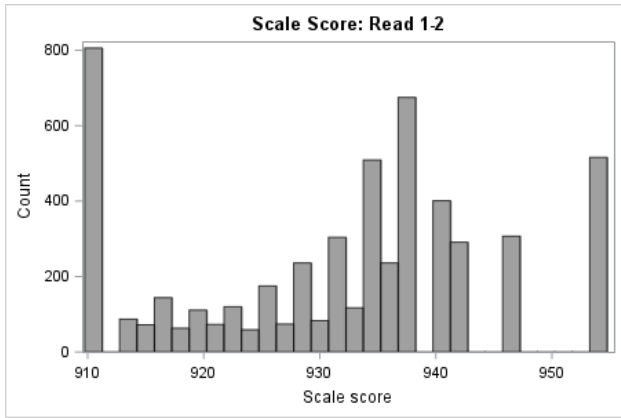
Figure 6.1.2B
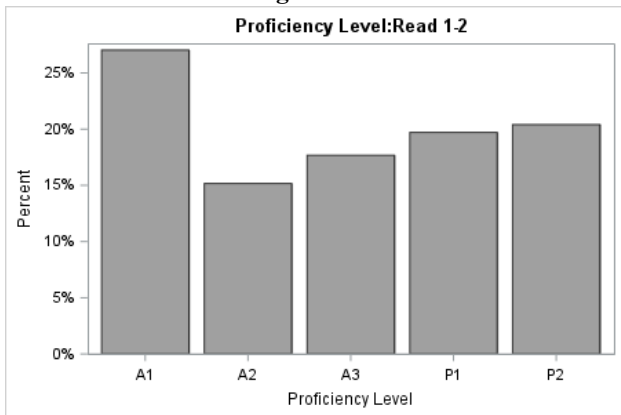
Scale Score: Read 1-2

Table 6.1.2A

Raw Score Descriptive Statistics: Read 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2649	0	36	21.34	12.38
2	2810	0	36	24.03	11.93
Total	5459	0	36	22.72	12.23

Table 6.1.2B

Scale Score Descriptive Statistics: Read 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2649	910	954	930.09	13.02
2	2810	910	954	933.10	13.07
Total	5459	910	954	931.64	13.13

Figure 6.1.2C

Proficiency Level: Read 1-2

Table 6.1.2C

Proficiency Level Distribution: Read 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	820	30.96	656	23.35	1476	27.04
A2	436	16.46	392	13.95	828	15.17
A3	480	18.12	485	17.26	965	17.68
P1	482	18.20	594	21.14	1076	19.71
P2	431	16.27	683	24.31	1114	20.41
Total	2649	100	2810	100	5459	100

Table 6.1.2D

Equating Summary: Read 1-2

No equating summary is presented because the Alternate ACCESS Series 503 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 6.1.2D
Test Characteristic Curve: Read 1-2 S402

Figure 6.1.2E
Test Information Function: Read 1-2 S402
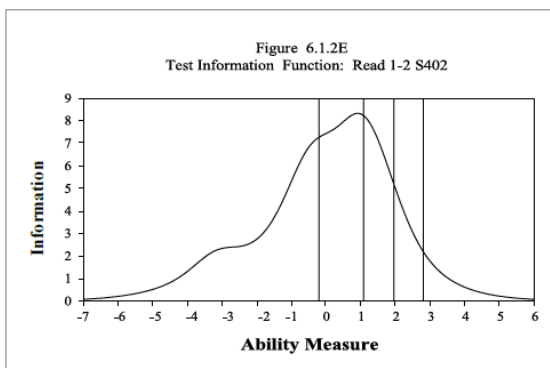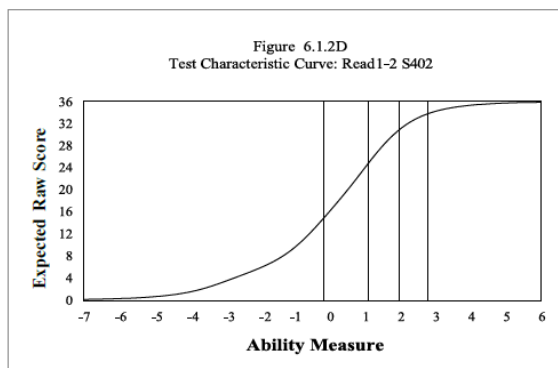
Table 6.1.2E

Reliability: Read 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
5459	9	0.948	2.792

Table 6.1.2F

Item Analysis Summary: Read 1-2

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	MOSR	9	0.00	66.42%	1.31	1.37
DIF Summary			Male/Female		Hispanic/Other	
	DIF Level		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA		4	5	4	5
	BB		0	0	0	0
	CC		0	0	0	0

Table 6.1.2G

Complete Item Analysis: Read 1-2

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Infit Mnsq	Outfit Mnsq	Point Biserial	M/F DIF Level	M/F Favored Group	H/O DIF Level	H/O Favored Group
1.R1_A1_103	-2.22		85.00%	3.02	3.61	0.60	AA	F	AA	H
2.R2_A2_103	-0.64	Yes	72.50%	1.39	1.11	0.83	AA	F	AA	O
3.R3_A2_103	-0.73	Yes	74.00%	1.26	1.04	0.83	AA	M	AA	H
4.R4_A3_103	-0.44	Yes	70.00%	1.18	0.98	0.86	AA	F	AA	O
5.R5_A3_103	0.73	Yes	59.00%	1.28	1.09	0.82	AA	M	AA	H
6.R6_P1_103	0.83	Yes	56.75%	0.94	0.83	0.85	AA	M	AA	O
7.R7_P1_103	1.01	Yes	54.25%	0.83	0.81	0.85	AA	F	AA	O
8.R8_P2_103	1.62	Yes	48.50%	0.82	0.87	0.81	AA	M	AA	H
9.R9_P2_103	0.94	Yes	56.75%	0.97	0.77	0.85	AA	M	AA	H

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	8.53	8.53	A1	6.44	6.44
1	A1	0.57	9.10	A1	0.32	6.76
2	A1	1.43	10.53	A1	1.21	7.97
3	A1	1.59	12.12	A1	1.00	8.97
4	A1	4.61	16.72	A1	3.95	12.92
5	A1	1.93	18.65	A1	1.28	14.20
6	A1	1.74	20.39	A1	0.93	15.12
7	A1	3.02	23.41	A1	2.28	17.40
8	A1	1.06	24.46	A1	1.25	18.65
9	A1	1.36	25.82	A1	0.50	19.15
10	A1	1.36	27.18	A1	0.89	20.04
11	A1	1.36	28.54	A1	1.32	21.35
12	A1	1.43	29.97	A1	1.07	22.42
13	A1	0.98	30.96	A1	0.93	23.35
14	A2	1.17	32.13	A2	1.00	24.34
15	A2	1.09	33.22	A2	1.14	25.48
16	A2	0.94	34.16	A2	1.10	26.58
17	A2	1.21	35.37	A2	0.93	27.51
18	A2	1.32	36.69	A2	1.39	28.90
19	A2	1.81	38.51	A2	1.35	30.25
20	A2	1.13	39.64	A2	1.25	31.49
21	A2	1.89	41.53	A2	1.25	32.74
22	A2	1.77	43.30	A2	1.28	34.02
23	A2	2.08	45.38	A2	1.60	35.62
24	A2	2.04	47.41	A2	1.67	37.30
25	A3	2.19	49.60	A3	1.60	38.90
26	A3	2.23	51.83	A3	2.06	40.96
27	A3	3.06	54.89	A3	2.99	43.95
28	A3	3.06	57.95	A3	2.67	46.62
29	A3	3.51	61.46	A3	3.38	50.00
30	A3	4.08	65.53	A3	4.56	54.56
31	P1	4.38	69.91	P1	4.59	59.15
32	P1	7.17	77.09	P1	8.54	67.69
33	P1	6.64	83.73	P1	8.01	75.69
34	P2	4.00	87.73	P2	6.58	82.28
35	P2	4.76	92.49	P2	6.44	88.72
36	P2	7.51	100.00	P2	11.28	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
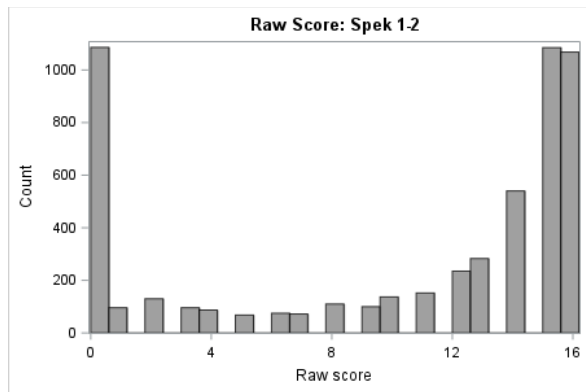
Table 6.1.3A

Raw Score Descriptive Statistics: Spek 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2626	0	16	9.43	6.36
2	2794	0	16	10.65	6.16
Total	5420	0	16	10.06	6.29

Figure 6.1.3B
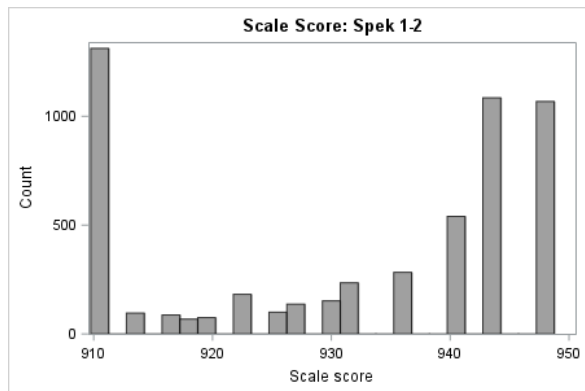
Table 6.1.3B

Scale Score Descriptive Statistics: Spek 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2626	910	948	930.03	14.95
2	2794	910	948	933.22	14.77
Total	5420	910	948	931.67	14.94

Figure 6.1.3C
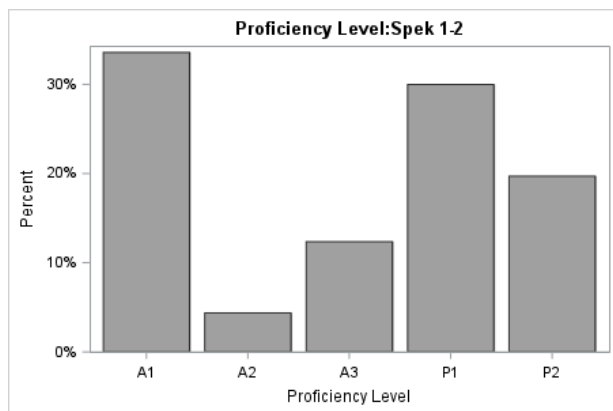
Table 6.1.3C

Proficiency Level Distribution: Spek 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	989	37.66	831	29.74	1820	33.58
A2	124	4.72	113	4.04	237	4.37
A3	356	13.56	314	11.24	670	12.36
P1	739	28.14	886	31.71	1625	29.98
P2	418	15.92	650	23.26	1068	19.70
Total	2626	100	2794	100	5420	100

Table 6.1.3D

Equating Summary: Spek 1-2

No equating summary is presented because the Alternate ACCESS Series 503 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
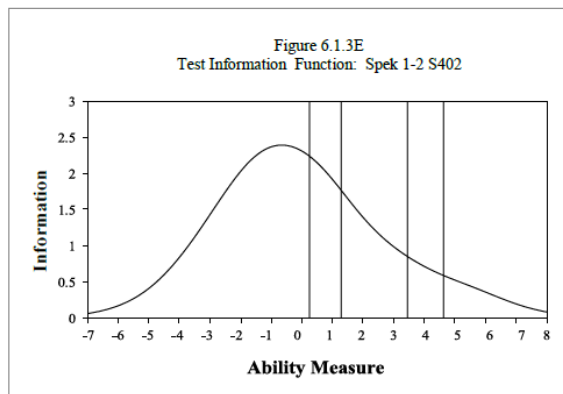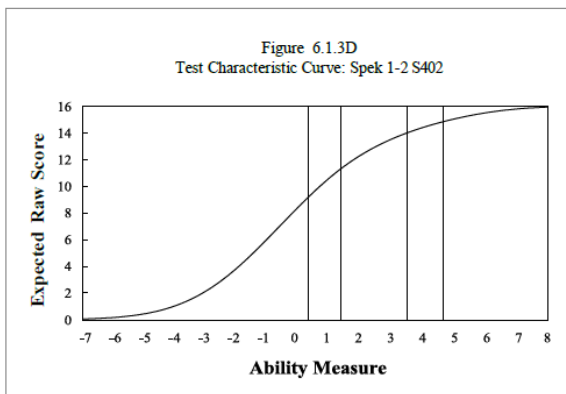
Table 6.1.3E

Reliability: Spek 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
5420	8	0.963	1.207

Table 6.1.3F

Item Analysis Summary: Spek 1-2

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	CR	8	0.00	67.13%	1.11	0.95
DIF Summary		Male/Female			Hispanic/Other	
	DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)	
	AA	5	3	5	3	
	BB	0	0	0	0	
	CC	0	0	0	0	

Table 6.1.3G

Complete Item Analysis: Spek 1-2

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Infit Mnsq	Outfit Mnsq	Point Biserial	M/F DIF Level	M/F Favored Group	H/O DIF Level	H/O Favored Group
1.S1_A1_103	-2.09		75.50%	1.61	1.77	0.87	AA	F	AA	O
2.S2_A2_103	-1.04		72.00%	1.44	1.37	0.89	AA	M	AA	O
3.S3_A3_103	-0.41	Yes	70.00%	1.01	0.75	0.92	AA	M	AA	O
4.S4_A1_103	-1.20	Yes	72.00%	1.11	0.99	0.91	AA	F	AA	O
5.S5_A2_103	0.00	Yes	68.50%	1.08	0.92	0.91	AA	M	AA	H
6.S6_A3_103	-0.23	Yes	68.50%	0.81	0.58	0.93	AA	F	AA	O
7.S7_P1_103	1.51	Yes	61.00%	1.01	0.70	0.86	AA	F	AA	O
8.S8_P2_103	4.55	Yes	39.50%	0.67	1.08	0.67	AA	M	AA	H

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	22.43	22.43	A1	17.79	17.79
1	A1	1.98	24.41	A1	1.57	19.36
2	A1	2.74	27.15	A1	2.08	21.44
3	A1	2.09	29.25	A1	1.47	22.91
4	A1	1.68	30.92	A1	1.54	24.45
5	A1	1.29	32.22	A1	1.22	25.66
6	A1	1.45	33.66	A1	1.32	26.99
7	A1	1.52	35.19	A1	1.15	28.13
8	A1	2.48	37.66	A1	1.61	29.74
9	A2	1.90	39.57	A2	1.79	31.53
10	A2	2.82	42.38	A2	2.25	33.79
11	A3	3.24	45.62	A3	2.40	36.18
12	A3	4.80	50.42	A3	3.90	40.09
13	A3	5.52	55.94	A3	4.94	45.03
14	P1	10.09	66.03	P1	9.84	54.87
15	P1	18.05	84.08	P1	21.87	76.74
16	P2	15.92	100.00	P2	23.26	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
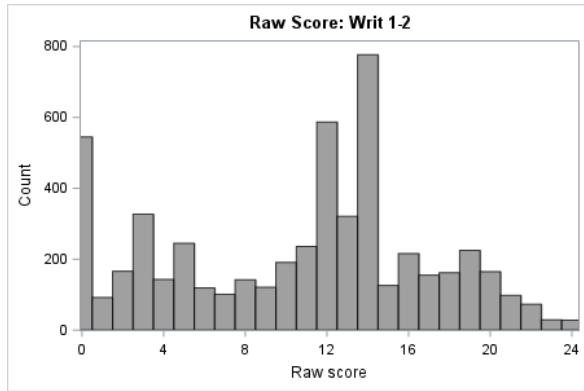
Table 6.1.4A

Raw Score Descriptive Statistics: Writ 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2616	0	24	9.83	6.19
2	2774	0	24	11.25	6.38
Total	5390	0	24	10.56	6.33

Figure 6.1.4B
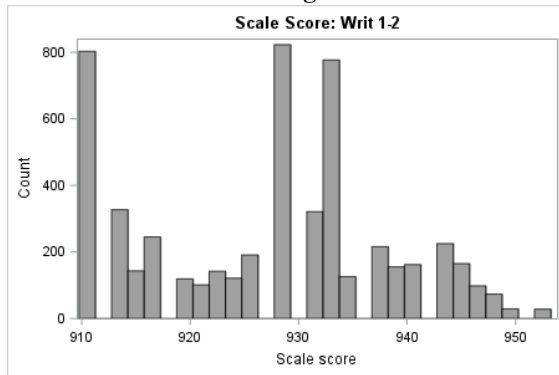
Table 6.1.4B

Scale Score Descriptive Statistics: Writ 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2616	910	953	925.81	11.02
2	2774	910	953	928.41	11.44
Total	5390	910	953	927.15	11.31

Figure 6.1.4C
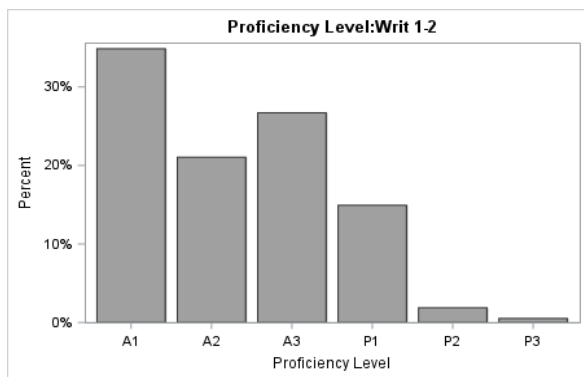
Table 6.1.4C

Proficiency Level Distribution: Writ 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	1024	39.14	856	30.86	1880	34.88
A2	568	21.71	567	20.44	1135	21.06
A3	671	25.65	769	27.72	1440	26.72
P1	310	11.85	495	17.84	805	14.94
P2	33	1.26	69	2.49	102	1.89
P3	10	0.38	18	0.65	28	0.52
Total	2616	100	2774	100	5390	100

Table 6.1.4D

Equating Summary: Writ 1-2

No equating summary is presented because the Alternate ACCESS Series 503 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 6.1.4D
Test Characteristic Curve: Writ 1-2 S402

Figure 6.1.4E
Test Information Function: Writ 1-2 S402
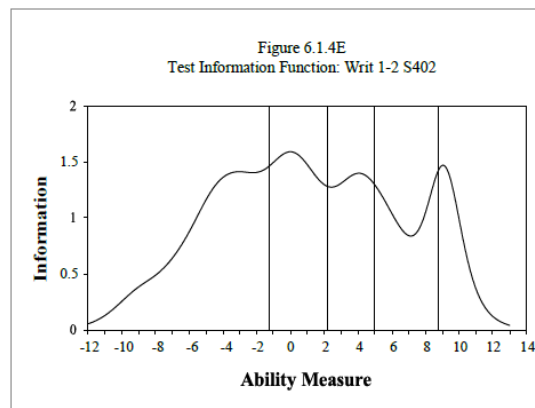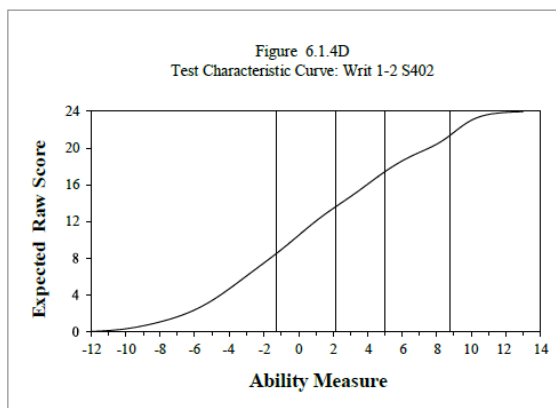
Table 6.1.4E

Reliability: Writ 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
5390	10	0.932	1.651

Table 6.1.4F

Item Analysis Summary: Writ 1-2

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	CR	10	0.00	52.65%	1.18	3.83
DIF Summary		Male/Female		Hispanic/Other		
	DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)	
	AA	6	4	5	5	
	BB	0	0	0	0	
	CC	0	0	0	0	

Table 6.1.4G

Complete Item Analysis: Writ 1-2

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Infit Mnsq	Outfit Mnsq	Point Biserial	DIF M/F DIF Level	DIF M/F Favored Group	DIF H/O DIF Level	DIF H/O Favored Group
1.W1_A1_103	-5.54		78.50%	2.00	9.90	0.69	AA	F	AA	H
2.W2_A2_103	-4.78	Yes	76.00%	1.42	9.90	0.76	AA	M	AA	O
3.W3_A3_103	-1.31	Yes	61.50%	1.41	7.92	0.81	AA	M	AA	O
4.W4_P1_103	1.73	Yes	41.00%	1.10	3.17	0.81	AA	F	AA	O
5.W5_A1_103	-2.63	Yes	67.50%	1.29	4.62	0.81	AA	F	AA	H
6.W6_A2_103	-2.18	Yes	66.00%	1.17	3.50	0.82	AA	M	AA	H
7.W7_A3_103	-0.34	Yes	54.00%	0.95	9.38	0.84	AA	M	AA	O
8.W8_P1_103	2.54	Yes	37.00%	0.73	0.96	0.83	AA	F	AA	O
9.W9_P3_103	6.84	Yes	12.75%	1.06	9.56	0.58	AA	M	AA	O
10.W10_P3_103	7.17	Yes	10.75%	1.00	9.90	0.53	AA	M	AA	H

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	11.12	11.12	A1	9.16	9.16
1	A1	1.95	13.07	A1	1.48	10.63
2	A1	3.52	16.59	A1	2.67	13.30
3	A1	7.07	23.66	A1	5.12	18.42
4	A1	2.94	26.61	A1	2.38	20.80
5	A1	4.78	31.38	A1	4.33	25.13
6	A1	2.91	34.29	A1	1.55	26.68
7	A1	1.91	36.20	A1	1.84	28.51
8	A1	2.94	39.14	A1	2.34	30.86
9	A2	2.41	41.55	A2	2.09	32.95
10	A2	3.94	45.49	A2	3.17	36.12
11	A2	4.09	49.58	A2	4.65	40.77
12	A2	11.28	60.86	A2	10.53	51.30
13	A3	5.81	66.67	A3	6.09	57.39
14	A3	14.03	80.70	A3	14.78	72.17
15	A3	2.29	82.99	A3	2.38	74.55
16	A3	3.52	86.51	A3	4.47	79.02
17	P1	2.75	89.26	P1	2.99	82.01
18	P1	2.52	91.78	P1	3.46	85.47
19	P1	3.17	94.95	P1	5.12	90.59
20	P1	2.18	97.13	P1	3.89	94.48
21	P1	1.22	98.36	P1	2.38	96.86
22	P2	0.99	99.35	P2	1.69	98.56
23	P2	0.27	99.62	P2	0.79	99.35
24	P3	0.38	100.00	P3	0.65	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	2610	910	946	930.24	12.44
2	2784	910	946	933.12	12.12
Total	5394	910	946	931.73	12.36
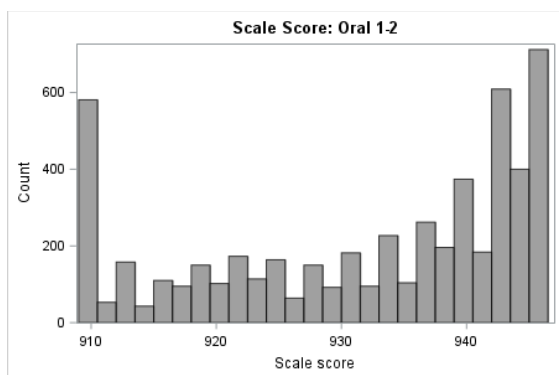
Figure 6.1.5B
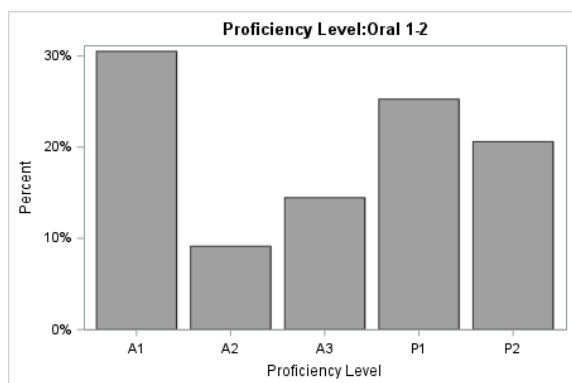
Table 6.1.5C

Proficiency Level Distribution: Oral 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	903	34.60	743	26.69	1646	30.52
A2	252	9.66	241	8.66	493	9.14
A3	418	16.02	362	13.00	780	14.46
P1	612	23.45	751	26.98	1363	25.27
P2	425	16.28	687	24.68	1112	20.62
Total	2610	100	2784	100	5394	100

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	129.258	0.941
Speaking	0.5	223.153	0.963
Oral		152.710	0.974

*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
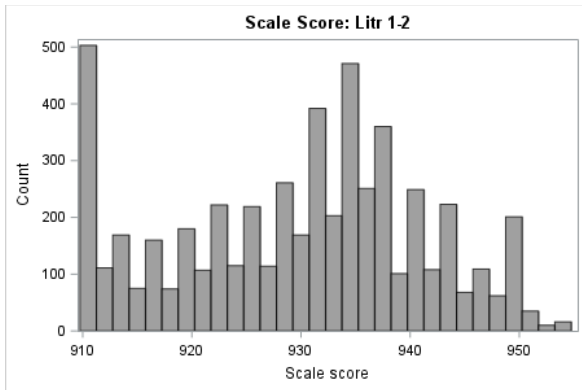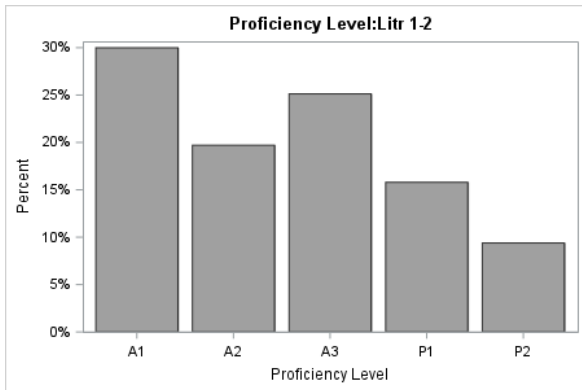
Scale Score: Litr 1-2

Table 6.1.6A

Scale Score Descriptive Statistics: Litr1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2593	910	954	928.30	11.19
2	2745	910	954	931.10	11.45
Total	5338	910	954	929.74	11.41

Figure 6.1.6B

Proficiency Level: Litr 1-2

Table 6.1.6C

Proficiency Level Distribution: Litr 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	888	34.25	713	25.97	1601	29.99
A2	548	21.13	504	18.36	1052	19.71
A3	643	24.80	698	25.43	1341	25.12
P1	336	12.96	507	18.47	843	15.79
P2	178	6.86	323	11.77	501	9.39
Total	2593	100	2745	100	5338	100

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	172.496	0.948
Writing	0.5	127.954	0.932
Literacy		130.214	0.966

*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
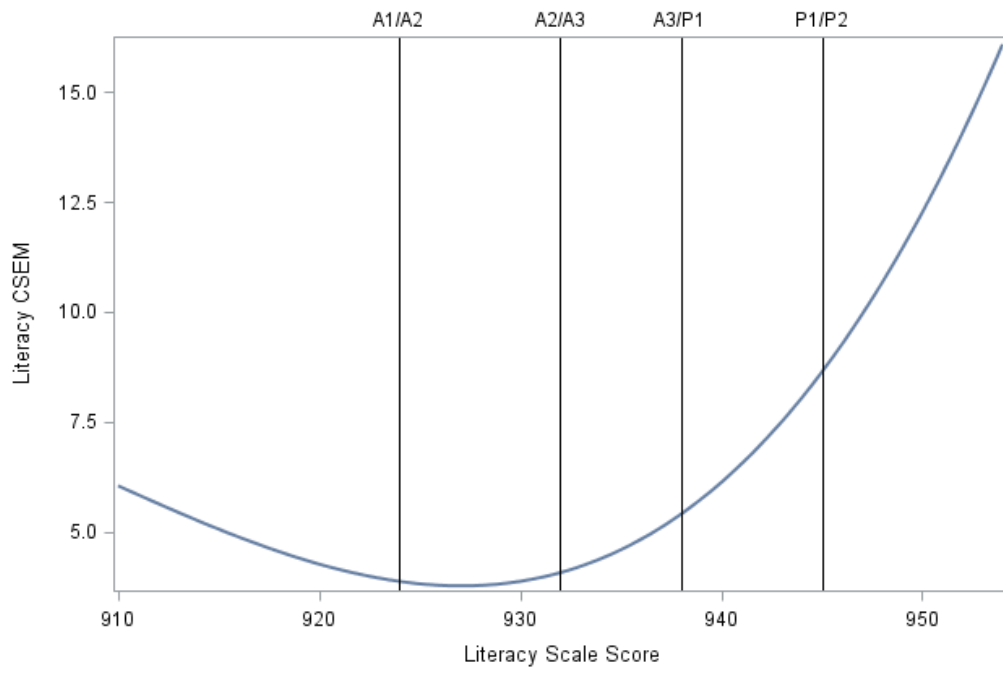
6.1.7 Comprehension Composite 1-2

Figure 6.1.7A
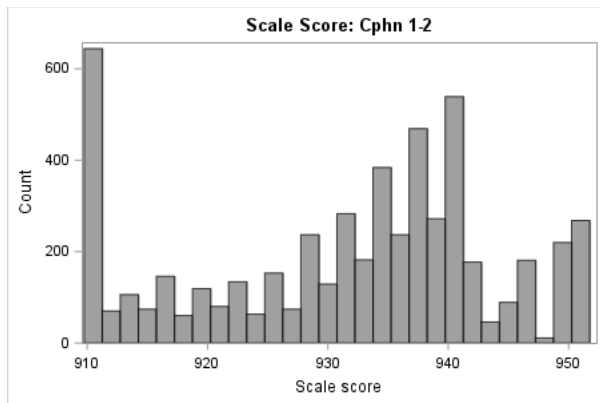
Figure 6.1.7B
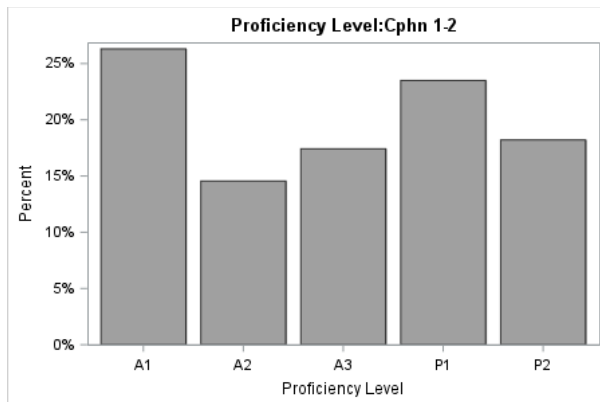
Table 6.1.7A

Scale Score Descriptive Statistics: Cphn 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2640	910	951	930.12	12.21
2	2807	910	951	932.97	12.09
Total	5447	910	951	931.59	12.23

Table 6.1.7C

Proficiency Level Distribution: Cphn 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	799	30.96	634	23.35	1433	27.04
A2	414	16.46	379	13.95	793	15.17
A3	473	18.12	476	17.26	949	17.68
P1	571	18.20	709	21.14	1280	19.71
P2	383	16.27	609	24.31	992	20.41
Total	2640	100	2807	100	5447	100

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	129.258	0.941
Reading	0.7	172.496	0.948
Comprehension		149.647	0.966

*Variances from students who had results in all four domains

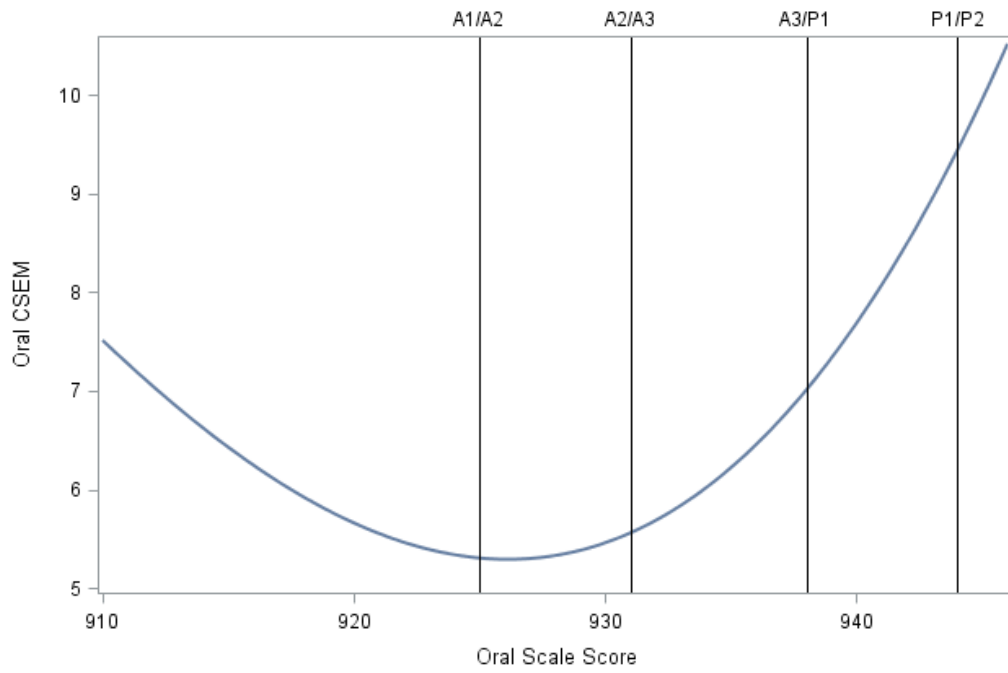Table 6.1.7F
n/a

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
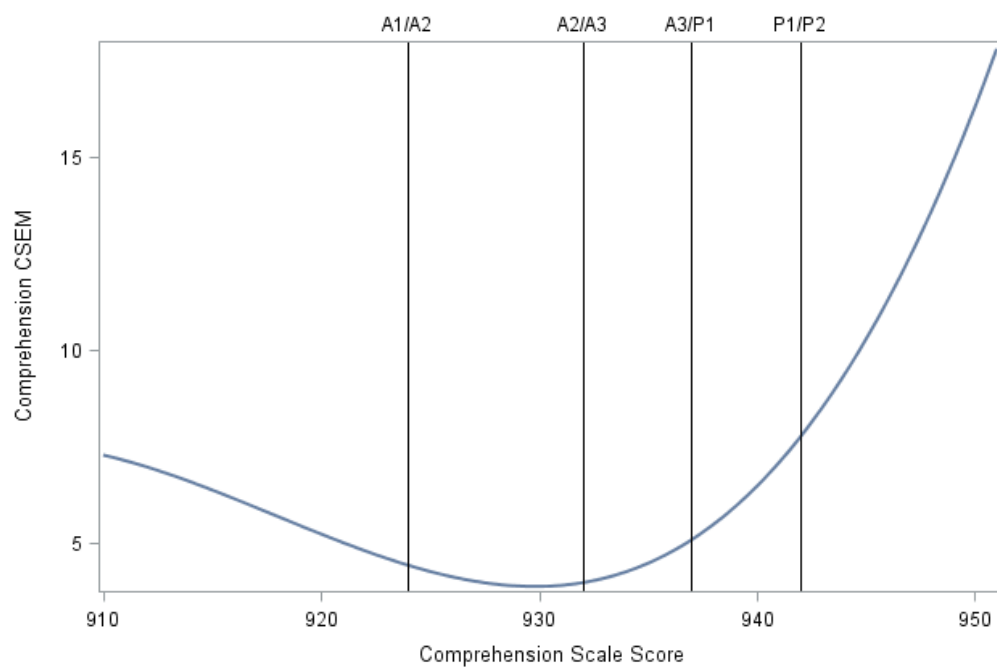
6.1.8 Overall Composite 1-2

Figure 6.1.8A
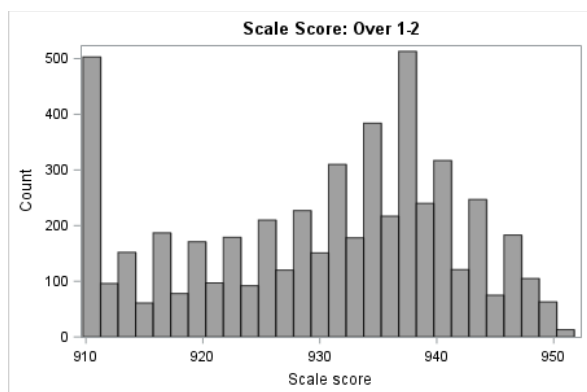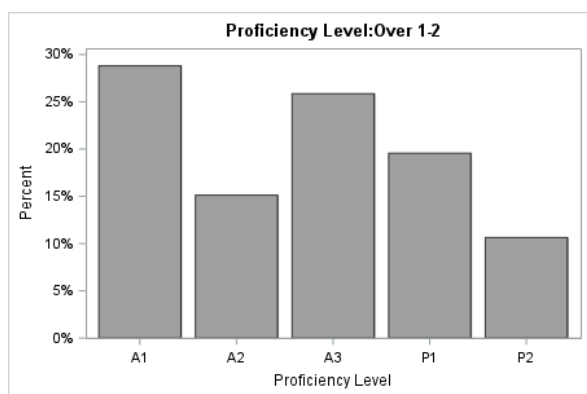
Table 6.1.8A

Scale Score Descriptive Statistics: Over 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	2565	910	951	928.71	11.13
2	2725	910	951	931.54	11.17
Total	5290	910	951	930.17	11.24

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	847	33.02	677	24.84	1524	28.81
A2	419	16.34	381	13.98	800	15.12
A3	665	25.93	703	25.80	1368	25.86
P1	436	17.00	599	21.98	1035	19.57
P2	198	7.72	365	13.39	563	10.64
Total	2565	100	2725	100	5290	100

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	129.258	0.941
Reading	0.35	172.496	0.948
Speaking	0.15	223.153	0.963
Writing	0.35	127.954	0.932
Overall Composite		126.262	0.980

 *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
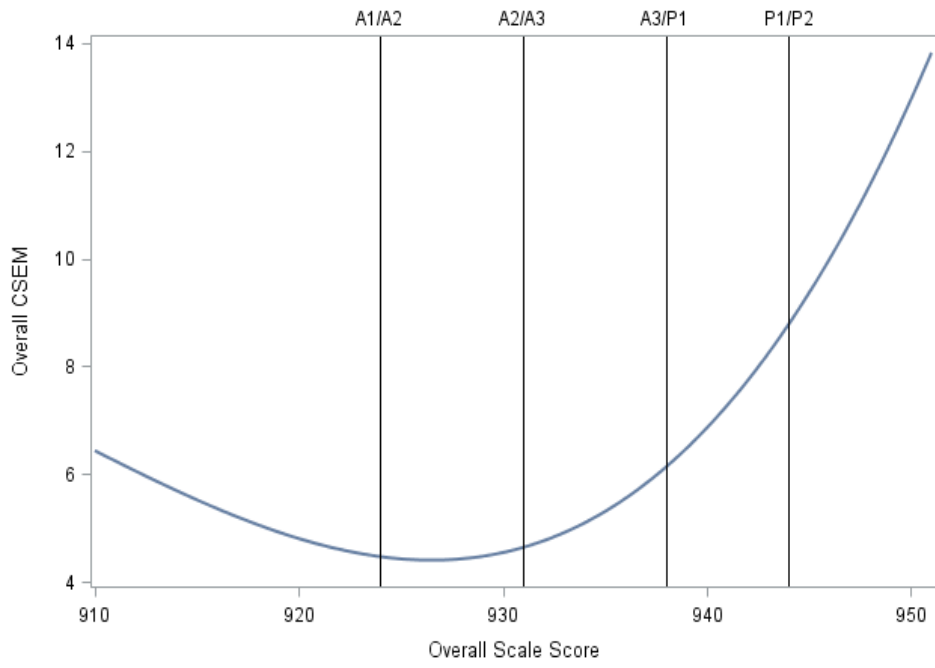
6.2   Grades: 3-5

6.2.1   Listening 3-5

Figure 6.2.1A
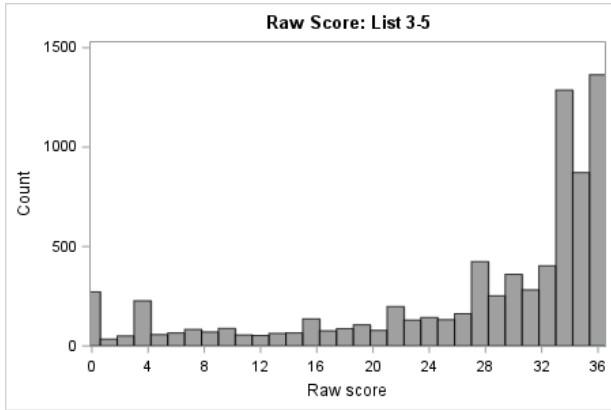
Figure 6.2.1B
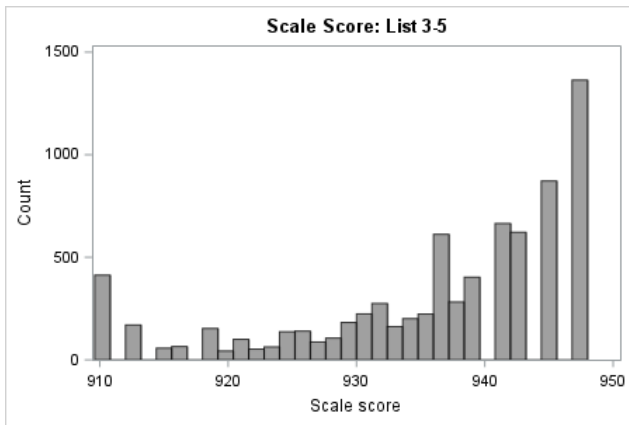
Figure 6.2.1C
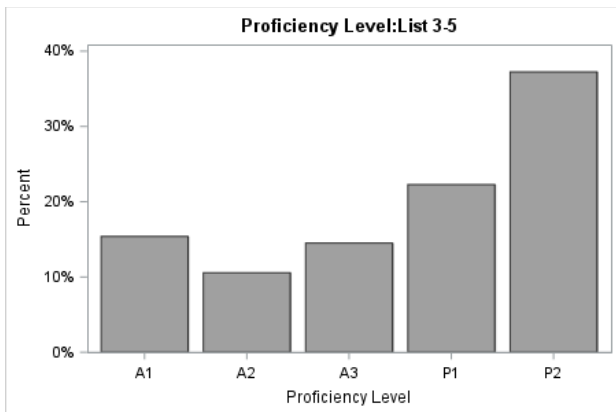
Table 6.2.1A

Raw Score Descriptive Statistics: List 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2728	0	36	25.55	11.11
4	2554	0	36	27.62	10.20
5	2391	0	36	28.12	10.18
Total	7673	0	36	27.04	10.58

Table 6.2.1B

Scale Score Descriptive Statistics: List 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2728	910	947	934.39	11.07
4	2554	910	947	936.58	10.33
5	2391	910	947	937.14	10.44
Total	7673	910	947	935.97	10.70

Table 6.2.1C

Proficiency Level Distribution: List 3-5

Level	Grade 3 Count	Grade 3 Percent	Grade 4 Count	Grade 4 Percent	Grade 5 Count	Grade 5 Percent	Total Count	Total Percent
A1	524	19.21	350	13.70	307	12.84	1181	15.39
A2	332	12.17	265	10.38	214	8.95	811	10.57
A3	427	15.65	367	14.37	319	13.34	1113	14.51
P1	597	21.88	579	22.67	534	22.33	1710	22.29
P2	848	31.09	993	38.88	1017	42.53	2858	37.25
Total	2728	100	2554	100	2391	100	7673	100

Table 6.2.1D

Equating Summary: List 3-5

No equating summary is presented because the Alternate ACCESS Series 503 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.

Figure 6.2.1D
Test Characteristic Curve: List 3-5 S103

Figure 6.2.1E
Test Information Function: List 3-5 S103
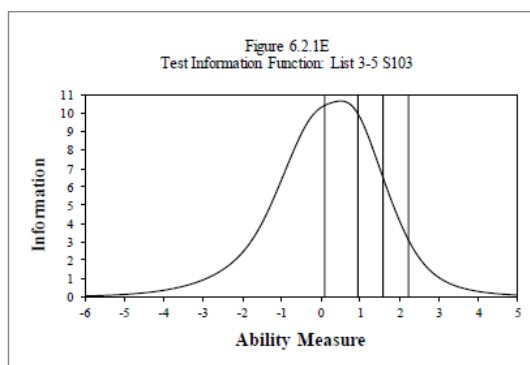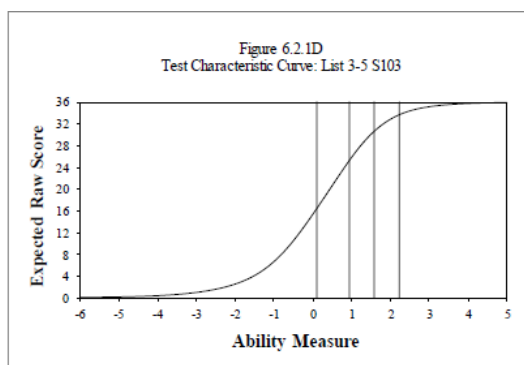
Table 6.2.1E

Reliability: List 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
7673	9	0.937	2.664

Table 6.2.1F

Item Analysis Summary: List 3-5

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	MOSR	9	0.00	79.97%	1.25	1.06
DIF Summary			Male/Female		Hispanic/Other	
	DIF Level		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA		4	5	6	3
	BB		0	0	0	0
	CC		0	0	0	0

Table 6.2.1G

Complete Item Analysis: List 3-5

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Fit Statistics Infit Mnsq	Fit Statistics Outfit Mnsq	Point Biserial	DIF M/F DIF Level	DIF M/F Favored Group	DIF H/O DIF Level	DIF H/O Favored Group
1. L1_A1_103	-1.12	Yes	90.50%	2.94	3.47	0.63	AA	F	AA	O
2. L2_A2_103	-0.31	Yes	84.25%	1.33	1.02	0.83	AA	M	AA	O
3. L3_A2_103	-0.45	Yes	85.75%	1.05	0.70	0.85	AA	F	AA	O
4. L4_A3_103	-0.43	Yes	84.25%	0.91	0.63	0.87	AA	F	AA	O
5. L5_A3_103	0.60	Yes	76.50%	1.14	0.99	0.82	AA	M	AA	O
6. L6_P1_103	0.64	Yes	73.25%	0.89	0.73	0.83	AA	F	AA	H
7. L7_P1_103	0.49	Yes	77.25%	0.85	0.66	0.87	AA	F	AA	H
8. L8_P2_103	0.76	Yes	72.50%	0.95	0.84	0.82	AA	M	AA	H
9. L9_P2_103	1.46	Yes	62.50%	1.00	1.00	0.74	AA	M	AA	H

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	4.18	4.18	A1	3.37	3.37	A1	3.01	3.01
1	A1	0.51	4.69	A1	0.43	3.80	A1	0.38	3.39
2	A1	0.84	5.54	A1	0.35	4.15	A1	0.75	4.14
3	A1	0.99	6.52	A1	0.55	4.70	A1	0.67	4.81
4	A1	2.82	9.35	A1	1.57	6.26	A1	2.22	7.03
5	A1	1.10	10.45	A1	0.78	7.05	A1	0.29	7.32
6	A1	0.73	11.18	A1	0.78	7.83	A1	1.05	8.36
7	A1	1.43	12.61	A1	0.98	8.81	A1	0.79	9.16
8	A1	1.28	13.89	A1	0.70	9.51	A1	0.71	9.87
9	A1	0.66	14.55	A1	0.59	10.10	A1	0.42	10.29
10	A1	0.66	15.21	A1	0.78	10.88	A1	0.29	10.58
11	A1	0.84	16.06	A1	0.82	11.71	A1	0.46	11.04
12	A1	0.84	16.90	A1	0.51	12.22	A1	0.67	11.71
13	A1	1.25	18.15	A1	0.78	13.00	A1	0.38	12.09
14	A1	1.06	19.21	A1	0.70	13.70	A1	0.75	12.84
15	A2	0.95	20.16	A2	0.86	14.57	A2	1.00	13.84
16	A2	0.84	21.00	A2	0.98	15.54	A2	0.67	14.51
17	A2	1.36	22.36	A2	1.06	16.60	A2	0.50	15.01
18	A2	1.50	23.86	A2	0.70	17.31	A2	1.17	16.19
19	A2	1.69	25.55	A2	1.37	18.68	A2	1.05	17.23
20	A2	1.14	26.69	A2	1.06	19.73	A2	0.84	18.07
21	A2	1.47	28.15	A2	1.10	20.83	A2	1.51	19.57
22	A2	1.14	29.29	A2	1.45	22.28	A2	1.09	20.66
23	A2	2.09	31.38	A2	1.80	24.08	A2	1.13	21.79
24	A3	2.13	33.50	A3	1.88	25.96	A3	1.55	23.34
25	A3	1.91	35.41	A3	1.49	27.45	A3	1.76	25.09
26	A3	1.83	37.24	A3	2.19	29.64	A3	2.34	27.44
27	A3	3.01	40.25	A3	2.58	32.22	A3	2.22	29.65
28	A3	3.34	43.59	A3	2.66	34.89	A3	2.68	32.33
29	A3	3.45	47.03	A3	3.56	38.45	A3	2.80	35.13
30	P1	5.06	52.09	P1	4.62	43.07	P1	4.35	39.48
31	P1	4.11	56.20	P1	3.37	46.44	P1	3.51	42.99
32	P1	4.66	60.85	P1	5.60	52.04	P1	5.56	48.56
33	P1	8.06	68.91	P1	9.08	61.12	P1	8.91	57.47
34	P2	7.70	76.61	P2	8.42	69.54	P2	8.24	65.70
35	P2	10.15	86.77	P2	11.43	80.97	P2	12.67	78.38
36	P2	13.23	100.00	P2	19.03	100.00	P2	21.62	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
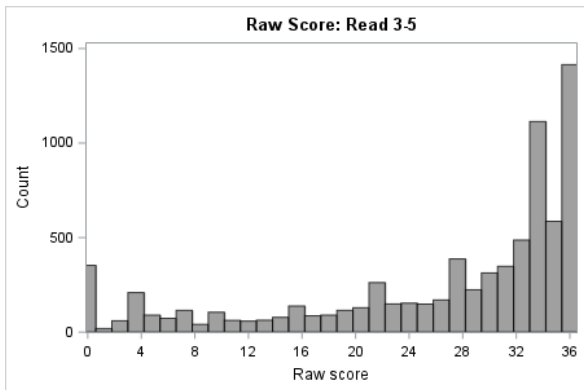
Figure 6.2.2B
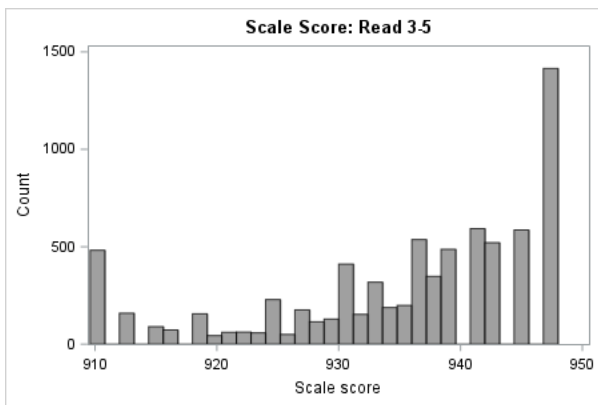
Figure 6.2.2C
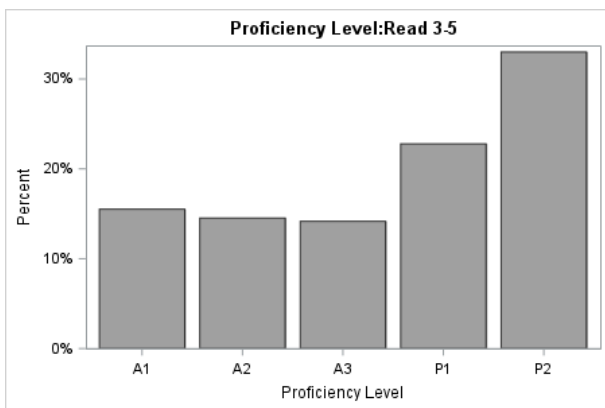
Table 6.2.2A

Raw Score Descriptive Statistics: Read 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2713	0	36	24.37	11.56
4	2547	0	36	26.87	10.44
5	2380	0	36	27.38	10.58
Total	7640	0	36	26.14	10.97

Table 6.2.2B

Scale Score Descriptive Statistics: Read 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2713	910	947	933.28	11.37
4	2547	910	947	935.85	10.44
5	2380	910	947	936.47	10.74
Total	7640	910	947	935.13	10.96

Table 6.2.2C

Proficiency Level Distribution: Read 3-5

Level	Grade 3 Count	Grade 3 Percent	Grade 4 Count	Grade 4 Percent	Grade 5 Count	Grade 5 Percent	Total Count	Total Percent
A1	538	19.83	336	13.19	311	13.07	1185	15.51
A2	454	16.73	360	14.13	296	12.44	1110	14.53
A3	383	14.12	365	14.33	334	14.03	1082	14.16
P1	640	23.59	603	23.67	498	20.92	1741	22.79
P2	698	25.73	883	34.67	941	39.54	2522	33.01
Total	2713	100	2547	100	2380	100	7640	100

Table 6.2.2D

Equating Summary: Read 3-5

No equating summary is presented because the Alternate ACCESS Series 503 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.

Figure 6.2.2D
Test Characteristic Curve: Read 3-5 S103

Figure 6.2.2E
Test Information Function: Read 3-5 S103
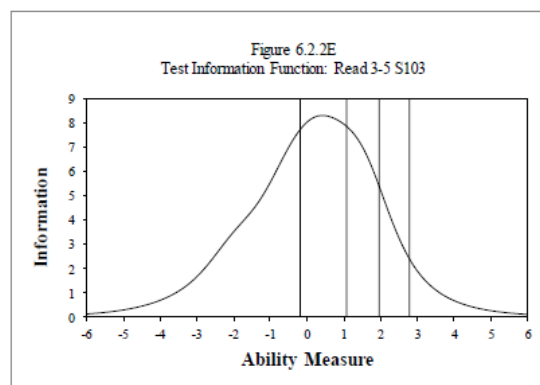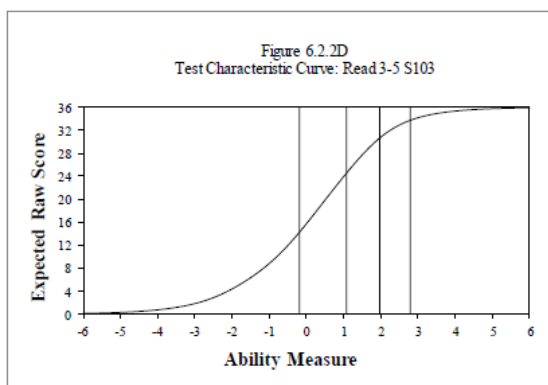
Table 6.2.2E

Reliability: Read 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
7640	9	0.948	2.792

Table 6.2.2F

Item Analysis Summary: Read 3-5

Item Summary				Average of % of Max. Possible Score Points		
	Item Type	No. of Items	Average Item Difficulty (in logits)		Average Infit Mean	Average Outfit Mean Square
	MOSR	9	0.00	75.94%	1.22	1.31
DIF Summary			Male/Female		Hispanic/Other	
	DIF Level		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA		5	4	4	5
	BB		0	0	0	0
	CC		0	0	0	0

Table 6.2.2G

Complete Item Analysis: Read 3-5

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Fit Statistics		Point Biserial	DIF			
				Infit Mnsq	Outfit Mnsq		M/F		H/O	
							DIF Level	Favored Group	DIF Level	Favored Group
1. R1_A1_103	-2.13	Yes	90.00%	2.96	4.89	0.61	AA	F	AA	O
2. R2_A2_103	-0.44	Yes	81.25%	1.56	1.17	0.83	AA	M	AA	O
3. R3_A2_103	-0.82	Yes	82.75%	1.27	0.91	0.85	AA	F	AA	H
4. R4_A3_103	0.07	Yes	79.75%	1.22	1.00	0.86	AA	M	AA	O
5. R5_A3_103	0.11	Yes	77.50%	0.98	0.70	0.89	AA	F	AA	H
6. R6_P1_103	1.14	Yes	67.75%	0.86	0.91	0.87	AA	F	AA	O
7. R7_P1_103	0.64	Yes	72.75%	0.76	0.60	0.90	AA	M	AA	O
8. R8_P2_103	1.55	Yes	61.00%	0.72	0.69	0.83	AA	M	AA	O
9. R9_P2_103	1.45	Yes	62.75%	0.82	0.80	0.83	AA	M	AA	H

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	5.57	5.57	A1	4.08	4.08	A1	4.12	4.12
1	A1	0.52	6.08	A1	0.08	4.16	A1	0.13	4.24
2	A1	1.25	7.34	A1	0.43	4.59	A1	0.63	4.87
3	A1	0.96	8.29	A1	0.43	5.03	A1	0.55	5.42
4	A1	2.54	10.84	A1	1.53	6.56	A1	2.14	7.56
5	A1	1.40	12.24	A1	1.22	7.77	A1	0.88	8.45
6	A1	1.22	13.45	A1	0.90	8.68	A1	0.71	9.16
7	A1	1.99	15.44	A1	1.10	9.78	A1	1.39	10.55
8	A1	0.77	16.22	A1	0.35	10.13	A1	0.46	11.01
9	A1	0.81	17.03	A1	0.63	10.76	A1	0.25	11.26
10	A1	0.92	17.95	A1	0.98	11.74	A1	0.46	11.72
11	A1	0.85	18.80	A1	0.82	12.56	A1	0.76	12.48
12	A1	1.03	19.83	A1	0.63	13.19	A1	0.59	13.07
13	A2	1.11	20.94	A2	0.90	14.10	A2	0.42	13.49
14	A2	1.25	22.19	A2	0.79	14.88	A2	1.01	14.50
15	A2	1.14	23.33	A2	1.53	16.41	A2	0.76	15.25
16	A2	0.70	24.03	A2	0.67	17.08	A2	0.59	15.84
17	A2	1.58	25.62	A2	1.02	18.10	A2	0.71	16.55
18	A2	1.11	26.72	A2	1.26	19.36	A2	1.18	17.73
19	A2	1.81	28.53	A2	1.41	20.77	A2	1.26	18.99
20	A2	1.88	30.41	A2	1.49	22.26	A2	1.68	20.67
21	A2	2.03	32.44	A2	1.53	23.79	A2	1.39	22.06
22	A2	1.88	34.32	A2	1.73	25.52	A2	1.68	23.74
23	A2	2.25	36.56	A2	1.81	27.33	A2	1.76	25.50
24	A3	2.29	38.85	A3	1.96	29.29	A3	1.72	27.23
25	A3	1.88	40.73	A3	1.92	31.21	A3	2.02	29.24
26	A3	2.36	43.09	A3	2.16	33.37	A3	2.14	31.39
27	A3	1.95	45.04	A3	2.83	36.20	A3	2.65	34.03
28	A3	2.73	47.77	A3	2.63	38.83	A3	2.44	36.47
29	A3	2.91	50.68	A3	2.83	41.66	A3	3.07	39.54
30	P1	4.90	55.58	P1	4.20	45.86	P1	3.07	42.61
31	P1	4.94	60.52	P1	5.03	50.88	P1	3.61	46.22
32	P1	6.49	67.01	P1	6.48	57.36	P1	6.13	52.35
33	P1	7.26	74.27	P1	7.97	65.33	P1	8.11	60.46
34	P2	5.64	79.91	P2	7.62	72.95	P2	7.31	67.77
35	P2	5.82	85.74	P2	8.36	81.31	P2	9.03	76.81
36	P2	14.26	100.00	P2	18.69	100.00	P2	23.19	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
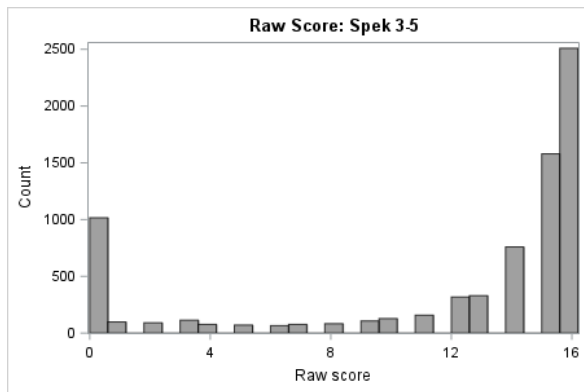
Figure 6.2.3B
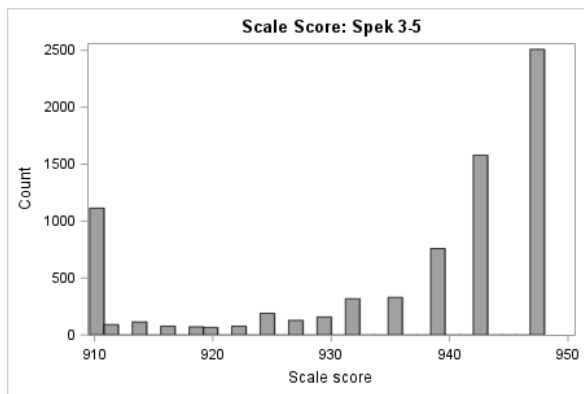
Figure 6.2.3C
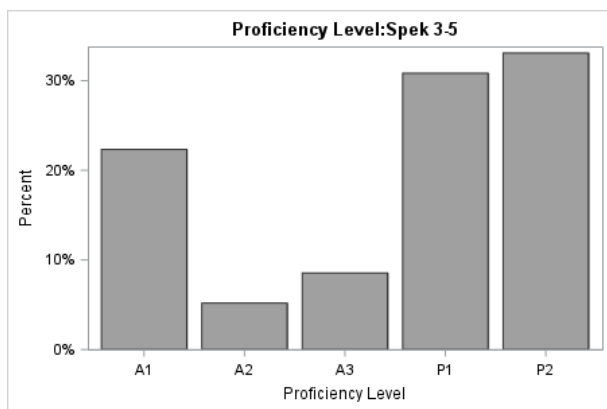
Table 6.2.3A

Raw Score Descriptive Statistics: Spek 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2693	0	16	11.22	5.97
4	2528	0	16	12.10	5.53
5	2360	0	16	12.12	5.64
Total	7581	0	16	11.80	5.74

Table 6.2.3B

Scale Score Descriptive Statistics: Spek 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2693	910	947	934.04	13.91
4	2528	910	947	936.21	13.12
5	2360	910	947	936.40	13.34
Total	7581	910	947	935.50	13.52

Table 6.2.3C

Proficiency Level Distribution: Spek 3-5

Level	Grade 3 Count	Grade 3 Percent	Grade 4 Count	Grade 4 Percent	Grade 5 Count	Grade 5 Percent	Total Count	Total Percent
A1	698	25.92	514	20.33	481	20.38	1693	22.33
A2	146	5.42	129	5.10	118	5.00	393	5.18
A3	264	9.80	202	7.99	182	7.71	648	8.55
P1	834	30.97	821	32.48	683	28.94	2338	30.84
P2	751	27.89	862	34.10	896	37.97	2509	33.10
Total	2693	100	2528	100	2360	100	7581	100

Table 6.2.3D

Equating Summary: Spek 3-5

No equating summary is presented because the Alternate ACCESS Series 503 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 6.2.3D
Test Characteristic Curve: Spek 3-5 S103

Figure 6.2.3E
Test Information Function: Spek 3-5 S103
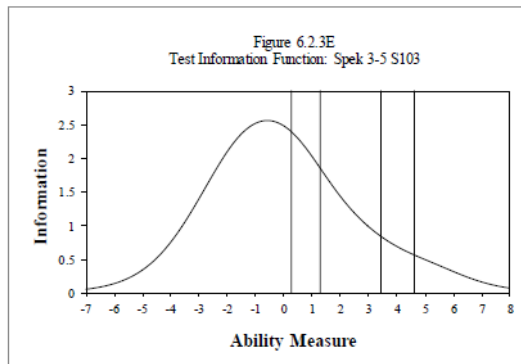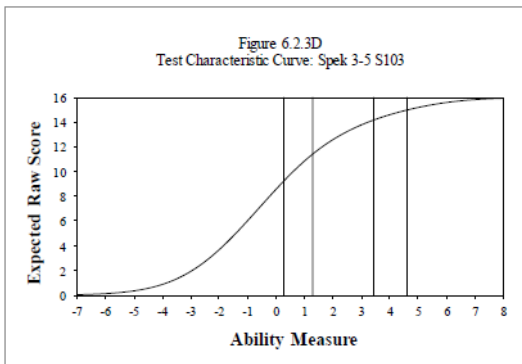
Table 6.2.3E

Reliability: Spek 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
7581	8	0.966	1.065

Table 6.2.3F

Item Analysis Summary: Spek 3-5

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	CR	8	0.00	75.88%	1.04	0.96
DIF Summary		Male/Female			Hispanic/Other	
	DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)	
	AA	5	3	3	5	
	BB	0	0	0	0	
	CC	0	0	0	0	

Table 6.2.3G

Complete Item Analysis: Spek 3-5

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Fit Statistics Infit Mnsq	Outfit Mnsq	Point Biserial	DIF M/F DIF Level	Favored Group	H/O DIF Level	Favored Group
1.S1_A1_103	-2.15	Yes	82.50%	1.57	1.06	0.89	AA	F	AA	O
2.S2_A2_103	-1.31	Yes	80.50%	1.46	1.61	0.90	AA	F	AA	H
3.S3_A3_103	-0.63	Yes	79.00%	1.03	0.70	0.93	AA	M	AA	O
4.S4_A1_103	-0.84	Yes	80.00%	1.00	1.00	0.92	AA	F	AA	O
5.S5_A2_103	-0.08	Yes	78.50%	0.98	0.94	0.93	AA	M	AA	H
6.S6_A3_103	-0.38	Yes	77.50%	0.75	0.50	0.94	AA	F	AA	H
7.S7_P1_103	1.16	Yes	72.50%	0.92	0.67	0.89	AA	M	AA	O
8.S8_P2_103	4.23	Yes	54.00%	0.65	0.93	0.70	AA	M	AA	H

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	15.78	15.78	A1	11.55	11.55	A1	12.67	12.67
1	A1	1.26	17.04	A1	1.54	13.09	A1	1.06	13.73
2	A1	1.45	18.49	A1	1.03	14.12	A1	1.10	14.83
3	A1	1.82	20.31	A1	1.38	15.51	A1	1.27	16.10
4	A1	1.00	21.31	A1	1.15	16.65	A1	0.89	16.99
5	A1	1.11	22.43	A1	0.79	17.44	A1	0.89	17.88
6	A1	0.93	23.36	A1	0.87	18.31	A1	0.81	18.69
7	A1	1.19	24.55	A1	0.99	19.30	A1	0.85	19.53
8	A1	1.37	25.92	A1	1.03	20.33	A1	0.85	20.38
9	A2	1.49	27.40	A2	1.42	21.76	A2	1.31	21.69
10	A2	1.71	29.11	A2	1.78	23.54	A2	1.57	23.26
11	A2	2.23	31.34	A2	1.90	25.44	A2	2.12	25.38
12	A3	4.90	36.24	A3	4.11	29.55	A3	3.47	28.86
13	A3	4.90	41.14	A3	3.88	33.43	A3	4.24	33.09
14	P1	10.17	51.32	P1	10.48	43.91	P1	9.32	42.42
15	P1	20.79	72.11	P1	21.99	65.90	P1	19.62	62.03
16	P2	27.89	100.00	P2	34.10	100.00	P2	37.97	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
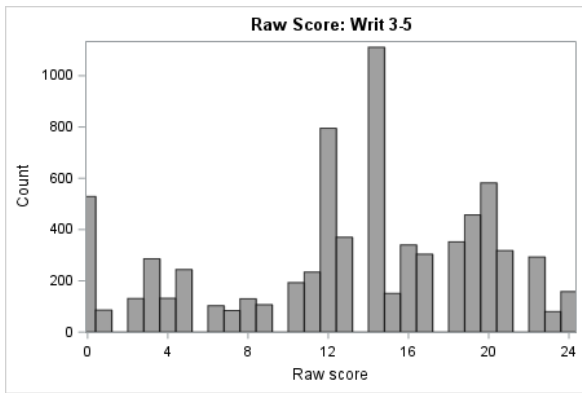
Table 6.2.4A

Raw Score Descriptive Statistics: Writ 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2705	0	24	12.18	6.47
4	2522	0	24	13.40	6.44
5	2343	0	24	13.71	6.62
Total	7570	0	24	13.06	6.54

Figure 6.2.4B
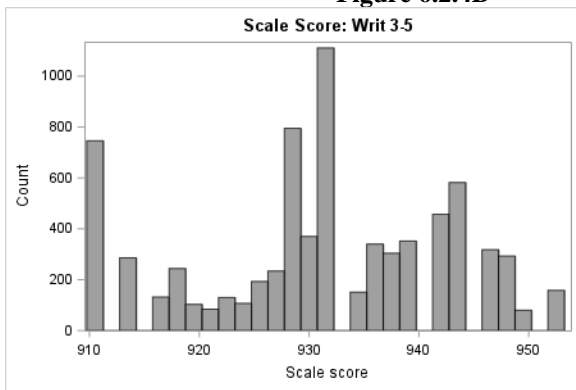
Table 6.2.4B

Scale Score Descriptive Statistics: Writ 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2705	910	953	929.51	11.21
4	2522	910	953	931.75	11.38
5	2343	910	953	932.37	11.74
Total	7570	910	953	931.14	11.50

Figure 6.2.4C
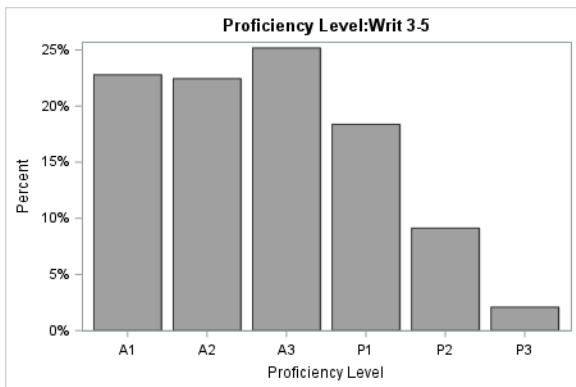
Table 6.2.4C

Proficiency Level Distribution: Writ 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	705	26.06	528	20.94	492	21.00	1725	22.79
A2	676	24.99	552	21.89	471	20.10	1699	22.44
A3	669	24.73	669	26.53	568	24.24	1906	25.18
P1	455	16.82	460	18.24	476	20.32	1391	18.38
P2	166	6.14	254	10.07	271	11.57	691	9.13
P3	34	1.26	59	2.34	65	2.77	158	2.09
Total	2705	100	2522	100	2343	100	7570	100

Table 6.2.4D

Equating Summary: Writ 3-5

No equating summary is presented because the Alternate ACCESS Series 503 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 6.2.4D
Test Characteristic Curve: Writ 3-5 S103

Figure 6.2.4E
Test Information Function: Writ 3-5 S103
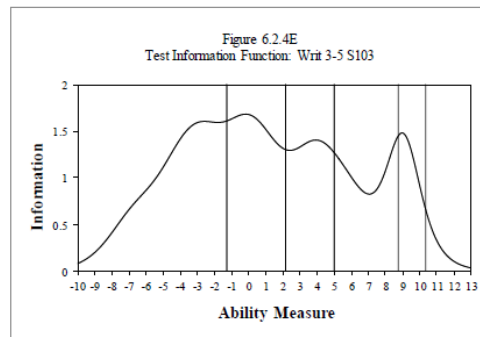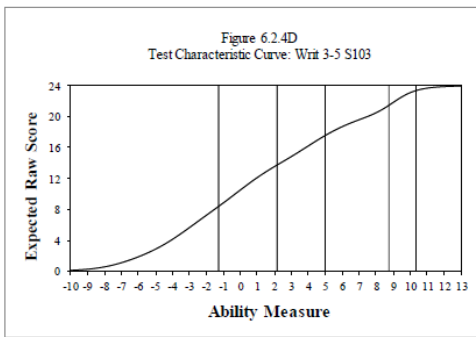
Table 6.2.4E

Reliability: Writ 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
7570	10	0.935	1.662

Table 6.2.4F

Item Analysis Summary: Writ 3-5

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	CR	10	0.00	63.93%	1.31	5.12
DIF Summary	DIF Level	Male/Female		Hispanic/Other	
		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA	4	6	4	6
	BB	0	0	0	0
	CC	0	0	0	0

Table 6.2.4G

Complete Item Analysis: Writ 3-5

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Fit Statistics		Point Biserial	DIF			
							M/F		H/O	
				Infit Mnsq	Outfit Mnsq		DIF Level	Favored Group	DIF Level	Favored Group
1.W1_A1_103	-5.15		85.00%	2.16	9.90	0.71	AA	M	AA	H
2.W2_A2_103	-4.34	Yes	83.50%	1.36	9.90	0.76	AA	F	AA	H
3.W3_A3_103	-1.34	Yes	74.00%	1.52	7.35	0.81	AA	M	AA	H
4.W4_P1_103	1.66	Yes	59.00%	1.12	2.45	0.85	AA	F	AA	O
5.W5_A1_103	-2.59	Yes	79.00%	1.31	8.91	0.79	AA	F	AA	H
6.W6_A2_103	-2.03		78.00%	1.23	5.78	0.80	AA	M	AA	H
7.W7_A3_103	-0.50	Yes	70.00%	1.28	5.37	0.82	AA	M	AA	H
8.W8_P1_103	2.33	Yes	55.50%	0.92	1.32	0.85	AA	F	AA	O
9.W9_P3_103	7.02	Yes	27.00%	0.99	9.90	0.68	AA	M	AA	H
10.W10_P3_103	6.84	Yes	27.50%	1.05	4.09	0.64	AA	F	AA	O

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	8.47	8.47	A1	6.23	6.23	A1	6.10	6.10
1	A1	1.33	9.80	A1	1.11	7.34	A1	0.94	7.04
2	A1	2.07	11.87	A1	1.31	8.64	A1	1.79	8.83
3	A1	4.14	16.01	A1	3.37	12.01	A1	3.80	12.63
4	A1	2.11	18.11	A1	1.67	13.68	A1	1.41	14.04
5	A1	3.14	21.26	A1	3.37	17.05	A1	3.16	17.20
6	A1	1.33	22.59	A1	1.35	18.40	A1	1.41	18.61
7	A1	1.29	23.88	A1	1.11	19.51	A1	0.90	19.50
8	A1	2.18	26.06	A1	1.43	20.94	A1	1.49	21.00
9	A2	1.37	27.43	A2	1.23	22.16	A2	1.66	22.66
10	A2	2.92	30.35	A2	2.97	25.14	A2	1.66	24.33
11	A2	3.29	33.64	A2	3.05	28.19	A2	2.90	27.23
12	A2	12.35	45.99	A2	10.07	38.26	A2	8.83	36.06
13	A2	5.06	51.05	A2	4.56	42.82	A2	5.04	41.10
14	A3	15.45	66.51	A3	15.31	58.13	A3	13.10	54.20
15	A3	2.11	68.61	A3	2.18	60.31	A3	1.66	55.87
16	A3	3.59	72.20	A3	4.84	65.15	A3	5.16	61.03
17	A3	3.59	75.79	A3	4.20	69.35	A3	4.31	65.34
18	P1	5.21	81.00	P1	4.12	73.47	P1	4.57	69.91
19	P1	5.62	86.62	P1	5.91	79.38	P1	6.66	76.57
20	P1	5.99	92.61	P1	8.21	87.59	P1	9.09	85.66
21	P2	2.85	95.45	P2	4.96	92.55	P2	4.95	90.61
22	P2	2.62	98.08	P2	3.93	96.47	P2	5.25	95.86
23	P2	0.67	98.74	P2	1.19	97.66	P2	1.37	97.23
24	P3	1.26	100.00	P3	2.34	100.00	P3	2.77	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

Scale Score: Oral 3-5

Table 6.2.5A

Scale Score Descriptive Statistics: Oral 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2682	910	947	934.44	11.57
4	2516	910	947	936.59	10.86
5	2353	910	947	936.99	10.95
Total	7551	910	947	935.95	11.20

Figure 6.2.5B
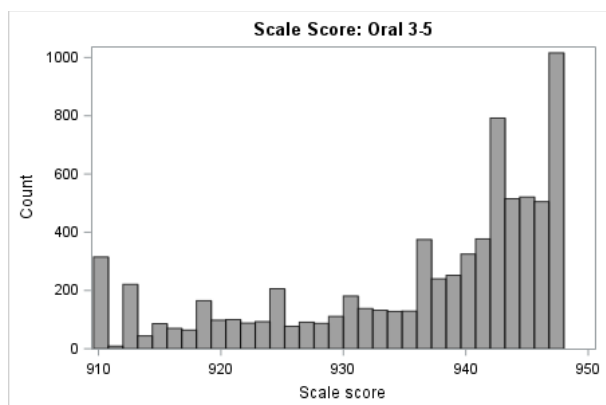
Proficiency Level: Oral 3-5
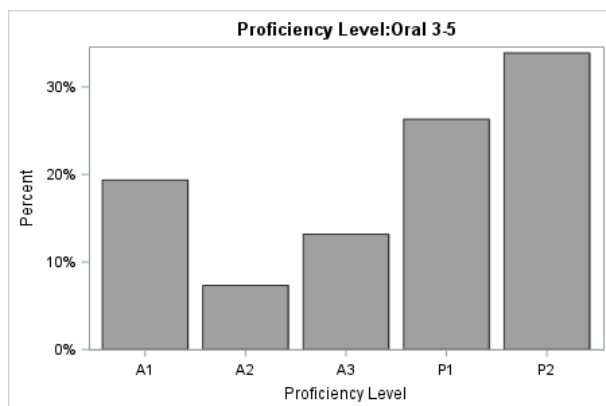
Table 6.2.5C

Proficiency Level Distribution: Oral 3-5

Level	Grade 3 Count	Grade 3 Percent	Grade 4 Count	Grade 4 Percent	Grade 5 Count	Grade 5 Percent	Total Count	Total Percent
A1	614	22.89	445	17.69	402	17.08	1461	19.35
A2	219	8.17	173	6.88	161	6.84	553	7.32
A3	391	14.58	324	12.88	279	11.86	994	13.16
P1	702	26.17	676	26.87	608	25.84	1986	26.30
P2	756	28.19	898	35.69	903	38.38	2557	33.86
Total	2682	100	2516	100	2353	100	7551	100

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	114.436	0.937
Speaking	0.5	182.746	0.966
Oral		125.423	0.973

*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
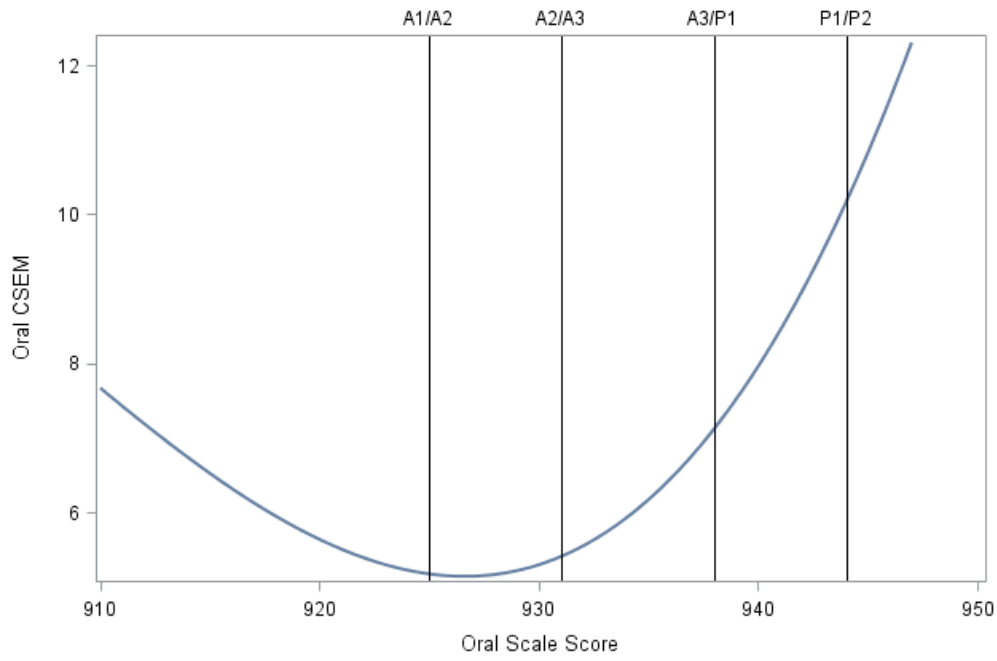
6.2.6 Literacy Composite 3-5

Figure 6.2.6A
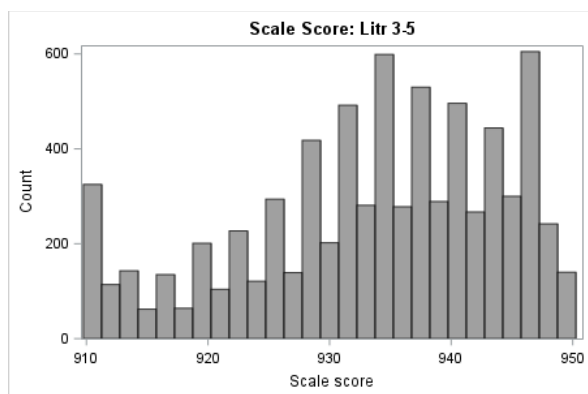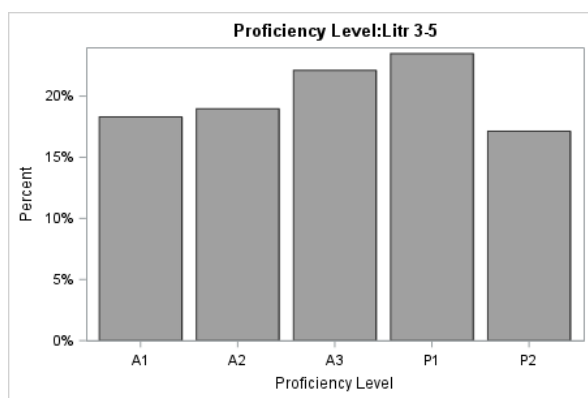
Scale Score: Litr 3-5

Table 6.2.6A

Scale Score Descriptive Statistics: Litr 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2673	910	950	931.73	10.60
4	2506	910	950	934.14	10.21
5	2333	910	950	934.78	10.51
Total	7512	910	950	933.48	10.53

Figure 6.2.6B

Proficiency Level: Litr 3-5

Table 6.2.6C

Proficiency Level Distribution: Litr 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	602	22.52	402	16.04	371	15.90	1375	18.30
A2	558	20.88	472	18.83	395	16.93	1425	18.97
A3	636	23.79	559	22.31	466	19.97	1661	22.11
P1	566	21.17	617	24.62	581	24.90	1764	23.48
P2	311	11.63	456	18.20	520	22.29	1287	17.13
Total	2673	100	2506	100	2333	100	7512	100

Table 6.2.6D
n/a

10

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	120.111	0.947
Writing	0.5	132.229	0.935
Literacy		110.854	0.966

 *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**			**Consistency**
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	
	A1/A2	0.972	0.016	0.012	0.960
	A2/A3	0.951	0.026	0.023	0.931
	A3/P1	0.941	0.016	0.043	0.918
	P1/P2	0.863	0.137	0.000	0.857

Figure 6.2.6F CSEM for Literacy Composite 3-5

6.2.7   Comprehension Composite 3-5

Figure 6.2.7A
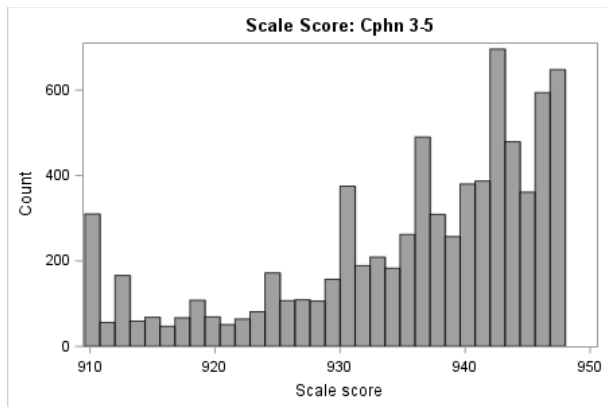
Figure 6.2.7B
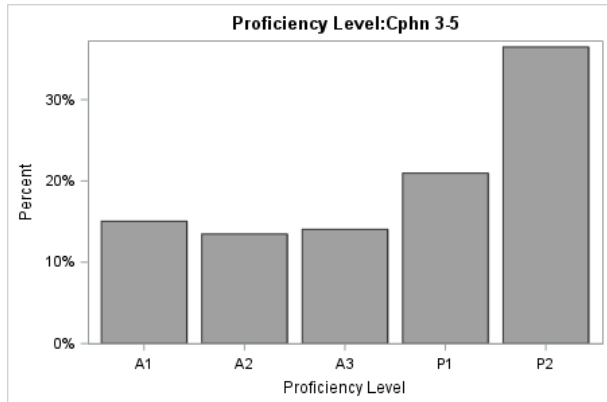
Table 6.2.7A

Scale Score Descriptive Statistics: Cphn 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2703	910	947	933.67	10.91
4	2538	910	947	936.12	10.06
5	2375	910	947	936.74	10.30
Total	7616	910	947	935.44	10.53

Table 6.2.7C

Proficiency Level Distribution: Cphn 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	514	19.83	330	13.19	302	13.07	1146	15.51
A2	421	16.73	335	14.13	270	12.44	1026	14.53
A3	396	14.12	362	14.33	312	14.03	1070	14.16
P1	600	23.59	544	23.67	452	20.92	1596	22.79
P2	772	25.73	967	34.67	1039	39.54	2778	33.01
Total	2703	100	2538	100	2375	100	7616	100

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	114.436	0.937
Reading	0.7	120.111	0.947
Comprehension		110.890	0.966

*Variances from students who had results in all four domains

Table 6.2.7F
n/a
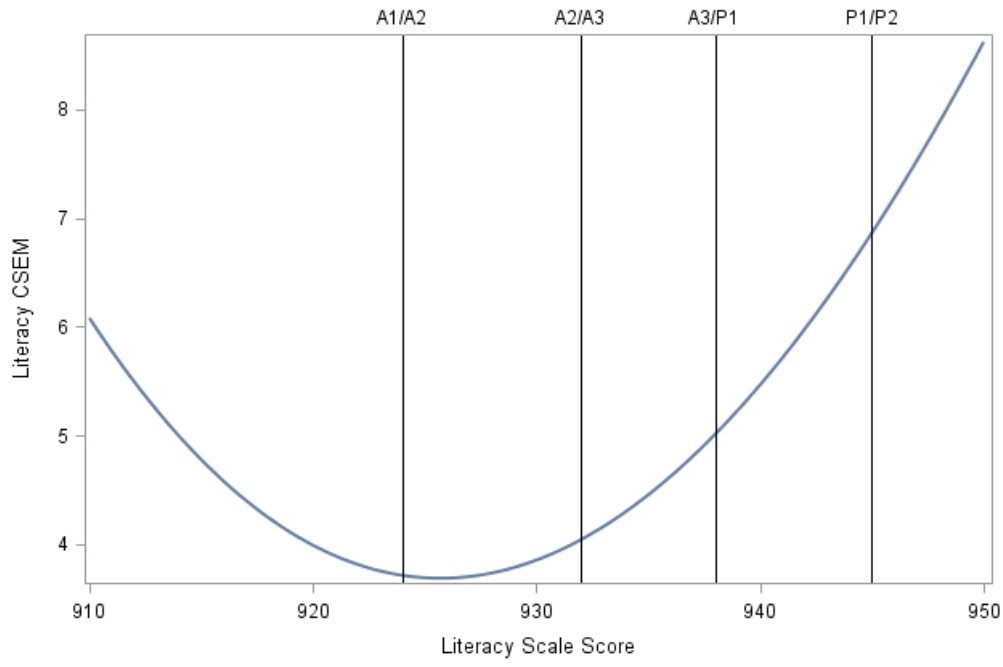
Table 6.2.7G
n/a

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

Scale Score: Over 3-5

Table 6.2.8A

Scale Score Descriptive Statistics: Over 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2645	910	949	932.37	10.51
4	2484	910	949	934.71	10.01
5	2314	910	949	935.31	10.24
Total	7443	910	949	934.07	10.34

Figure 6.2.8B
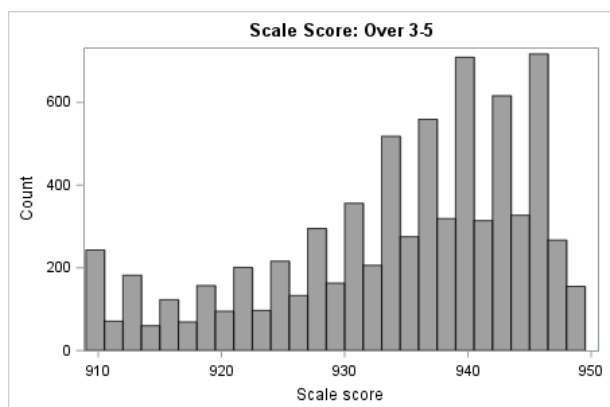
Proficiency Level: Over 3-5
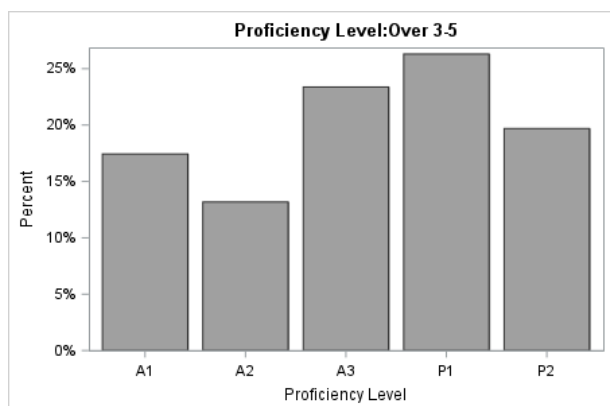
Table 6.2.8C

Proficiency Level Distribution: Over 3-5

Level	Grade 3 Count	Grade 3 Percent	Grade 4 Count	Grade 4 Percent	Grade 5 Count	Grade 5 Percent	Total Count	Total Percent
A1	564	21.32	378	15.22	356	15.38	1298	17.44
A2	391	14.78	320	12.88	270	11.67	981	13.18
A3	662	25.03	585	23.55	493	21.31	1740	23.38
P1	661	24.99	680	27.38	617	26.66	1958	26.31
P2	367	13.88	521	20.97	578	24.98	1466	19.70
Total	2645	100	2484	100	2314	100	7443	100

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	114.436	0.937
Reading	0.35	120.111	0.947
Speaking	0.15	182.746	0.966
Writing	0.35	132.229	0.935
Overall Composite		106.909	0.980

 *Variances from students who had results in all four domains

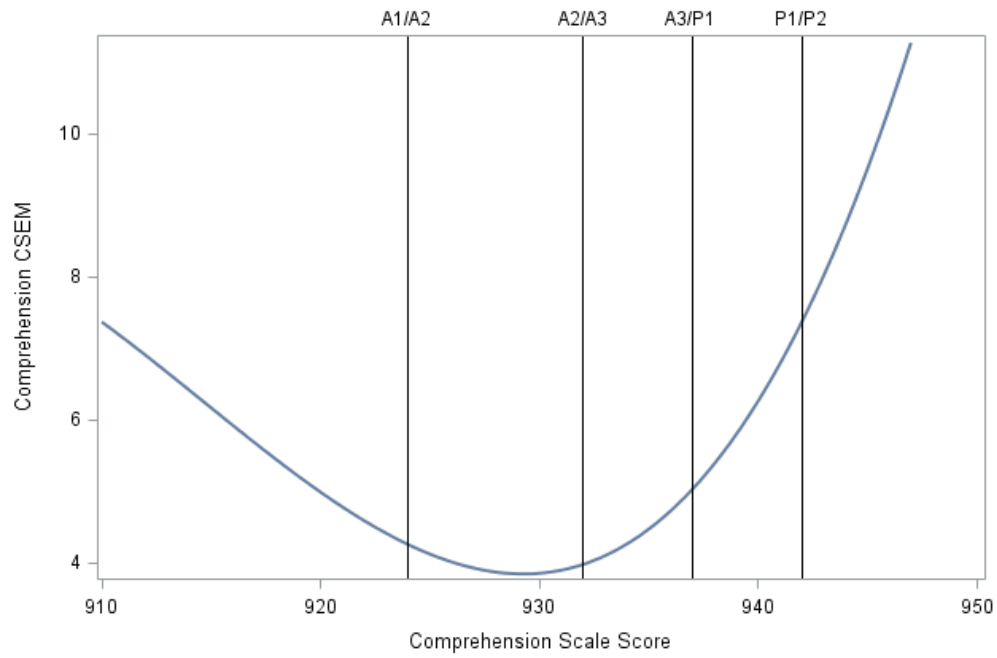Table 6.2.8F
n/a

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			
			False Positives	False Negatives	
	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
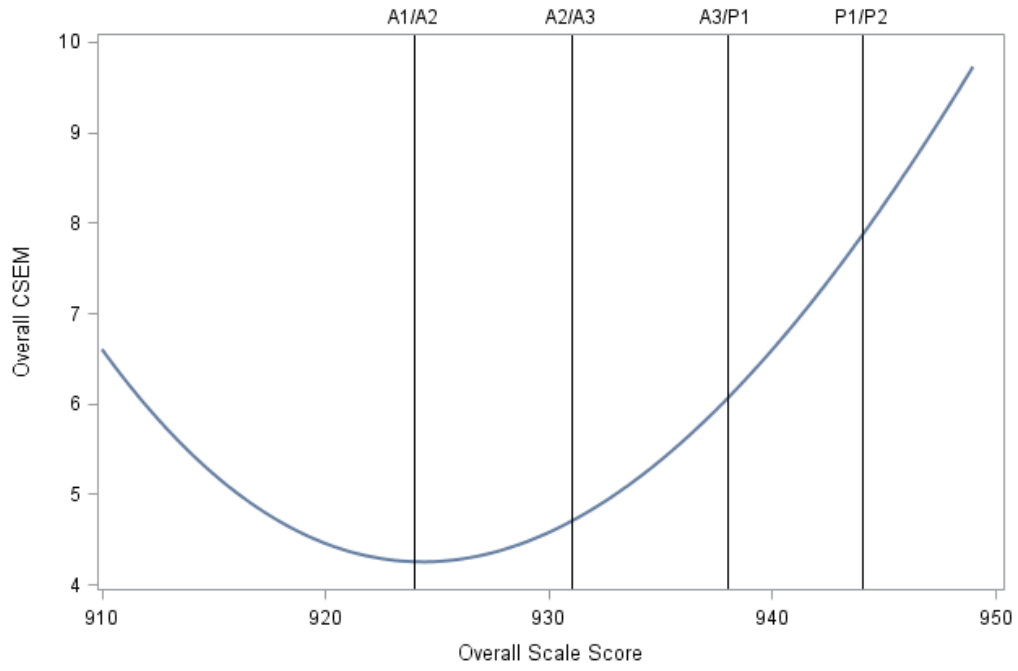
6.3   Grades: 6-8

6.3.1   Listening 6-8

Figure 6.3.1A
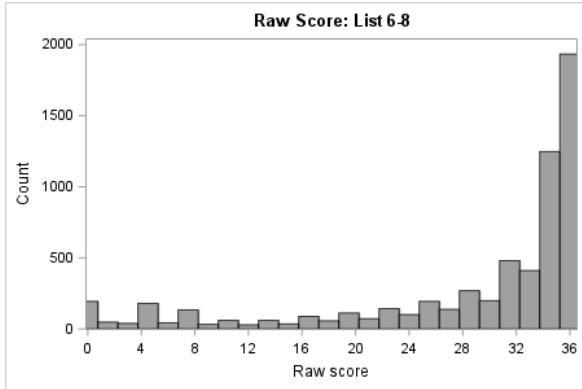
Figure 6.3.1B
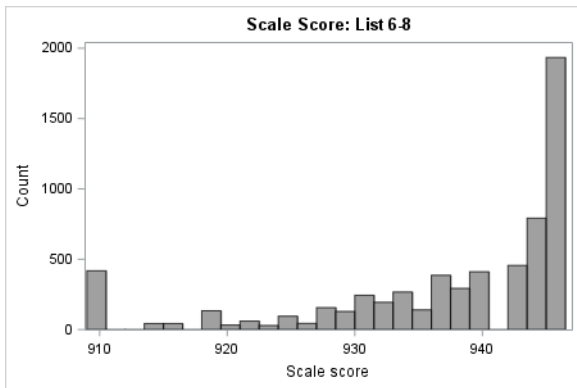
Figure 6.3.1C
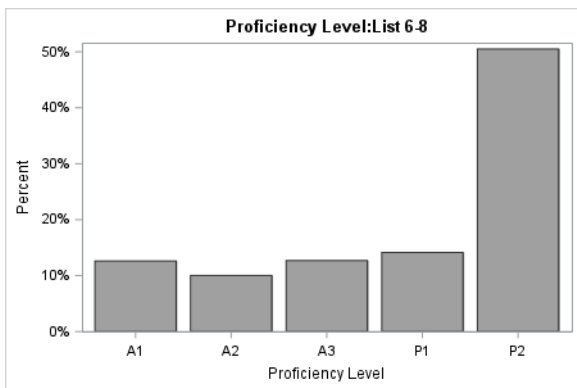
Table 6.3.1A

Raw Score Descriptive Statistics: List 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2164	0	36	28.01	10.54
7	2093	0	36	28.86	10.35
8	2043	0	36	29.12	10.15
Total	6300	0	36	28.65	10.36

Table 6.3.1B

Scale Score Descriptive Statistics: List 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2164	910	946	936.40	10.61
7	2093	910	946	937.42	10.53
8	2043	910	946	937.69	10.34
Total	6300	910	946	937.16	10.51

Table 6.3.1C

Proficiency Level Distribution: List 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	298	13.77	262	12.52	235	11.50	795	12.62
A2	240	11.09	195	9.32	197	9.64	632	10.03
A3	298	13.77	255	12.18	247	12.09	800	12.70
P1	332	15.34	281	13.43	278	13.61	891	14.14
P2	996	46.03	1100	52.56	1086	53.16	3182	50.51
Total	2164	100	2093	100	2043	100	6300	100

Table 6.3.1D

Equating Summary: List 6-8

No equating summary is presented because the Alternate ACCESS Series 503 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.

Figure 6.3.1D
Test Characteristic Curve: List 6-8 S103

Figure 6.3.1E
Test Information Function: List 6-8 S103
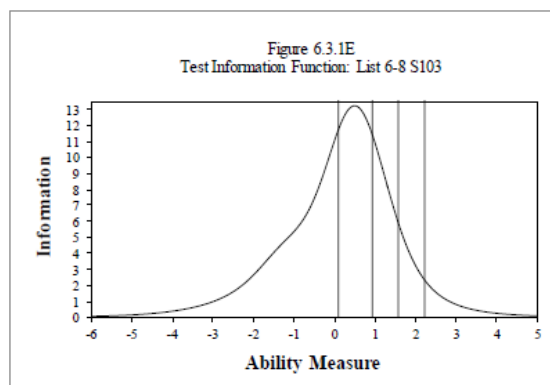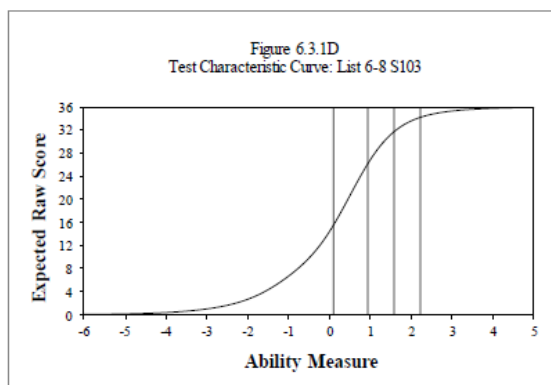
Table 6.3.1E

Reliability: List 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
6300	9	0.945	2.423

Table 6.3.1F

Item Analysis Summary: List 6-8

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	MOSR	9	0.00	81.89%	1.21	1.41
DIF Summary			Male/Female		Hispanic/Other	
	DIF Level		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA		4	5	5	4
	BB		0	0	0	0
	CC		0	0	0	0

Table 6.3.1G

Complete Item Analysis: List 6-8

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Infit Mnsq	Outfit Mnsq	Point Biserial	DIF M/F DIF Level	DIF M/F Favored Group	DIF H/O DIF Level	DIF H/O Favored Group
1. L1_A1_103	-1.53	Yes	91.00%	2.93	3.99	0.64	AA	M	AA	O
2. L2_A2_103	0.17	Yes	82.25%	1.16	0.97	0.85	AA	M	AA	H
3. L3_A2_103	0.05	Yes	83.00%	1.06	0.93	0.87	AA	F	AA	O
4. L4_A3_103	-0.04	Yes	82.50%	1.02	0.89	0.87	AA	M	AA	O
5. L5_A3_103	0.21	Yes	80.75%	0.90	0.84	0.88	AA	M	AA	O
6. L6_P1_103	0.95	Yes	74.50%	1.18	1.13	0.81	AA	F	AA	H
7. L7_P1_103	0.55	Yes	78.75%	0.78	0.70	0.89	AA	F	AA	H
8. L8_P2_103	0.37	Yes	78.00%	0.78	0.62	0.89	AA	F	AA	H
9. L9_P2_103	0.59	Yes	77.75%	0.87	0.70	0.87	AA	M	AA	O

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	3.28	3.28	A1	3.01	3.01	A1	2.94	2.94
1	A1	0.18	3.47	A1	0.19	3.20	A1	0.29	3.23
2	A1	0.60	4.07	A1	0.57	3.77	A1	0.44	3.67
3	A1	0.88	4.94	A1	0.62	4.40	A1	0.34	4.01
4	A1	2.22	7.16	A1	2.20	6.59	A1	2.10	6.12
5	A1	0.65	7.81	A1	0.57	7.17	A1	0.83	6.95
6	A1	0.79	8.60	A1	0.76	7.93	A1	0.49	7.44
7	A1	1.29	9.89	A1	1.86	9.79	A1	1.57	9.01
8	A1	0.79	10.67	A1	0.38	10.18	A1	0.49	9.50
9	A1	0.42	11.09	A1	0.62	10.80	A1	0.49	9.99
10	A1	0.46	11.55	A1	0.53	11.32	A1	0.34	10.33
11	A1	0.69	12.25	A1	0.33	11.66	A1	0.49	10.82
12	A1	0.69	12.94	A1	0.38	12.04	A1	0.29	11.11
13	A1	0.83	13.77	A1	0.48	12.52	A1	0.39	11.50
14	A2	0.46	14.23	A2	0.29	12.80	A2	0.39	11.89
15	A2	0.51	14.74	A2	0.62	13.43	A2	0.54	12.43
16	A2	1.02	15.76	A2	0.43	13.86	A2	0.64	13.07
17	A2	0.60	16.36	A2	0.62	14.48	A2	0.88	13.95
18	A2	0.97	17.33	A2	1.00	15.48	A2	0.73	14.68
19	A2	1.20	18.53	A2	0.57	16.05	A2	0.83	15.52
20	A2	0.88	19.41	A2	0.96	17.01	A2	0.88	16.40
21	A2	1.20	20.61	A2	1.15	18.16	A2	1.08	17.47
22	A2	1.02	21.63	A2	1.15	19.30	A2	0.88	18.36
23	A2	1.52	23.15	A2	1.05	20.35	A2	1.17	19.53
24	A2	1.71	24.86	A2	1.48	21.83	A2	1.62	21.15
25	A3	1.62	26.48	A3	1.53	23.36	A3	1.08	22.22
26	A3	1.94	28.42	A3	1.34	24.70	A3	1.71	23.94
27	A3	2.36	30.78	A3	2.10	26.80	A3	2.10	26.04
28	A3	2.36	33.13	A3	1.67	28.48	A3	2.10	28.14
29	A3	2.26	35.40	A3	2.15	30.63	A3	2.25	30.40
30	A3	3.23	38.63	A3	3.39	34.02	A3	2.84	33.24
31	P1	3.47	42.10	P1	3.20	37.22	P1	2.20	35.44
32	P1	5.36	47.46	P1	4.01	41.23	P1	4.55	39.99
33	P1	6.52	53.97	P1	6.21	47.44	P1	6.85	46.84
34	P2	7.16	61.14	P2	7.45	54.90	P2	7.10	53.94
35	P2	12.62	73.75	P2	12.71	67.61	P2	12.38	66.32
36	P2	26.25	100.00	P2	32.39	100.00	P2	33.68	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
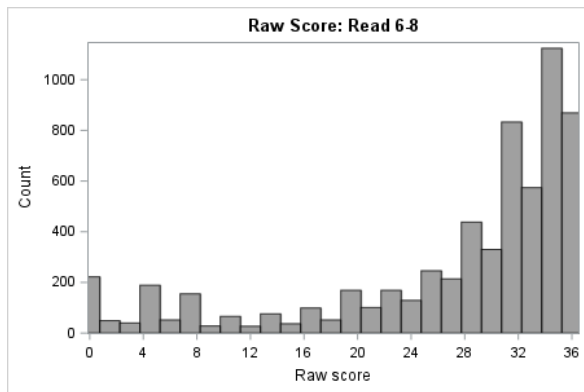
Figure 6.3.2B
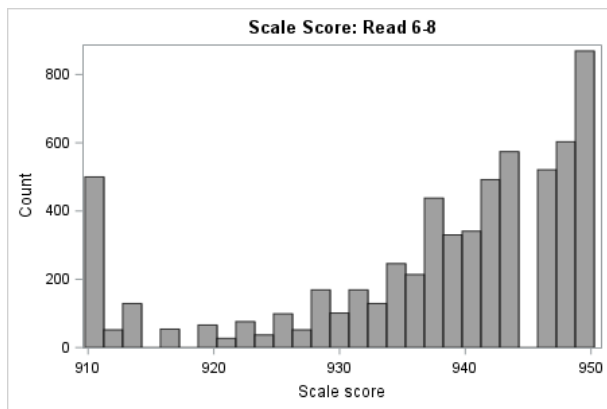
Figure 6.3.2C
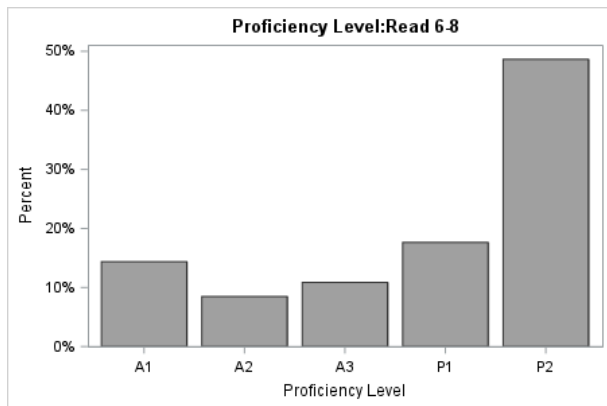
Table 6.3.2A

Raw Score Descriptive Statistics: Read 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2159	0	36	26.50	10.35
7	2086	0	36	27.26	10.12
8	2043	0	36	27.38	10.27
Total	6288	0	36	27.04	10.26

Table 6.3.2B

Scale Score Descriptive Statistics: Read 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2159	910	950	936.74	11.86
7	2086	910	950	937.74	11.77
8	2043	910	950	937.97	11.98
Total	6288	910	950	937.47	11.88

Table 6.3.2C

Proficiency Level Distribution: Read 6-8

Level	Grade 6 Count	Grade 6 Percent	Grade 7 Count	Grade 7 Percent	Grade 8 Count	Grade 8 Percent	Total Count	Total Percent
A1	327	15.15	288	13.81	289	14.15	904	14.38
A2	194	8.99	168	8.05	171	8.37	533	8.48
A3	246	11.39	243	11.65	194	9.50	683	10.86
P1	422	19.55	367	17.59	320	15.66	1109	17.64
P2	970	44.93	1020	48.90	1069	52.33	3059	48.65
Total	2159	100	2086	100	2043	100	6288	100

Table 6.3.2D

Equating Summary: Read 6-8

No equating summary is presented because the Alternate ACCESS Series 503 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 6.3.2D
Test Characteristic Curve: Read 6-8 S103

Figure 6.3.2E
Test Information Function: Read 6-8 S103
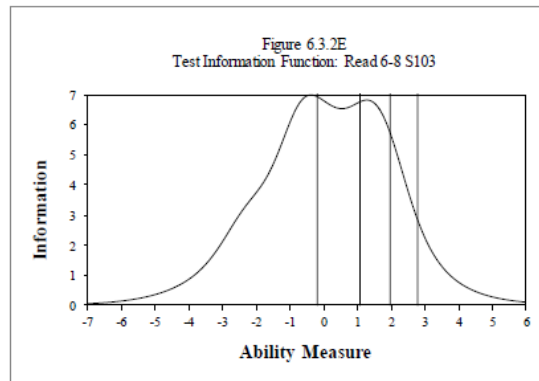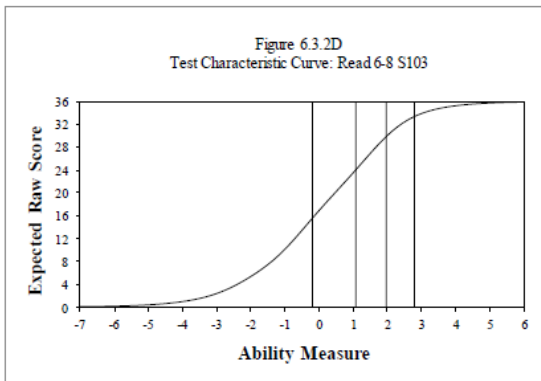
Table 6.3.2E

Reliability: Read 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
6288	9	0.941	2.496

Table 6.3.2F

Item Analysis Summary: Read 6-8

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	MOSR	9	0.00	76.22%	1.23	1.46
DIF Summary			Male/Female		Hispanic/Other	
	DIF Level		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA		4	5	4	5
	BB		0	0	0	0
	CC		0	0	0	0

Table 6.3.2G

Complete Item Analysis: Read 6-8

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Infit Mnsq	Outfit Mnsq	Point Biserial	DIF M/F DIF Level	Favored Group	DIF H/O DIF Level	Favored Group
1.R1_A1_103	-2.50		91.00%	3.11	3.15	0.66	AA	F	AA	O
2.R2_A2_103	-0.76	Yes	84.50%	1.32	0.73	0.87	AA	M	AA	O
3.R3_A2_103	-0.93	Yes	84.50%	1.12	0.71	0.87	AA	F	AA	O
4.R4_A3_103	0.07	Yes	77.75%	1.18	0.92	0.87	AA	M	AA	O
5.R5_A3_103	-0.72	Yes	83.75%	0.90	0.52	0.89	AA	F	AA	O
6.R6_P1_103	1.17	Yes	69.50%	0.95	0.86	0.83	AA	F	AA	H
7.R7_P1_103	1.09	Yes	70.00%	0.96	0.91	0.83	AA	F	AA	H
8.R8_P2_103	1.66	Yes	63.00%	0.89	0.98	0.79	AA	F	AA	H
9.R9_P2_103	1.81	Yes	61.25%	0.92	0.95	0.76	AA	M	AA	H

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	3.94	3.94	A1	3.16	3.16	A1	3.48	3.48
1	A1	0.46	4.40	A1	0.29	3.45	A1	0.34	3.82
2	A1	0.46	4.86	A1	0.53	3.98	A1	0.24	4.06
3	A1	0.79	5.65	A1	0.62	4.60	A1	0.49	4.55
4	A1	2.08	7.74	A1	2.68	7.29	A1	2.59	7.15
5	A1	0.65	8.38	A1	0.43	7.72	A1	0.59	7.73
6	A1	1.16	9.54	A1	0.62	8.34	A1	0.69	8.42
7	A1	1.81	11.35	A1	2.01	10.35	A1	2.35	10.77
8	A1	0.56	11.90	A1	0.38	10.74	A1	0.29	11.06
9	A1	0.56	12.46	A1	0.43	11.17	A1	0.34	11.40
10	A1	0.32	12.78	A1	0.67	11.84	A1	0.24	11.65
11	A1	0.83	13.62	A1	0.48	12.32	A1	0.59	12.24
12	A1	0.51	14.13	A1	0.43	12.75	A1	0.34	12.58
13	A1	0.60	14.73	A1	0.34	13.09	A1	0.88	13.46
14	A1	0.42	15.15	A1	0.72	13.81	A1	0.69	14.15
15	A2	0.69	15.84	A2	0.62	14.43	A2	0.44	14.59
16	A2	0.74	16.58	A2	0.72	15.15	A2	0.93	15.52
17	A2	0.79	17.37	A2	0.77	15.92	A2	0.78	16.30
18	A2	0.83	18.20	A2	0.81	16.73	A2	0.83	17.13
19	A2	1.11	19.31	A2	0.77	17.50	A2	1.17	18.31
20	A2	1.99	21.31	A2	1.53	19.03	A2	1.47	19.77
21	A2	1.71	23.02	A2	1.53	20.57	A2	1.57	21.34
22	A2	1.11	24.13	A2	1.29	21.86	A2	1.17	22.52
23	A3	1.85	25.98	A3	1.20	23.06	A3	1.42	23.94
24	A3	1.99	27.98	A3	2.16	25.22	A3	2.01	25.94
25	A3	1.67	29.64	A3	2.21	27.42	A3	1.47	27.41
26	A3	2.36	32.01	A3	2.21	29.63	A3	1.81	29.22
27	A3	3.52	35.53	A3	3.88	33.51	A3	2.79	32.01
28	P1	3.57	39.09	P1	2.73	36.24	P1	2.89	34.90
29	P1	4.21	43.31	P1	3.69	39.93	P1	3.77	38.67
30	P1	5.65	48.96	P1	5.56	45.49	P1	4.50	43.17
31	P1	6.11	55.07	P1	5.61	51.10	P1	4.50	47.67
32	P2	8.52	63.59	P2	6.81	57.91	P2	8.13	55.80
33	P2	9.08	72.67	P2	9.54	67.45	P2	8.76	64.56
34	P2	8.15	80.82	P2	7.81	75.26	P2	8.91	73.47
35	P2	8.06	88.88	P2	10.55	85.81	P2	10.23	83.70
36	P2	11.12	100.00	P2	14.19	100.00	P2	16.30	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
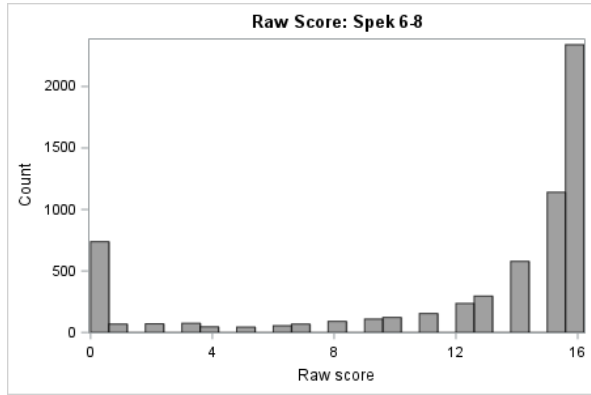
Figure 6.3.3B
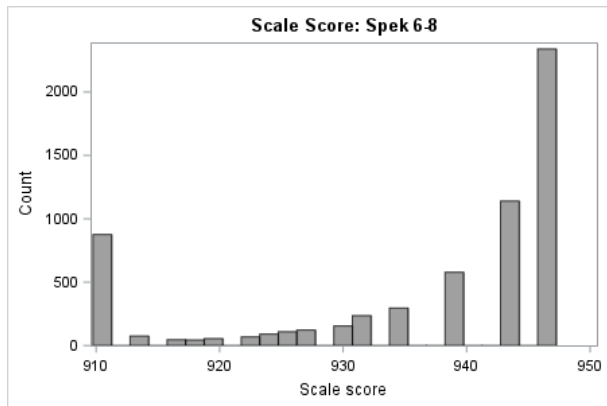
Figure 6.3.3C
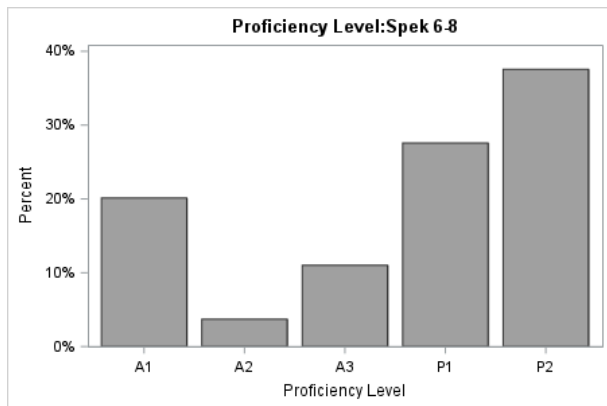
Table 6.3.3A

Raw Score Descriptive Statistics: Spek 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2141	0	16	12.01	5.52
7	2065	0	16	12.32	5.40
8	2023	0	16	12.07	5.63
Total	6229	0	16	12.13	5.52

Table 6.3.3B

Scale Score Descriptive Statistics: Spek 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2141	910	947	935.96	13.09
7	2065	910	947	936.76	12.92
8	2023	910	947	936.27	13.38
Total	6229	910	947	936.32	13.13

Table 6.3.3C

Proficiency Level Distribution: Spek 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	443	20.69	380	18.40	431	21.30	1254	20.13
A2	84	3.92	91	4.41	56	2.77	231	3.71
A3	256	11.96	225	10.90	205	10.13	686	11.01
P1	611	28.54	563	27.26	544	26.89	1718	27.58
P2	747	34.89	806	39.03	787	38.90	2340	37.57
Total	2141	100	2065	100	2023	100	6229	100

Table 6.3.3D

Equating Summary: Spek 6-8

No equating summary is presented because the Alternate ACCESS Series 503 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 6.3.3D
Test Characteristic Curve: Spek 6-8 S103

Figure 6.3.3E
Test Information Function: Spek 6-8 S103
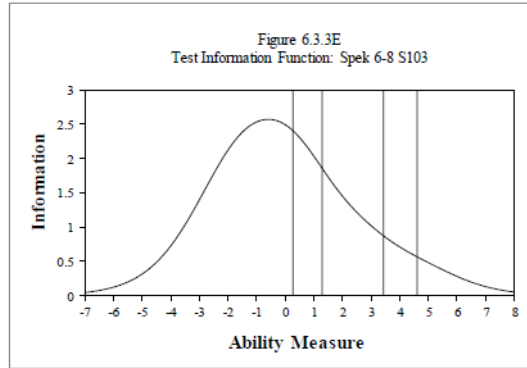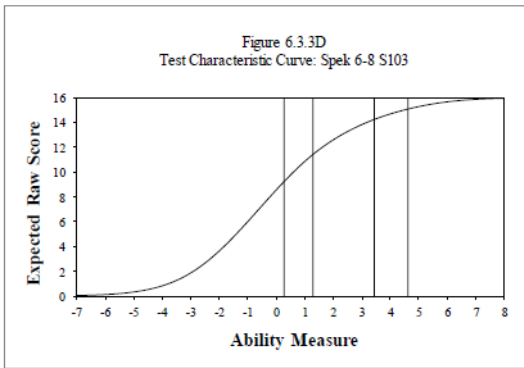
Table 6.3.3E

Reliability: Spek 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
6229	8	0.965	1.030

Table 6.3.3F

Item Analysis Summary: Spek 6-8

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	CR	8	0.00	77.19%	1.02	0.96
DIF Summary			Male/Female		Hispanic/Other	
	DIF Level		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA		5	3	3	5
	BB		0	0	0	0
	CC		0	0	0	0

Table 6.3.3G

Complete Item Analysis: Spek 6-8

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Fit Statistics Infit Mnsq	Outfit Mnsq	Point Biserial	DIF M/F DIF Level	Favored Group	DIF H/O DIF Level	Favored Group
1.S1_A1_103	-1.91	Yes	84.00%	1.44	1.15	0.89	AA	F	AA	H
2.S2_A2_103	-1.23		82.50%	1.32	1.24	0.90	AA	M	AA	O
3.S3_A3_103	-0.52	Yes	80.50%	1.03	0.98	0.92	AA	F	AA	O
4.S4_A1_103	-0.62		81.00%	1.05	0.91	0.92	AA	F	AA	H
5.S5_A2_103	-0.71	Yes	81.50%	0.85	0.77	0.93	AA	F	AA	O
6.S6_A3_103	-0.47	Yes	79.50%	0.71	0.52	0.94	AA	F	AA	O
7.S7_P1_103	1.47	Yes	73.50%	0.91	0.76	0.87	AA	M	AA	O
8.S8_P2_103	3.95	Yes	58.50%	0.75	1.03	0.70	AA	M	AA	H

.

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	12.38	12.38	A1	10.75	10.75	A1	12.41	12.41
1	A1	1.03	13.40	A1	1.26	12.01	A1	0.99	13.40
2	A1	0.93	14.34	A1	1.16	13.17	A1	1.29	14.68
3	A1	0.89	15.23	A1	1.26	14.43	A1	1.48	16.16
4	A1	0.65	15.88	A1	0.73	15.16	A1	0.84	17.00
5	A1	0.61	16.49	A1	0.82	15.98	A1	0.69	17.70
6	A1	0.98	17.47	A1	0.82	16.80	A1	0.84	18.54
7	A1	1.35	18.82	A1	0.63	17.43	A1	1.29	19.82
8	A1	1.87	20.69	A1	0.97	18.40	A1	1.48	21.30
9	A2	1.91	22.61	A2	2.03	20.44	A2	1.29	22.59
10	A2	2.01	24.61	A2	2.37	22.81	A2	1.48	24.07
11	A3	2.80	27.42	A3	1.89	24.70	A3	2.72	26.79
12	A3	4.25	31.67	A3	3.87	28.57	A3	3.21	30.00
13	A3	4.90	36.57	A3	5.13	33.70	A3	4.20	34.21
14	P1	9.34	45.91	P1	9.20	42.91	P1	9.29	43.50
15	P1	19.20	65.11	P1	18.06	60.97	P1	17.60	61.10
16	P2	34.89	100.00	P2	39.03	100.00	P2	38.90	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			Consistency
	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
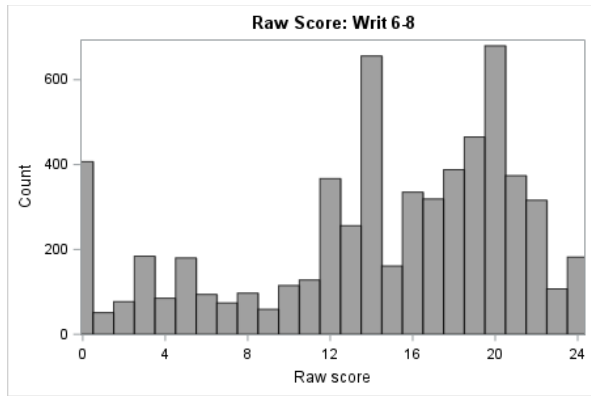
Figure 6.3.4B
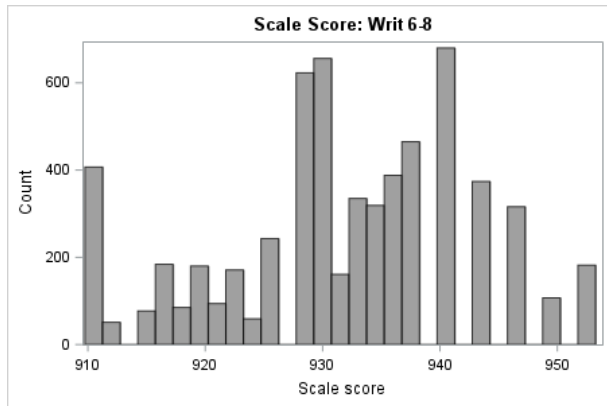
Figure 6.3.4C
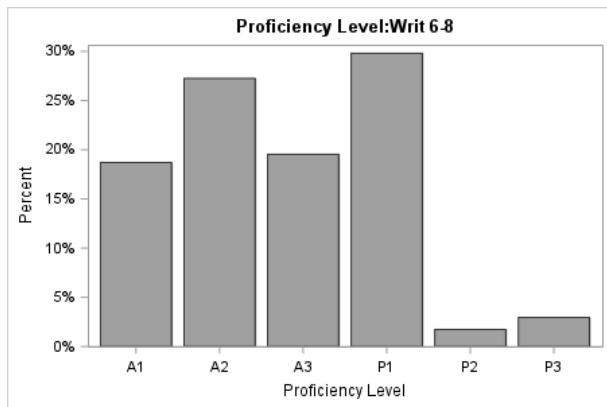
Table 6.3.4A

Raw Score Descriptive Statistics: Writ 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2126	0	24	13.94	6.61
7	2032	0	24	14.45	6.73
8	1999	0	24	14.56	6.73
Total	6157	0	24	14.31	6.69

Table 6.3.4B

Scale Score Descriptive Statistics: Writ 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2126	910	953	931.31	10.25
7	2032	910	953	932.25	10.54
8	1999	910	953	932.39	10.57
Total	6157	910	953	931.97	10.46

Table 6.3.4C

Proficiency Level Distribution: Writ 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	408	19.19	381	18.75	363	18.16	1152	18.71
A2	646	30.39	520	25.59	512	25.61	1678	27.25
A3	417	19.61	391	19.24	395	19.76	1203	19.54
P1	568	26.72	639	31.45	628	31.42	1835	29.80
P2	32	1.51	40	1.97	35	1.75	107	1.74
P3	55	2.59	61	3.00	66	3.30	182	2.96
Total	2126	100	2032	100	1999	100	6157	100

Table 6.3.4D

Equating Summary: Writ 6-8

No equating summary is presented because the Alternate ACCESS Series 503 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 6.3.4D
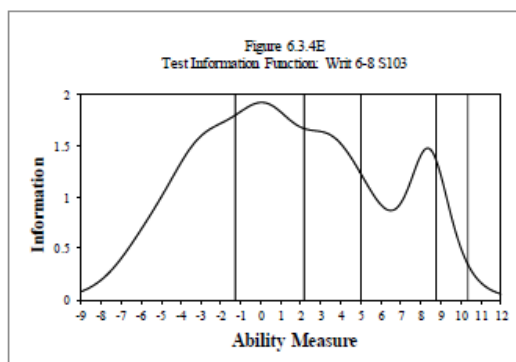Test Characteristic Curve: Writ 6-8 S103

Figure 6.3.4E
Test Information Function: Writ 6-8 S103
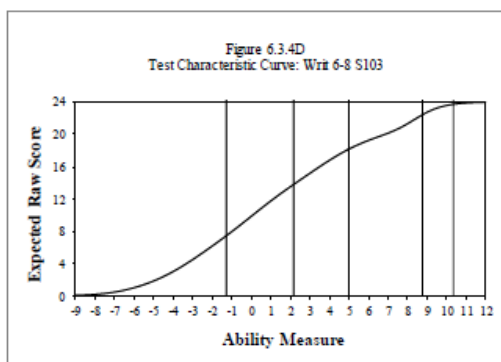
Table 6.3.4E

Reliability: Writ 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
6157	10	0.940	1.633

Table 6.3.4F

Item Analysis Summary: Writ 6-8

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	CR	10	0.00	67.23%	1.51	5.89
DIF Summary			Male/Female		Hispanic/Other	
	DIF Level		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA		6	4	6	4
	BB		0	0	0	0
	CC		0	0	0	0

Table 6.3.4G

Complete Item Analysis: Writ 6-8

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Infit Mnsq	Outfit Mnsq	Point Biserial	DIF M/F DIF Level	DIF M/F Favored Group	DIF H/O DIF Level	DIF H/O Favored Group
1.W1_A1_103	-4.12		85.00%	2.57	9.90	0.74	AA	F	AA	H
2.W2_A2_103	-2.91		83.00%	1.67	9.90	0.80	AA	F	AA	H
3.W3_A3_103	-1.32	Yes	78.50%	1.81	9.90	0.81	AA	F	AA	H
4.W4_P1_103	1.61	Yes	65.50%	1.34	2.05	0.84	AA	M	AA	O
5.W5_A1_103	-1.50		79.50%	1.44	7.30	0.82	AA	M	AA	H
6.W6_A2_103	-1.35		79.00%	1.33	3.08	0.83	AA	M	AA	H
7.W7_A3_103	-0.13		74.50%	1.43	6.94	0.84	AA	M	AA	O
8.W8_P1_103	1.83	Yes	64.50%	1.01	1.16	0.87	AA	M	AA	O
9.W9_P3_103	6.46	Yes	34.00%	1.04	3.94	0.71	AA	F	AA	O
10.W10_P3_103	6.16	Yes	34.75%	1.03	5.07	0.65	AA	F	AA	H

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	6.87	6.87	A1	6.45	6.45	A1	6.50	6.50
1	A1	0.85	7.71	A1	0.74	7.19	A1	0.90	7.40
2	A1	1.41	9.13	A1	1.33	8.51	A1	1.00	8.40
3	A1	2.92	12.04	A1	2.90	11.42	A1	3.15	11.56
4	A1	1.41	13.45	A1	1.38	12.80	A1	1.35	12.91
5	A1	2.73	16.18	A1	3.35	16.14	A1	2.70	15.61
6	A1	1.65	17.83	A1	1.53	17.67	A1	1.40	17.01
7	A1	1.36	19.19	A1	1.08	18.75	A1	1.15	18.16
8	A2	1.74	20.93	A2	0.94	19.69	A2	2.05	20.21
9	A2	0.80	21.73	A2	1.08	20.77	A2	1.00	21.21
10	A2	2.30	24.04	A2	2.02	22.79	A2	1.25	22.46
11	A2	2.26	26.29	A2	2.17	24.95	A2	1.80	24.26
12	A2	6.68	32.97	A2	6.05	31.00	A2	5.10	29.36
13	A2	4.37	37.35	A2	4.43	35.43	A2	3.65	33.02
14	A2	12.23	49.58	A2	8.91	44.34	A2	10.76	43.77
15	A3	2.63	52.21	A3	2.51	46.85	A3	2.70	46.47
16	A3	5.88	58.09	A3	5.61	52.46	A3	4.80	51.28
17	A3	4.61	62.70	A3	5.61	58.07	A3	5.35	56.63
18	A3	6.49	69.19	A3	5.51	63.58	A3	6.90	63.53
19	P1	8.23	77.42	P1	7.28	70.87	P1	7.10	70.64
20	P1	8.98	86.41	P1	12.30	83.17	P1	11.96	82.59
21	P1	4.84	91.25	P1	6.50	89.67	P1	6.95	89.54
22	P1	4.66	95.91	P1	5.36	95.03	P1	5.40	94.95
23	P2	1.51	97.41	P2	1.97	97.00	P2	1.75	96.70
24	P3	2.59	100.00	P3	3.00	100.00	P3	3.30	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

Scale Score: Oral 6-8

Figure 6.3.5B
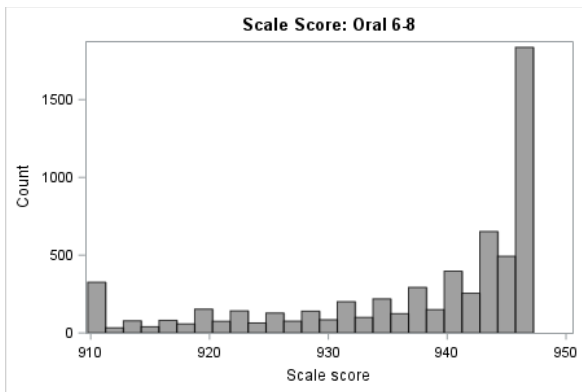
Proficiency Level: Oral 6-8
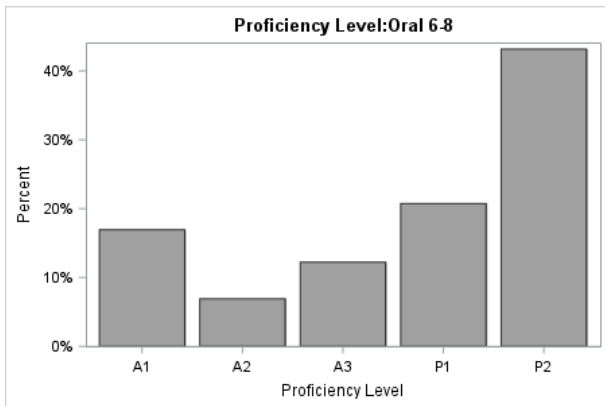
Table 6.3.5A

Scale Score Descriptive Statistics: Oral 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2136	910	947	936.60	11.22
7	2053	910	947	937.59	11.08
8	2010	910	947	937.44	11.27
Total	6199	910	947	937.20	11.20

Table 6.3.5C

Proficiency Level Distribution: Oral 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	386	18.07	327	15.93	337	16.77	1050	16.94
A2	158	7.40	131	6.38	140	6.97	429	6.92
A3	278	13.01	256	12.47	224	11.14	758	12.23
P1	470	22.00	417	20.31	399	19.85	1286	20.75
P2	844	39.51	922	44.91	910	45.27	2676	43.17
Total	2136	100	2053	100	2010	100	6199	100

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	110.469	0.945
Speaking	0.5	172.474	0.965
Oral		125.375	0.976

*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
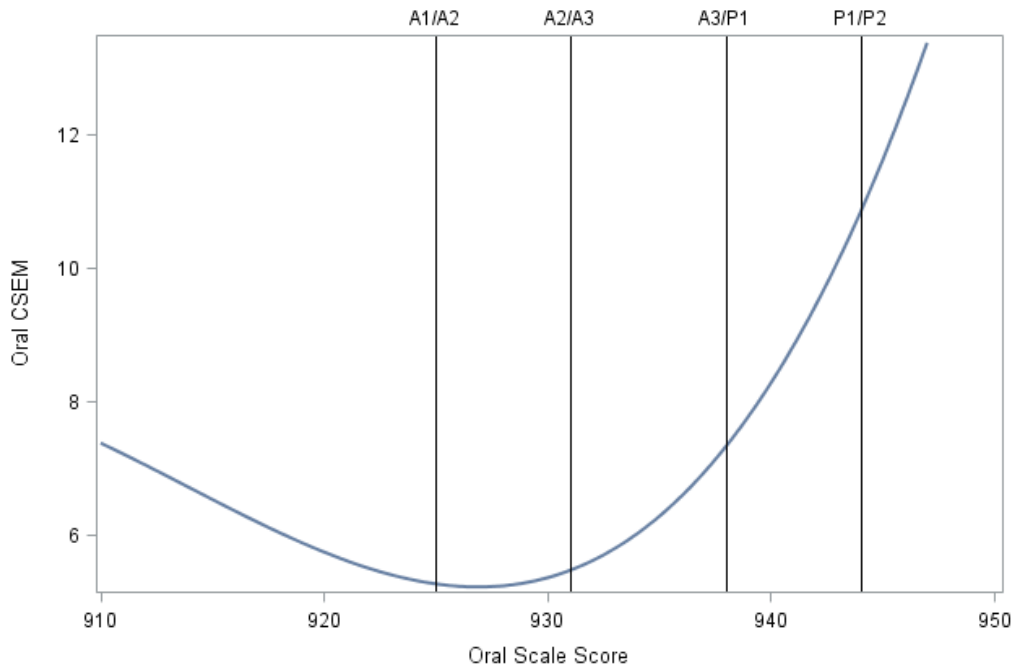
6.3.6   Literacy Composite 6-8

Figure 6.3.6A
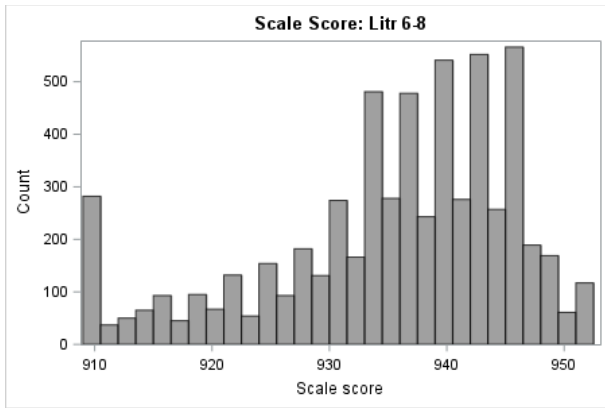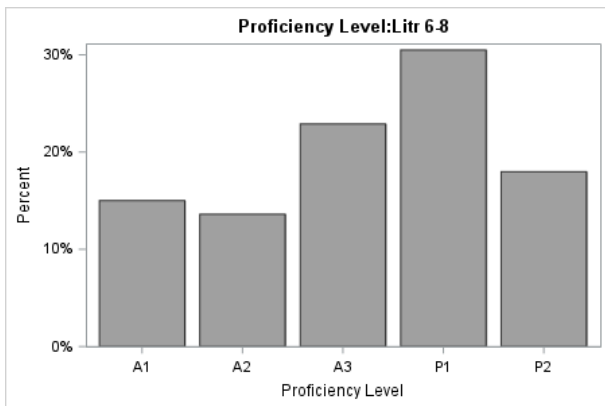
Table 6.3.6A

Scale Score Descriptive Statistics: Litr 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2117	910	952	934.31	10.26
7	2021	910	952	935.37	10.33
8	1990	910	952	935.51	10.62
Total	6128	910	952	935.05	10.42

Figure 6.3.6B

Table 6.3.6C

Proficiency Level Distribution: Litr 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	342	16.15	284	14.05	294	14.77	920	15.01
A2	302	14.27	281	13.90	251	12.61	834	13.61
A3	535	25.27	462	22.86	406	20.40	1403	22.89
P1	645	30.47	587	29.05	637	32.01	1869	30.50
P3	293	13.84	407	20.14	402	20.20	1102	17.98
Total	2117	100	2021	100	1990	100	6128	100

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	141.054	0.941
Writing	0.5	109.363	0.940
Literacy		108.483	0.966

*Variances from students who had results in all four domains

Table 6.3.6F
n/a

Table 6.3.6G
n/a

Table 6.3.6H
n/a

Table 6.3.6I
n/a

Table 6.3.6J

Accuracy and Consistency of Classification Indices: Litr 6-8

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.638	0.616		0.503	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	A1	0.922		0.883	
	A2	0.740		0.636	
	A3	0.794		0.696	
	P1	0.499		0.503	
	P2	-		0.557	
Indices at Cut Points		**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
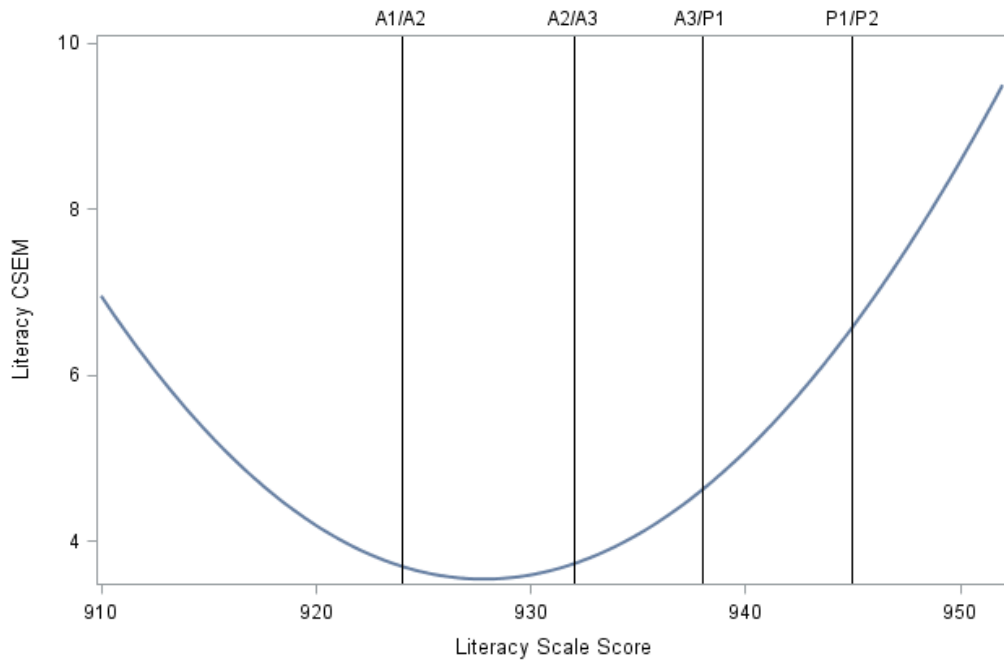
6.3.7 Comprehension Composite 6-8

Figure 6.3.7A
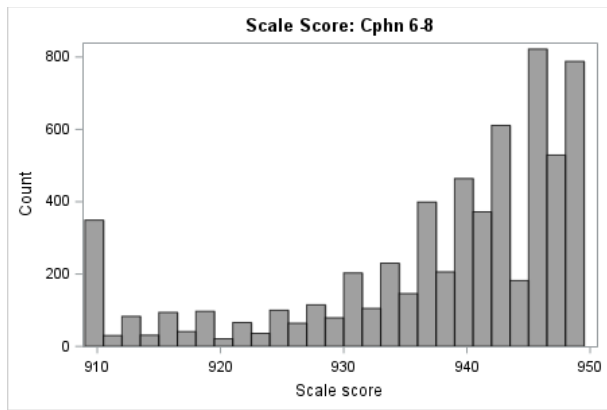
Figure 6.3.7B

Table 6.3.7A

Scale Score Descriptive Statistics: Cphn 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2152	910	949	936.70	11.13
7	2077	910	949	937.74	11.03
8	2034	910	949	937.96	11.15
Total	6263	910	949	937.46	11.12
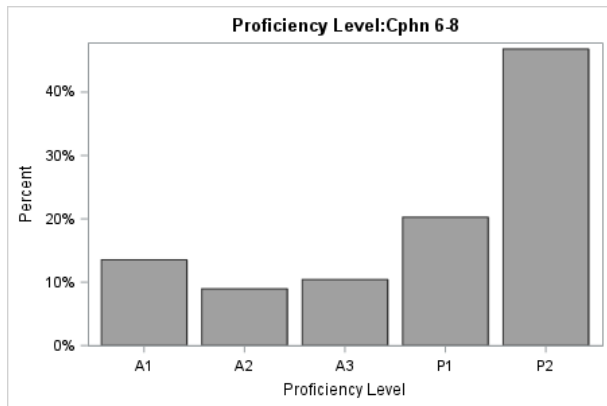
Table 6.3.7C

Proficiency Level Distribution: Cphn 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	317	15.15	269	13.81	262	14.15	848	14.38
A2	198	8.99	180	8.05	183	8.37	561	8.48
A3	240	11.39	222	11.65	191	9.50	653	10.86
P1	481	19.55	416	17.59	372	15.66	1269	17.64
P2	916	44.93	990	48.90	1026	52.33	2932	48.65
Total	2152	100	2077	100	2034	100	6263	100

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	110.469	0.945
Reading	0.7	141.054	0.941
Comprehension		123.600	0.962

*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
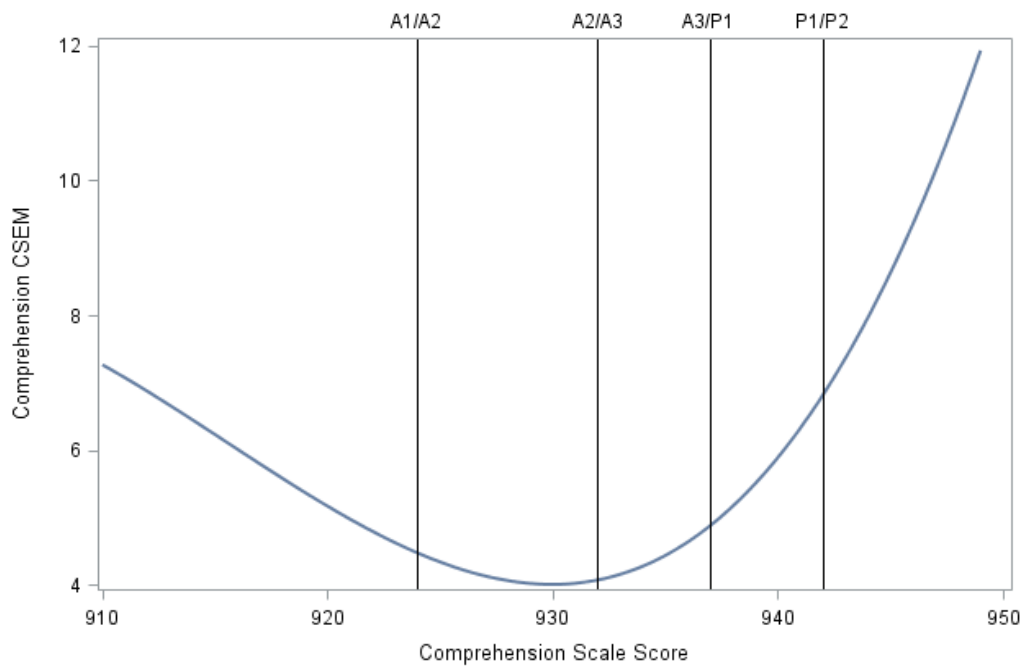
6.3.8   Overall Composite 6-8

Figure 6.3.8A

Scale Score: Over 6-8

Figure 6.3.8B

Proficiency Level:Over 6-8

Table 6.3.8A

Scale Score Descriptive Statistics: Over 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2101	910	950	934.79	10.18
7	2002	910	950	935.87	10.18
8	1969	910	950	935.95	10.41
Total	6072	910	950	935.52	10.27
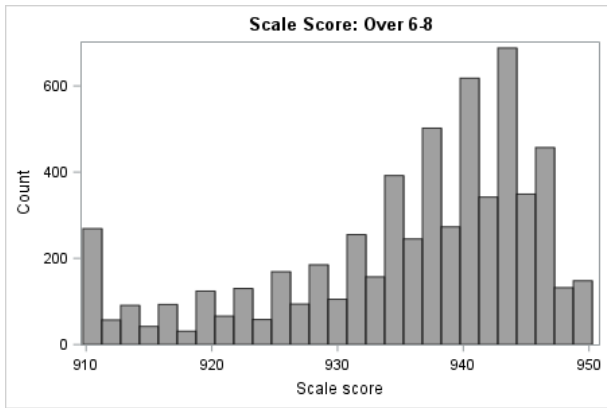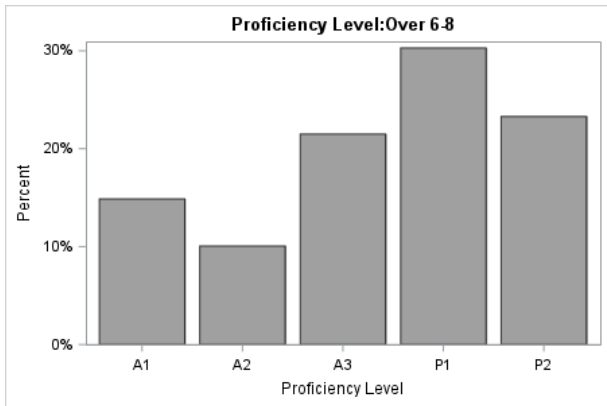
Table 6.3.8C

Proficiency Level Distribution: Over 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	330	15.71	286	14.29	287	14.58	903	14.87
A2	231	10.99	191	9.54	189	9.60	611	10.06
A3	497	23.66	430	21.48	378	19.20	1305	21.49
P1	648	30.84	598	29.87	593	30.12	1839	30.29
P2	395	18.80	497	24.83	522	26.51	1414	23.29
Total	2101	100	2002	100	1969	100	6072	100

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	110.469	0.945
Reading	0.35	141.054	0.941
Speaking	0.15	172.474	0.965
Writing	0.35	109.363	0.940
Overall Composite		105.377	0.980

*Variances from students who had results in all four domains

Table 6.3.8F
n/a

Table 6.3.8G
n/a

Table 6.3.8H
n/a

Table 6.3.8I
n/a

Table 6.3.8J

Accuracy and Consistency of Classification Indices: Over 6-8

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.738	0.657		0.552	

Conditional on Level	Level	Accuracy		Consistency	
	A1	0.954		0.932	
	A2	0.733		0.630	
	A3	0.887		0.829	
	P1	0.597		0.486	
	P2	0.684		0.637	

Indices at Cut Points		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
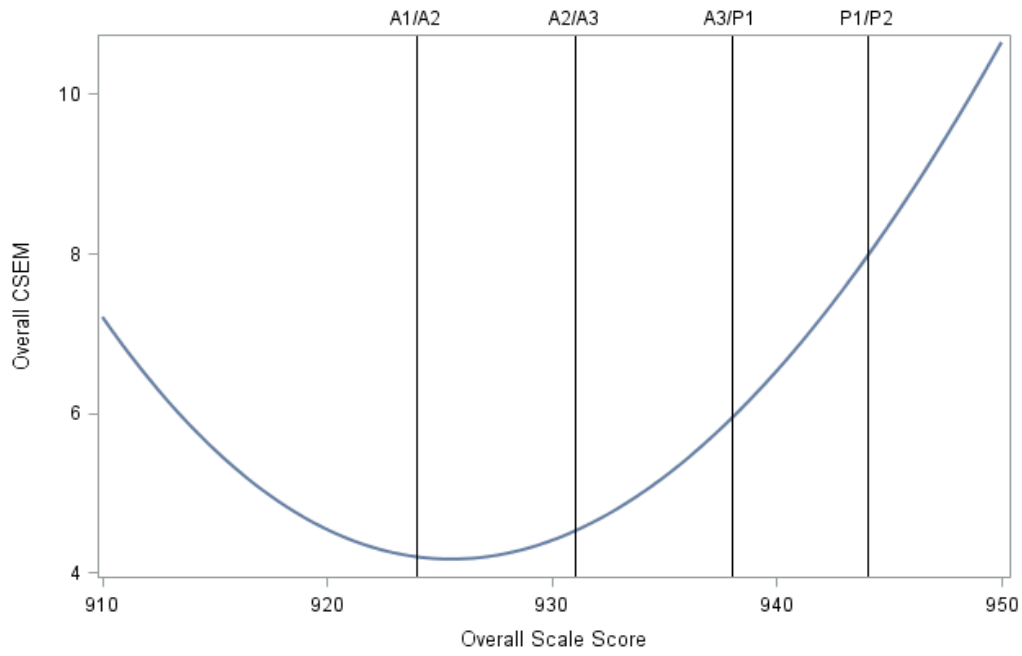
6.4   Grades: 9-12

6.4.1   Listening 9-12

Figure 6.4.1A
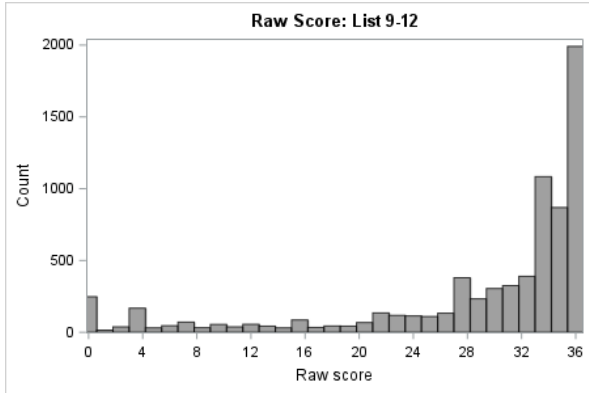
Figure 6.4.1B
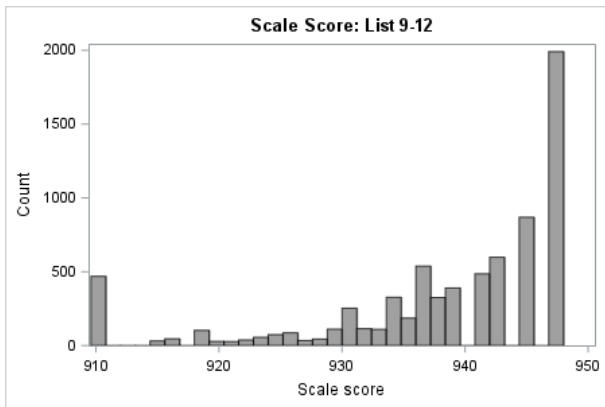
Figure 6.4.1C
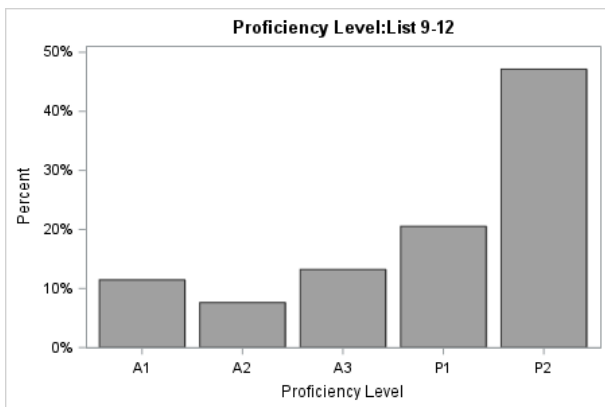
Table 6.4.1A

Raw Score Descriptive Statistics: List 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1780	0	36	28.13	10.11
10	1590	0	36	28.67	10.13
11	1541	0	36	29.40	9.56
12	2423	0	36	29.03	9.89
Total	7334	0	36	28.81	9.94

Table 6.4.1B

Scale Score Descriptive Statistics: List 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1780	910	947	937.23	10.41
10	1590	910	947	937.88	10.52
11	1541	910	947	938.56	10.03
12	2423	910	943	938.28	10.29
Total	7334	910	947	938.00	10.32

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	227	12.75	189	11.89	155	10.06	270	11.14	841	11.47
A2	150	8.43	118	7.42	106	6.88	185	7.64	559	7.62
A3	269	15.11	218	13.71	196	12.72	287	11.84	970	13.23
P1	371	20.84	314	19.75	325	21.09	496	20.47	1506	20.53
P2	763	42.87	751	47.23	759	49.25	1185	48.91	3458	47.15
Total	1780	100	1590	100	1541	100	2423	100	7334	100

Table 6.4.1D

Equating Summary: List 9-12

No equating summary is presented because the Alternate ACCESS Series 503 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.

Figure 6.4.1D
Test Characteristic Curve: List 9-12 S103
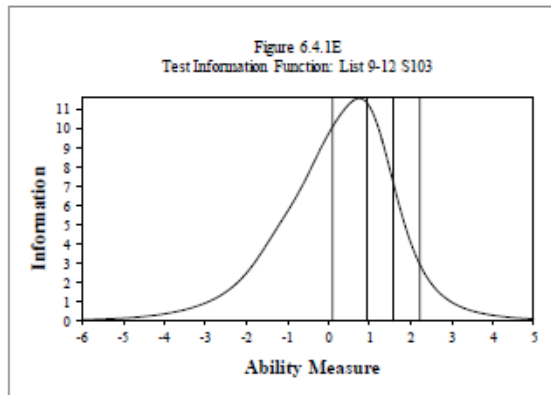

Figure 6.4.1E
Test Information Function: List 9-12 S103
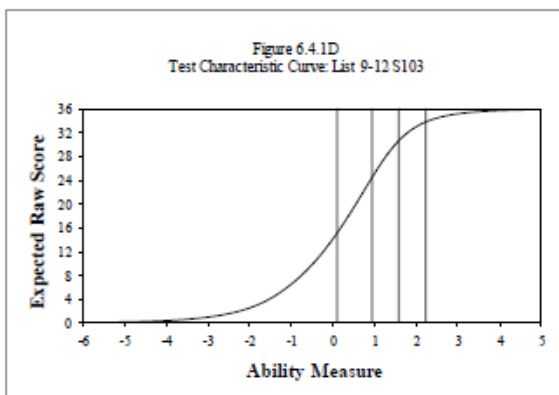
Table 6.4.1E

Reliability: List 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
7334	9	0.942	2.393

Table 6.4.1F

Item Analysis Summary: List 9-12

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	MOSR	9	0.00	82.83%	1.16	1.28
DIF Summary			Male/Female		Hispanic/Other	
	DIF Level		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA		6	3	3	6
	BB		0	0	0	0
	CC		0	0	0	0

Table 6.4.1G

Complete Item Analysis: List 9-12

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Infit Mnsq	Outfit Mnsq	Point Biserial	DIF M/F DIF Level	DIF M/F Favored Group	DIF H/O DIF Level	DIF H/O Favored Group
1. L1_A1_103	-1.38	Yes	88.00%	2.97	3.83	0.70	AA	M	AA	O
2. L2_A2_103	0.11	Yes	78.75%	1.24	1.00	0.86	AA	F	AA	O
3. L3_A2_103	-0.44	Yes	82.50%	1.00	0.73	0.89	AA	F	AA	O
4. L4_A3_103	0.71	Yes	73.25%	0.88	0.81	0.87	AA	F	AA	O
5. L5_A3_103	-0.23	Yes	80.50%	0.74	0.50	0.91	AA	M	AA	O
6. L6_P1_103	0.74	Yes	72.25%	0.71	0.69	0.89	AA	F	AA	H
7. L7_P1_103	0.77	Yes	72.25%	0.83	0.83	0.87	AA	F	AA	H
8. L8_P2_103	1.02	Yes	68.25%	0.91	0.86	0.83	AA	M	AA	H
9. L9_P2_103	0.77	Yes	70.25%	0.84	0.76	0.85	AA	M	AA	O

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	3.15	3.15	A1	4.03	4.03	A1	3.31	3.31	A1	3.18	3.18
1	A1	0.34	3.48	A1	0.25	4.28	A1	0.06	3.37	A1	0.17	3.34
2	A1	0.45	3.93	A1	0.69	4.97	A1	0.39	3.76	A1	0.54	3.88
3	A1	0.79	4.72	A1	0.25	5.22	A1	0.39	4.15	A1	0.33	4.21
4	A1	1.63	6.35	A1	1.95	7.17	A1	2.01	6.16	A1	1.86	6.07
5	A1	0.51	6.85	A1	0.13	7.30	A1	0.32	6.49	A1	0.62	6.69
6	A1	0.90	7.75	A1	0.38	7.67	A1	0.65	7.14	A1	0.58	7.26
7	A1	1.52	9.27	A1	0.88	8.55	A1	0.39	7.53	A1	0.99	8.25
8	A1	0.22	9.49	A1	0.57	9.12	A1	0.58	8.11	A1	0.41	8.67
9	A1	0.51	10.00	A1	0.44	9.56	A1	0.26	8.37	A1	0.33	9.00
10	A1	0.34	10.34	A1	0.25	9.81	A1	0.39	8.76	A1	0.45	9.45
11	A1	0.62	10.96	A1	0.50	10.31	A1	0.45	9.21	A1	0.50	9.95
12	A1	0.84	11.80	A1	0.75	11.07	A1	0.71	9.93	A1	0.74	10.69
13	A1	0.96	12.75	A1	0.82	11.89	A1	0.13	10.06	A1	0.45	11.14
14	A2	0.34	13.09	A2	0.50	12.39	A2	0.26	10.32	A2	0.54	11.68
15	A2	0.84	13.93	A2	0.88	13.27	A2	0.32	10.64	A2	0.54	12.22
16	A2	0.51	14.44	A2	0.38	13.65	A2	0.45	11.10	A2	0.70	12.92
17	A2	0.73	15.17	A2	0.44	14.09	A2	0.26	11.36	A2	0.41	13.33
18	A2	0.51	15.67	A2	0.38	14.47	A2	0.39	11.75	A2	0.95	14.28
19	A2	0.56	16.24	A2	0.69	15.16	A2	0.52	12.26	A2	0.58	14.86
20	A2	1.18	17.42	A2	0.63	15.79	A2	1.30	13.56	A2	0.70	15.56
21	A2	1.85	19.27	A2	0.69	16.48	A2	0.91	14.47	A2	1.11	16.67
22	A2	0.67	19.94	A2	0.69	17.17	A2	0.84	15.31	A2	0.58	17.25
23	A2	1.24	21.18	A2	2.14	19.31	A2	1.62	16.94	A2	1.53	18.78
24	A3	2.02	23.20	A3	1.82	21.13	A3	0.91	17.85	A3	1.44	20.22
25	A3	1.52	24.72	A3	1.51	22.64	A3	1.88	19.73	A3	1.24	21.46
26	A3	2.25	26.97	A3	2.26	24.91	A3	1.17	20.90	A3	1.61	23.07
27	A3	3.48	30.45	A3	2.89	27.80	A3	2.73	23.62	A3	1.82	24.89
28	A3	2.53	32.98	A3	2.45	30.25	A3	2.86	26.48	A3	2.39	27.28
29	A3	3.31	36.29	A3	2.77	33.02	A3	3.18	29.66	A3	3.34	30.62
30	P1	4.44	40.73	P1	4.53	37.55	P1	3.70	33.35	P1	4.00	34.63
31	P1	4.49	45.22	P1	4.09	41.64	P1	4.61	37.96	P1	4.50	39.13
32	P1	5.67	50.90	P1	4.65	46.29	P1	5.32	43.28	P1	5.49	44.61
33	P1	6.24	57.13	P1	6.48	52.77	P1	7.46	50.75	P1	6.48	51.09
34	P2	7.87	65.00	P2	7.67	60.44	P2	9.02	59.77	P2	8.13	59.22
35	P2	11.29	76.29	P2	11.57	72.01	P2	12.52	72.29	P2	12.01	71.23
36	P2	23.71	100.00	P2	27.99	100.00	P2	27.71	100.00	P2	28.77	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
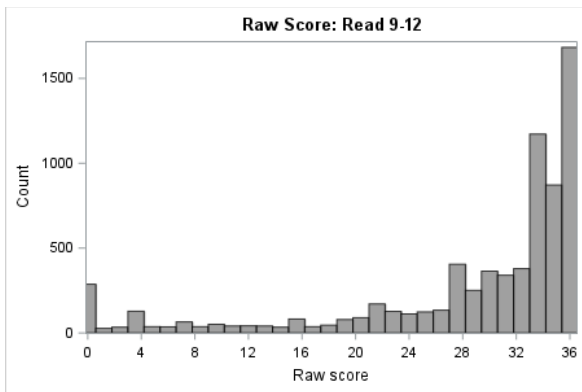
Figure 6.4.2A

Figure 6.4.2B
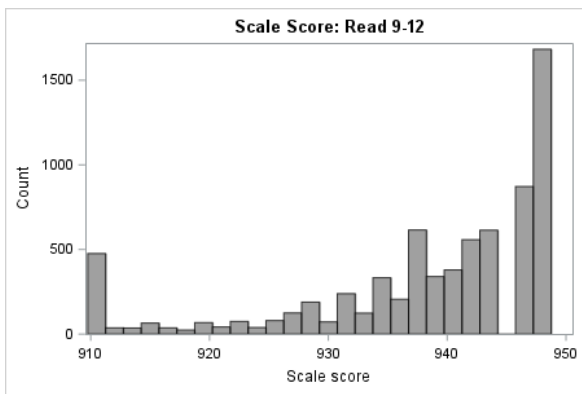
Figure 6.4.2C
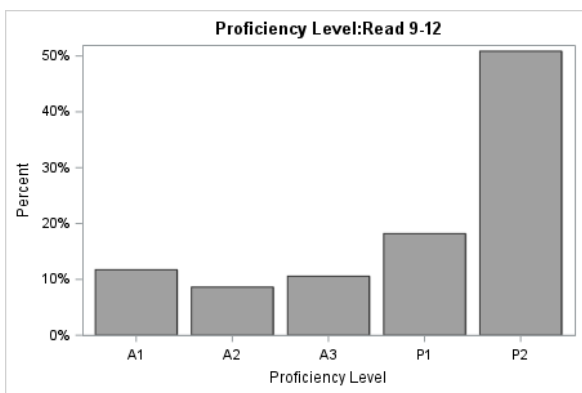
Table 6.4.2A

Raw Score Descriptive Statistics: Read 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1774	0	36	27.85	10.00
10	1596	0	36	28.50	9.96
11	1537	0	36	29.05	9.51
12	2421	0	36	28.47	10.11
Total	7328	0	36	28.45	9.94

Table 6.4.2B

Scale Score Descriptive Statistics: Read 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1774	910	948	937.30	10.88
10	1596	910	948	938.12	10.82
11	1537	910	948	938.74	10.50
12	2421	910	948	938.17	11.12
Total	7328	910	948	938.07	10.88

Table 6.4.2C

Proficiency Level Distribution: Read 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	223	12.57	177	11.09	162	10.54	298	12.31	860	11.74
A2	177	9.98	140	8.77	112	7.29	203	8.38	632	8.62
A3	203	11.44	170	10.65	153	9.95	249	10.29	775	10.58
P1	342	19.28	291	18.23	296	19.26	405	16.73	1334	18.20
P2	829	46.73	818	51.25	814	52.96	1266	52.29	3727	50.86
Total	1774	100	1596	100	1537	100	2421	100	7328	100

Table 6.4.2D

Equating Summary: Read 9-12

No equating summary is presented because the Alternate ACCESS Series 503 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 6.4.2D
Test Characteristic Curve: Read 9-12 S103

Figure 6.4.2E
Test Information Function: Read 9-12 S103
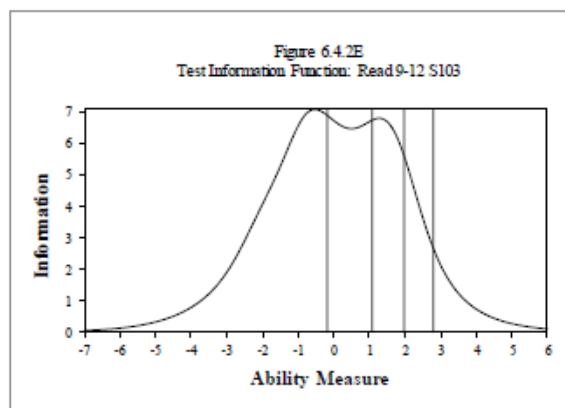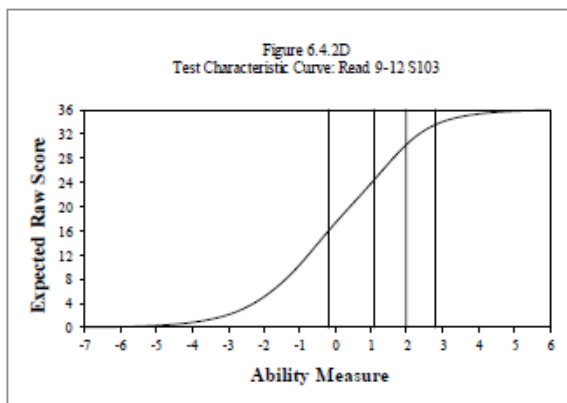
Table 6.4.2E

Reliability: Read 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
7328	9	0.945	2.325

Table 6.4.2F

Item Analysis Summary: Read 9-12

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	MOSR	9	0.00	79.19%	1.29	1.39
DIF Summary			Male/Female		Hispanic/Other	
	DIF Level		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA		7	2	5	4
	BB		0	0	0	0
	CC		0	0	0	0

Table 6.4.2G

Complete Item Analysis: Read 9-12

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Infit Mnsq	Outfit Mnsq	Point Biserial	DIF M/F DIF Level	Favored Group	DIF H/O DIF Level	Favored Group
1.R1_A1_103	-2.17		87.75%	3.45	3.63	0.72	AA	M	AA	O
2.R2_A2_103	-0.81	Yes	82.75%	1.32	0.93	0.88	AA	M	AA	O
3.R3_A2_103	-0.94	Yes	83.25%	1.08	0.70	0.88	AA	F	AA	H
4.R4_A3_103	0.01	Yes	78.00%	1.19	0.95	0.89	AA	F	AA	H
5.R5_A3_103	-0.93	Yes	82.00%	0.93	0.58	0.90	AA	M	AA	H
6.R6_P1_103	1.29	Yes	68.00%	0.91	0.85	0.85	AA	M	AA	O
7.R7_P1_103	0.95	Yes	69.00%	0.94	0.98	0.87	AA	M	AA	O
8.R8_P2_103	1.56	Yes	62.50%	0.85	0.96	0.81	AA	M	AA	O
9.R9_P2_103	1.74	Yes	60.50%	0.97	1.06	0.78	AA	F	AA	H

.

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	3.95	3.95	A1	4.57	4.57	A1	3.25	3.25	A1	3.88	3.88
1	A1	0.34	4.28	A1	0.63	5.20	A1	0.20	3.45	A1	0.33	4.21
2	A1	0.39	4.68	A1	0.50	5.70	A1	0.39	3.84	A1	0.54	4.75
3	A1	0.23	4.90	A1	0.44	6.14	A1	0.26	4.10	A1	0.54	5.29
4	A1	1.47	6.37	A1	0.88	7.02	A1	1.30	5.40	A1	1.65	6.94
5	A1	0.68	7.05	A1	0.56	7.58	A1	0.52	5.92	A1	0.33	7.27
6	A1	0.56	7.61	A1	0.38	7.96	A1	0.46	6.38	A1	0.54	7.81
7	A1	1.13	8.74	A1	0.38	8.33	A1	1.17	7.55	A1	0.87	8.67
8	A1	0.51	9.24	A1	0.50	8.83	A1	0.39	7.94	A1	0.58	9.25
9	A1	0.28	9.53	A1	0.19	9.02	A1	0.33	8.26	A1	0.45	9.71
10	A1	0.51	10.03	A1	0.50	9.52	A1	0.33	8.59	A1	0.21	9.91
11	A1	0.73	10.77	A1	0.44	9.96	A1	0.39	8.98	A1	0.62	10.53
12	A1	0.45	11.22	A1	0.31	10.28	A1	0.52	9.50	A1	0.87	11.40
13	A1	0.68	11.89	A1	0.38	10.65	A1	0.65	10.15	A1	0.54	11.94
14	A1	0.68	12.57	A1	0.44	11.09	A1	0.39	10.54	A1	0.37	12.31
15	A2	0.56	13.13	A2	0.63	11.72	A2	0.46	11.00	A2	0.50	12.80
16	A2	0.62	13.75	A2	0.50	12.22	A2	0.46	11.45	A2	0.70	13.51
17	A2	1.01	14.77	A2	0.38	12.59	A2	0.33	11.78	A2	0.33	13.84
18	A2	0.79	15.56	A2	0.56	13.16	A2	0.65	12.43	A2	0.54	14.37
19	A2	1.13	16.69	A2	0.88	14.04	A2	0.98	13.40	A2	1.24	15.61
20	A2	1.18	17.87	A2	1.07	15.10	A2	1.43	14.83	A2	1.24	16.85
21	A2	1.52	19.39	A2	1.75	16.85	A2	1.30	16.14	A2	0.99	17.84
22	A2	1.30	20.69	A2	1.25	18.11	A2	0.39	16.53	A2	0.95	18.79
23	A2	1.86	22.55	A2	1.75	19.86	A2	1.30	17.83	A2	1.90	20.69
24	A3	1.80	24.35	A3	1.63	21.49	A3	1.63	19.45	A3	1.20	21.89
25	A3	1.75	26.10	A3	1.94	23.43	A3	1.69	21.15	A3	1.49	23.38
26	A3	2.31	28.41	A3	2.01	25.44	A3	1.50	22.64	A3	1.57	24.95
27	A3	2.71	31.12	A3	2.51	27.94	A3	2.80	25.44	A3	2.81	27.76
28	A3	2.87	33.99	A3	2.57	30.51	A3	2.34	27.78	A3	3.22	30.98
29	P1	3.33	37.32	P1	3.38	33.90	P1	3.64	31.42	P1	3.39	34.37
30	P1	5.47	42.78	P1	4.64	38.53	P1	5.27	36.69	P1	4.63	38.99
31	P1	4.85	47.63	P1	5.08	43.61	P1	4.68	41.38	P1	4.17	43.16
32	P1	5.64	53.27	P1	5.14	48.75	P1	5.66	47.04	P1	4.54	47.71
33	P2	8.85	62.12	P2	7.64	56.39	P2	6.57	53.61	P2	7.35	55.06
34	P2	7.61	69.73	P2	9.65	66.04	P2	9.89	63.50	P2	7.15	62.21
35	P2	11.22	80.95	P2	12.22	78.26	P2	11.84	75.34	P2	12.23	74.43
36	P2	19.05	100.00	P2	21.74	100.00	P2	24.66	100.00	P2	25.57	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
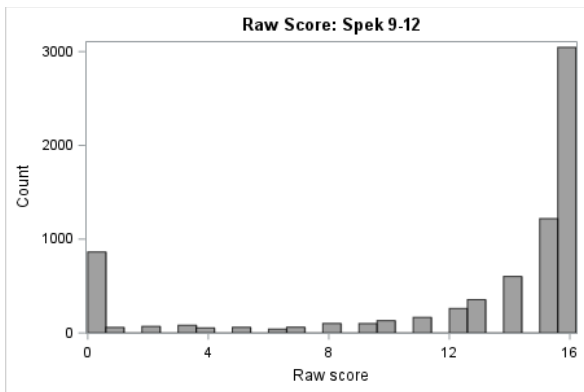
Table 6.4.3A

Raw Score Descriptive Statistics: Spek 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1757	0	16	12.20	5.57
10	1587	0	16	12.41	5.38
11	1525	0	16	12.61	5.25
12	2401	0	16	12.21	5.65
Total	7270	0	16	12.33	5.49

Figure 6.4.3B
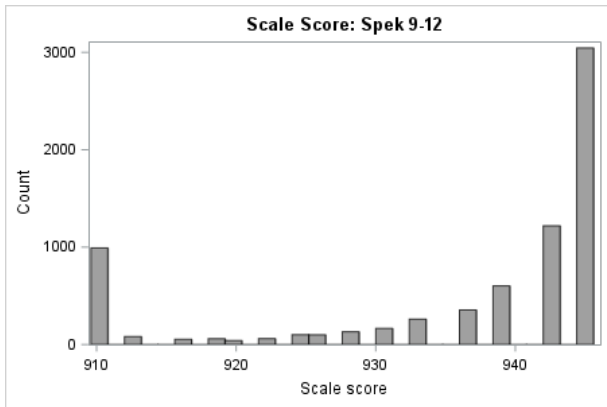
Table 6.4.3B

Scale Score Descriptive Statistics: Spek 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1757	910	945	935.71	12.50
10	1587	910	945	936.17	12.17
11	1525	910	945	936.61	11.87
12	2401	910	945	935.77	12.74
Total	7270	910	945	936.02	12.38

Figure 6.4.3C
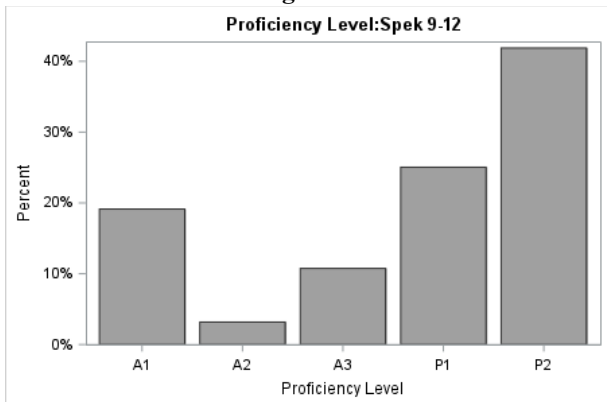
Table 6.4.3C

Proficiency Level Distribution: Spek 9-12

Level	Grade 9 Count	Grade 9 Percent	Grade 10 Count	Grade 10 Percent	Grade 11 Count	Grade 11 Percent	Grade 12 Count	Grade 12 Percent	Total Count	Total Percent
A1	348	19.81	301	18.97	248	16.26	494	20.57	1391	19.13
A2	49	2.79	55	3.47	63	4.13	64	2.67	231	3.18
A3	208	11.84	164	10.33	167	10.95	243	10.12	782	10.76
P1	446	25.38	412	25.96	405	26.56	558	23.24	1821	25.05
P2	706	40.18	655	41.27	642	42.10	1042	43.40	3045	41.88
Total	1757	100	1587	100	1525	100	2401	100	7270	100

Table 6.4.3D

Equating Summary: Spek 9-12

No equating summary is presented because the Alternate ACCESS Series 503 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).
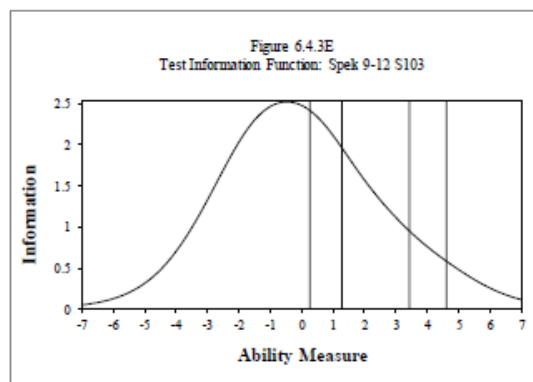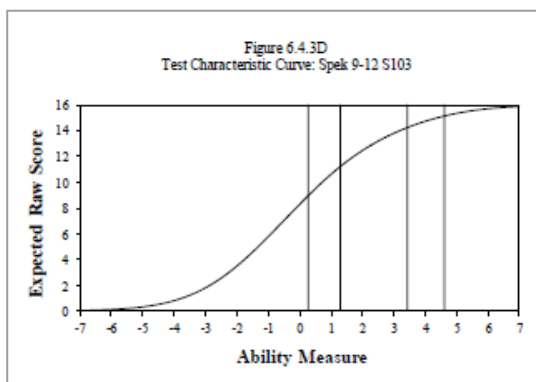

Figure 6.4.3D
Test Characteristic Curve: Spek 9-12 S103

Figure 6.4.3E
Test Information Function: Spek 9-12 S103

Table 6.4.3E

Reliability: Spek 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
7270	8	0.968	0.987

Table 6.4.3F

Item Analysis Summary: Spek 9-12

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	CR	8	0.00	76.94%	0.97	0.84
DIF Summary		Male/Female			Hispanic/Other	
	DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)	
	AA	4	4	5	3	
	BB	0	0	0	0	
	CC	0	0	0	0	

Table 6.4.3G

Complete Item Analysis: Spek 9-12

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Infit Mnsq	Outfit Mnsq	Point Biserial	DIF M/F DIF Level	DIF M/F Favored Group	DIF H/O DIF Level	DIF H/O Favored Group
1.S1_A1_103	-2.09		80.50%	1.43	1.65	0.90	AA	F	AA	O
2.S2_A2_103	-1.16		79.00%	1.27	1.23	0.91	AA	M	AA	H
3.S3_A3_103	-0.43	Yes	77.00%	0.98	0.82	0.93	AA	F	AA	H
4.S4_A1_103	-0.69	Yes	77.00%	0.90	0.90	0.93	AA	M	AA	O
5.S5_A2_103	-0.47	Yes	77.00%	0.84	0.73	0.94	AA	F	AA	H
6.S6_A3_103	-0.16	Yes	76.00%	0.75	0.64	0.94	AA	M	AA	O
7.S7_P1_103	1.74	Yes	69.00%	0.88	0.76	0.87	AA	M	AA	H
8.S8_P2_103	3.64	Yes	56.00%	0.78	0.90	0.73	AA	M	AA	H

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	12.92	12.92	A1	10.96	10.96	A1	10.49	10.49	A1	12.58	12.58
1	A1	0.68	13.60	A1	0.95	11.91	A1	0.85	11.34	A1	0.79	13.37
2	A1	0.85	14.46	A1	0.76	12.67	A1	0.92	12.26	A1	1.21	14.58
3	A1	0.68	15.14	A1	1.32	13.99	A1	0.98	13.25	A1	1.42	15.99
4	A1	0.74	15.88	A1	0.76	14.74	A1	0.52	13.77	A1	0.87	16.87
5	A1	1.02	16.90	A1	0.57	15.31	A1	0.79	14.56	A1	0.87	17.74
6	A1	0.85	17.76	A1	0.76	16.07	A1	0.13	14.69	A1	0.50	18.24
7	A1	0.85	18.61	A1	0.82	16.89	A1	0.72	15.41	A1	0.92	19.16
8	A1	1.20	19.81	A1	2.08	18.97	A1	0.85	16.26	A1	1.42	20.57
9	A2	1.14	20.94	A2	1.13	20.10	A2	2.10	18.36	A2	1.25	21.82
10	A2	1.65	22.60	A2	2.33	22.43	A2	2.03	20.39	A2	1.42	23.24
11	A3	2.96	25.55	A3	1.89	24.32	A3	1.97	22.36	A3	2.25	25.49
12	A3	3.53	29.08	A3	3.28	27.60	A3	3.54	25.90	A3	3.87	29.36
13	A3	5.35	34.43	A3	5.17	32.77	A3	5.44	31.34	A3	4.00	33.36
14	P1	8.54	42.97	P1	8.44	41.21	P1	9.77	41.11	P1	7.04	40.40
15	P1	16.85	59.82	P1	17.52	58.73	P1	16.79	57.90	P1	16.20	56.60
16	P2	40.18	100.00	P2	41.27	100.00	P2	42.10	100.00	P2	43.40	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
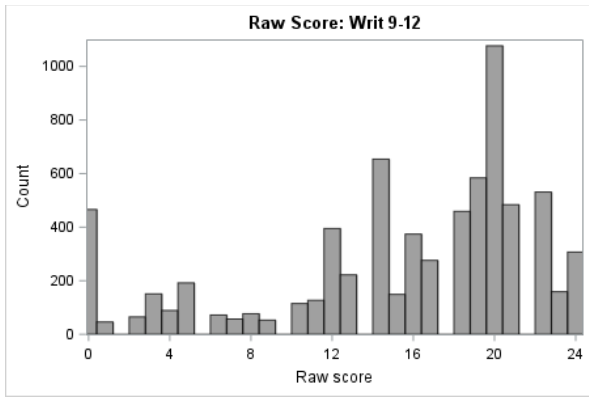
Figure 6.4.4A

Figure 6.4.4B
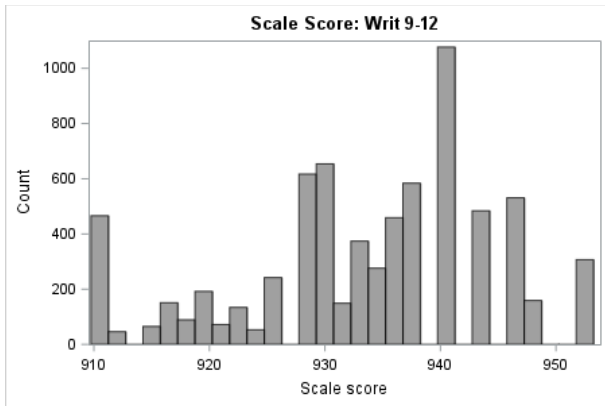
Figure 6.4.4C
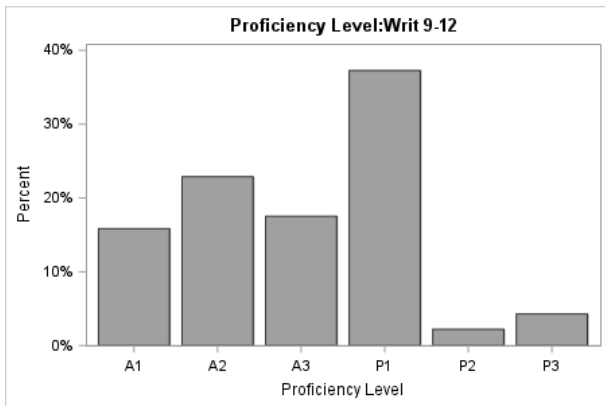
Table 6.4.4A

Raw Score Descriptive Statistics: Writ 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1739	0	24	14.84	6.69
10	1563	0	24	15.49	6.64
11	1511	0	24	15.59	6.65
12	2368	0	24	15.33	6.74
Total	7181	0	24	15.30	6.69

Table 6.4.4B

Scale Score Descriptive Statistics: Writ 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1739	910	953	932.92	10.72
10	1563	910	953	933.95	10.74
11	1511	910	953	934.18	10.86
12	2368	910	953	933.78	10.94
Total	7181	910	953	933.69	10.83

Table 6.4.4C

Proficiency Level Distribution: Writ 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	301	17.31	236	15.10	225	14.89	376	15.88	1138	15.85
A2	433	24.90	342	21.88	336	22.24	532	22.47	1643	22.88
A3	331	19.03	263	16.83	262	17.34	402	16.98	1258	17.52
P1	572	32.89	632	40.44	570	37.72	902	38.09	2676	37.27
P2	37	2.13	28	1.79	45	2.98	49	2.07	159	2.21
P3	65	3.74	62	3.97	73	4.83	107	4.52	307	4.28
Total	1739	100	1563	100	1511	100	2368	100	7181	100

Table 6.4.4D

Equating Summary: Writ 9-12

No equating summary is presented because the Alternate ACCESS Series 503 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).
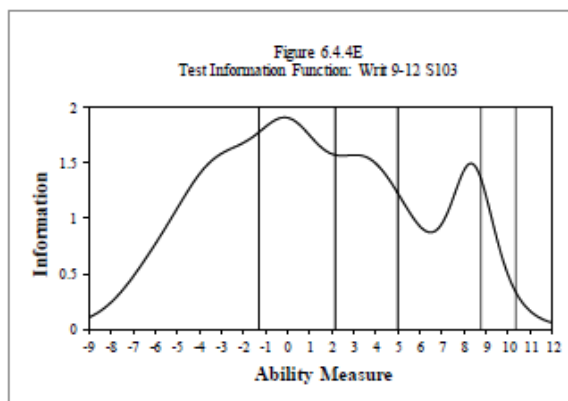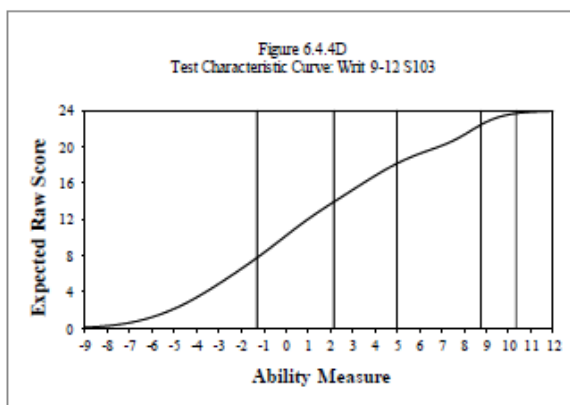

Figure 6.4.4D
Test Characteristic Curve: Writ 9-12 S103

Figure 6.4.4E
Test Information Function: Writ 9-12 S103

Table 6.4.4E

Reliability: Writ 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
7181	10	0.943	1.596

Table 6.4.4F

Item Analysis Summary: Writ 9-12

Item Summary	Item Type	No. of Items	Average Item Difficulty (in logits)	Average of % of Max. Possible Score Points	Average Infit Mean	Average Outfit Mean Square
	CR	10	0.00	68.75%	1.42	5.48
DIF Summary	DIF Level	Male/Female		Hispanic/Other	
		Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	AA	4	6	4	6
	BB	0	0	0	0
	CC	0	0	0	0

Table 6.4.4G

Complete Item Analysis: Writ 9-12

Name	Item Difficulty (in logits)	Anchored?	% of Max. Possible Score Points	Fit Statistics		Point Biserial	DIF			
							M/F		H/O	
				Infit Mnsq	Outfit Mnsq		DIF Level	Favored Group	DIF Level	Favored Group
1.W1_A1_103	-4.41		83.50%	2.88	9.90	0.75	AA	M	AA	O
2.W2_A2_103	-3.09		81.00%	1.73	9.90	0.81	AA	F	AA	H
3.W3_A3_103	-1.35		77.00%	1.69	6.60	0.83	AA	M	AA	O
4.W4_P1_103	1.73	Yes	62.50%	1.43	2.18	0.84	AA	F	AA	O
5.W5_A1_103	-2.00	Yes	78.00%	1.40	9.90	0.83	AA	M	AA	H
6.W6_A2_103	-1.77	Yes	77.50%	1.25	8.39	0.84	AA	M	AA	H
7.W7_A3_103	-0.30		73.00%	1.50	6.32	0.84	AA	M	AA	O
8.W8_P1_103	1.83	Yes	62.50%	1.07	1.27	0.87	AA	F	AA	H
9.W9_P3_103	6.28	Yes	33.25%	1.03	3.13	0.71	AA	F	AA	O
10.W10_P3_103	6.27	Yes	33.75%	1.04	2.70	0.66	AA	M	AA	O

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	6.27	6.27	A1	6.53	6.53	A1	6.49	6.49	A1	6.63	6.63
1	A1	0.75	7.02	A1	0.70	7.23	A1	0.53	7.02	A1	0.59	7.22
2	A1	0.92	7.94	A1	0.83	8.06	A1	0.73	7.74	A1	1.06	8.28
3	A1	2.30	10.24	A1	2.11	10.17	A1	1.79	9.53	A1	2.15	10.43
4	A1	1.50	11.73	A1	0.96	11.13	A1	1.32	10.85	A1	1.18	11.61
5	A1	3.16	14.89	A1	2.50	13.63	A1	2.18	13.04	A1	2.74	14.36
6	A1	1.44	16.33	A1	1.02	14.65	A1	1.06	14.10	A1	0.63	14.99
7	A1	0.98	17.31	A1	0.45	15.10	A1	0.79	14.89	A1	0.89	15.88
8	A2	1.04	18.34	A2	0.96	16.06	A2	0.79	15.68	A2	1.35	17.23
9	A2	0.92	19.26	A2	0.58	16.63	A2	0.66	16.35	A2	0.76	17.99
10	A2	2.13	21.39	A2	1.66	18.30	A2	1.26	17.60	A2	1.39	19.38
11	A2	1.96	23.35	A2	1.34	19.64	A2	2.12	19.72	A2	1.69	21.07
12	A2	5.81	29.15	A2	5.37	25.02	A2	5.43	25.15	A2	5.41	26.48
13	A2	3.45	32.60	A2	2.75	27.77	A2	3.11	28.26	A2	3.04	29.52
14	A2	9.60	42.21	A2	9.21	36.98	A2	8.87	37.13	A2	8.83	38.34
15	A3	2.47	44.68	A3	2.18	39.16	A3	1.46	38.58	A3	2.11	40.46
16	A3	5.18	49.86	A3	4.99	44.15	A3	5.03	43.61	A3	5.49	45.95
17	A3	3.80	53.65	A3	4.09	48.24	A3	4.30	47.92	A3	3.42	49.37
18	A3	7.59	61.24	A3	5.57	53.81	A3	6.55	54.47	A3	5.95	55.32
19	P1	7.30	68.55	P1	9.09	62.89	P1	9.00	63.47	P1	7.56	62.88
20	P1	12.59	81.14	P1	17.15	80.04	P1	13.83	77.30	P1	16.09	78.97
21	P1	6.90	88.04	P1	5.82	85.86	P1	7.02	84.32	P1	7.05	86.02
22	P1	6.10	94.13	P1	8.38	94.24	P1	7.88	92.19	P1	7.39	93.41
23	P2	2.13	96.26	P2	1.79	96.03	P2	2.98	95.17	P2	2.07	95.48
24	P3	3.74	100.00	P3	3.97	100.00	P3	4.83	100.00	P3	4.52	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
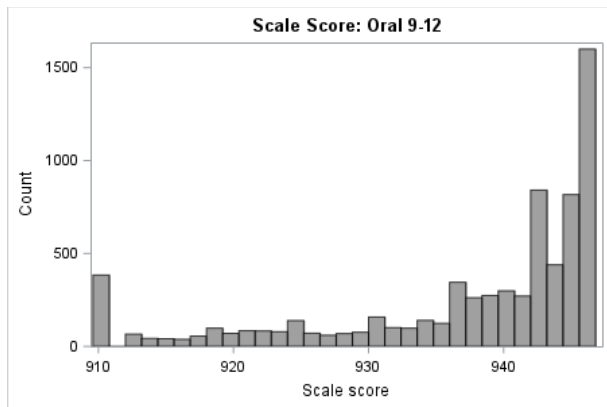
Table 6.4.5A

Scale Score Descriptive Statistics: Oral 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1747	910	946	936.71	10.67
10	1573	910	946	937.27	10.60
11	1515	910	946	937.87	10.17
12	2390	910	946	937.27	10.69
Total	7225	910	946	937.26	10.56

Figure 6.4.5B
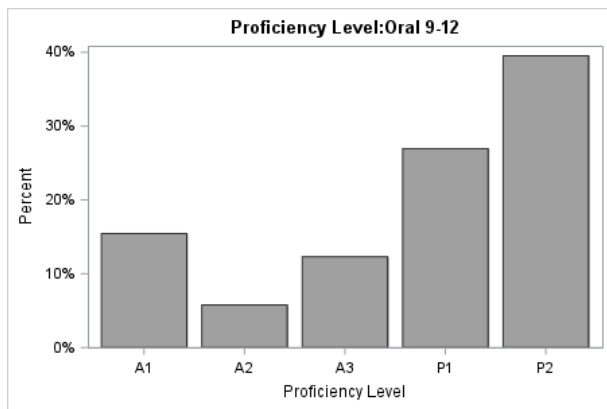
Table 6.4.5C

Proficiency Level Distribution: Oral 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	280	16.03	242	15.38	201	13.27	393	16.44	1116	15.45
A2	115	6.58	90	5.72	78	5.15	134	5.61	417	5.77
A3	247	14.14	196	12.46	194	12.81	252	10.54	889	12.30
P1	473	27.07	411	26.13	421	27.79	642	26.86	1947	26.95
P2	632	36.18	634	40.31	621	40.99	969	40.54	2856	39.53
Total	1747	100	1573	100	1515	100	2390	100	7225	100

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	106.546	0.942
Speaking	0.5	153.344	0.968
Oral		111.553	0.975

*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
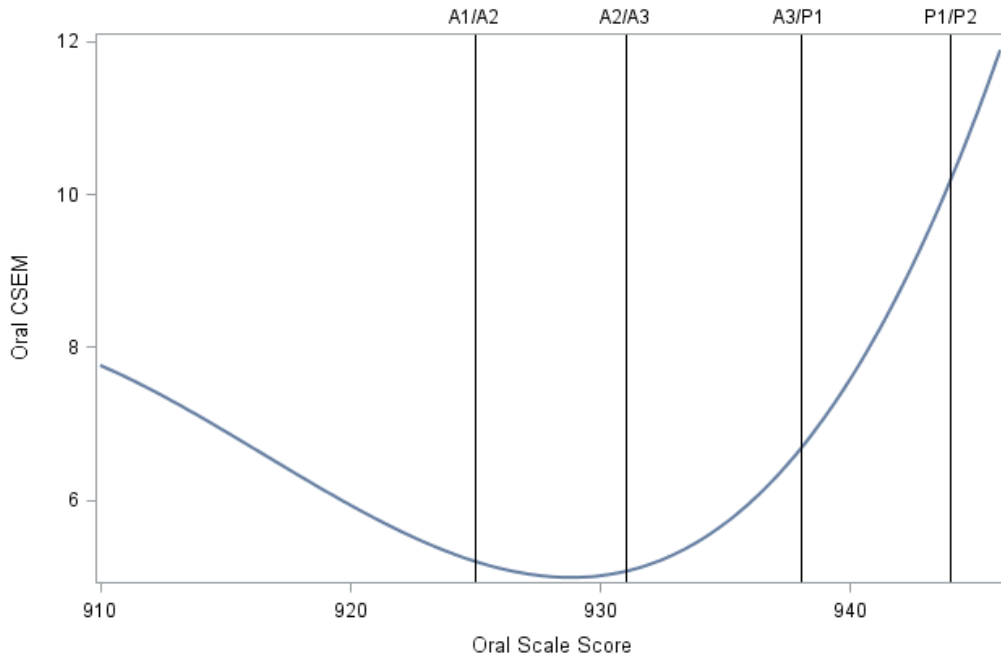
6.4.6   Literacy Composite 9-12

Figure 6.4.6A

Scale Score: Litr 9-12
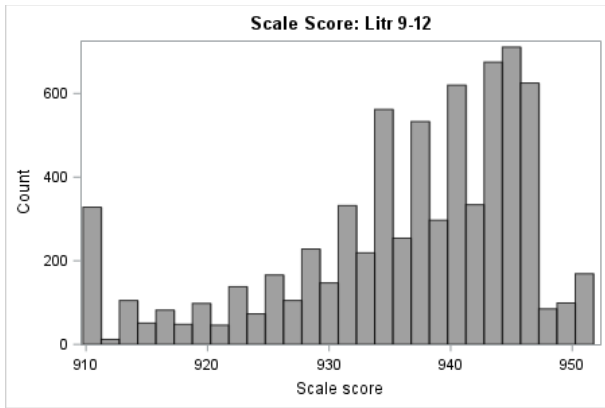
Table 6.4.6A

Scale Score Descriptive Statistics: Litr 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1728	910	951	935.42	10.07
10	1554	910	951	936.41	10.05
11	1504	910	951	936.76	10.03
12	2356	910	951	936.28	10.40
Total	7142	910	951	936.20	10.18

Figure 6.4.6B

Proficiency Level: Litr 9-12
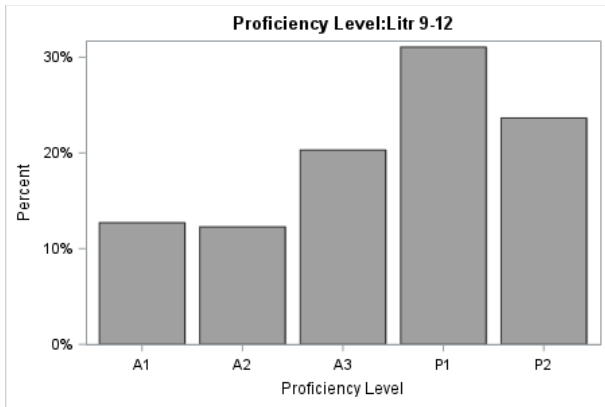
Table 6.4.6C

Proficiency Level Distribution: Litr 9-12

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	229	13.25	190	12.23	172	11.44	317	13.46	908	12.71
A2	255	14.76	190	12.23	163	10.84	269	11.42	877	12.28
A3	367	21.24	306	19.69	323	21.48	454	19.27	1450	20.30
P1	556	32.18	501	32.24	454	30.19	707	30.01	2218	31.06
P2	321	18.58	367	23.62	392	26.06	609	25.85	1689	23.65
Total	1728	100	1554	100	1504	100	2356	100	7142	100

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	118.279	0.945
Writing	0.5	117.385	0.943
Literacy		103.549	0.968

*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
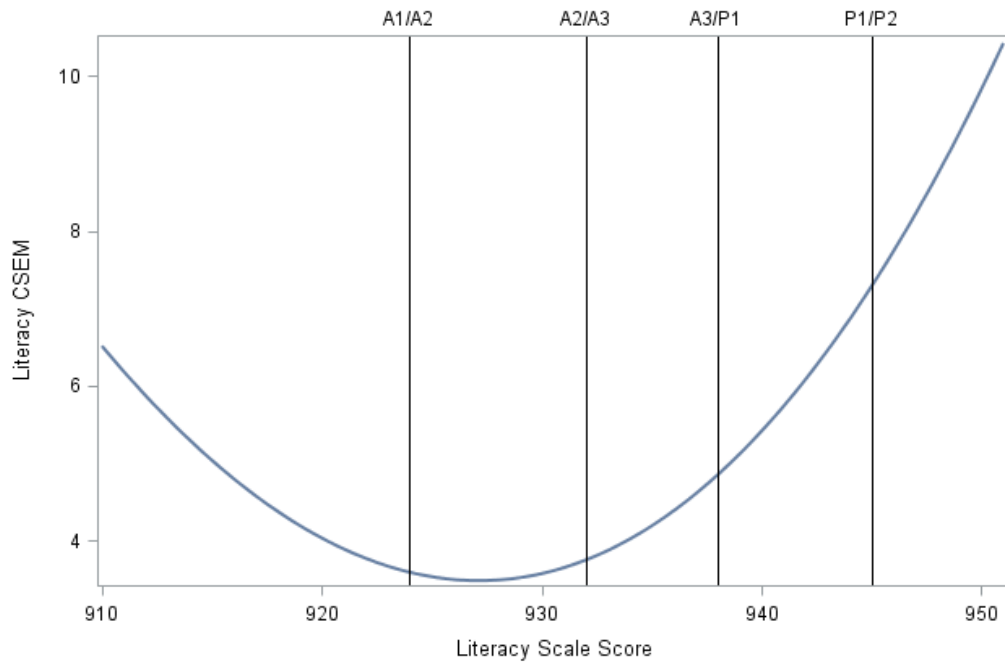
6.4.7 Comprehension Composite 9-12

Figure 6.4.7A

Scale Score: Cphn 9-12
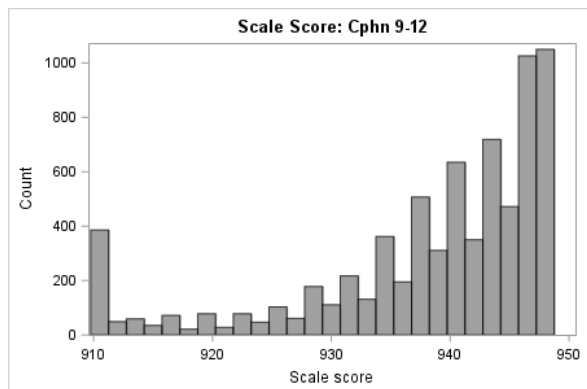
Table 6.4.7A

Scale Score Descriptive Statistics: Cphn 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1766	910	948	937.42	10.40
10	1582	910	948	938.16	10.45
11	1524	910	948	938.82	10.07
12	2409	910	948	938.33	10.63
Total	7281	910	948	938.17	10.43

Figure 6.4.7B

Proficiency Level:Cphn 9-12
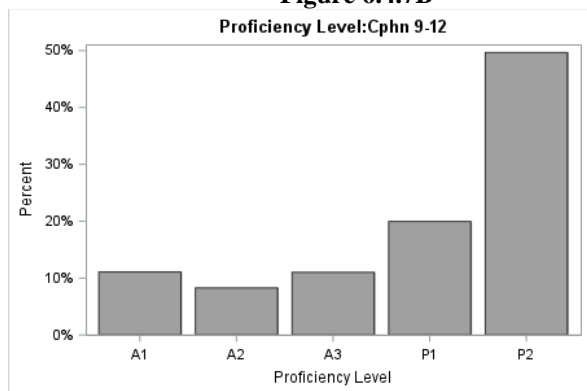
Table 6.4.7C

Proficiency Level Distribution: Cphn 9-12

Level	Grade 9 Count	Grade 9 Percent	Grade 10 Count	Grade 10 Percent	Grade 11 Count	Grade 11 Percent	Grade 12 Count	Grade 12 Percent	Total Count	Total Percent
A1	209	12.57	174	11.09	145	10.54	278	12.31	806	11.74
A2	174	9.98	124	8.77	116	7.29	189	8.38	603	8.62
A3	221	11.44	190	10.65	151	9.95	240	10.29	802	10.58
P1	371	19.28	306	18.23	303	19.26	473	16.73	1453	18.20
P2	791	46.73	788	51.25	809	52.96	1229	52.29	3617	50.86
Total	1766	100	1582	100	1524	100	2409	100	7281	100

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	106.546	0.942
Reading	0.7	118.279	0.945
Comprehension		108.757	0.966

*Variances from students who had results in all four domains
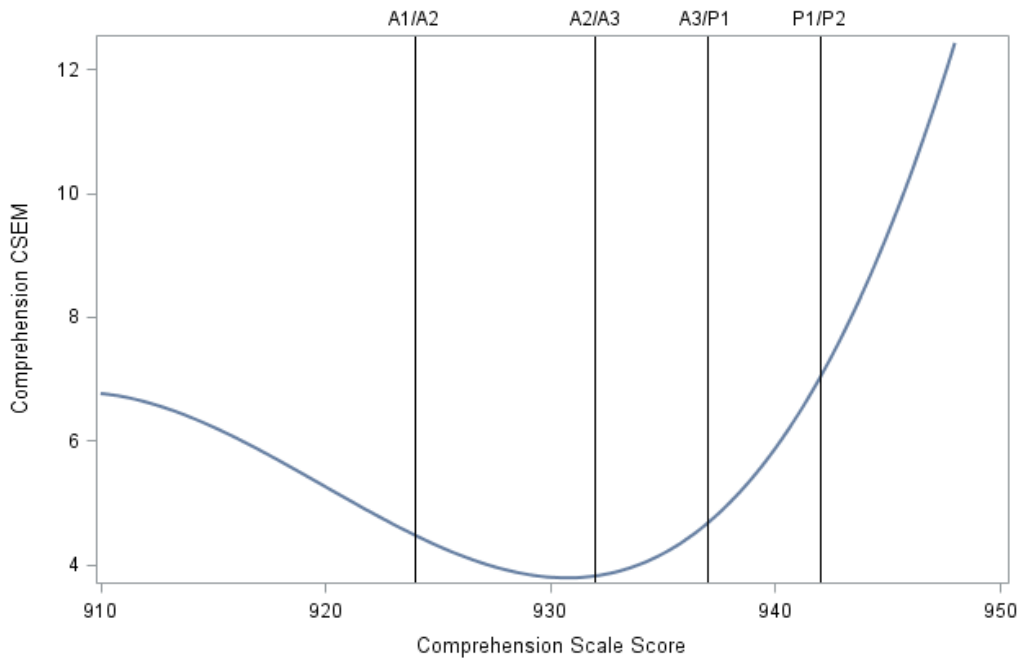

Table 6.4.7F
n/a

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
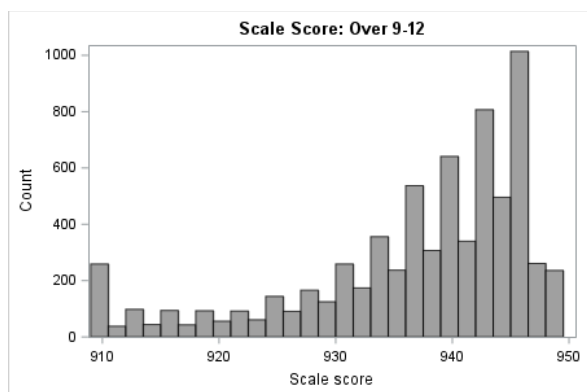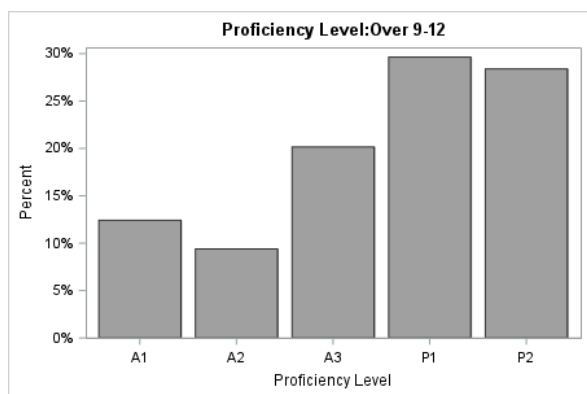
6.4.8 Overall Composite 9-12

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	1713	910	949	935.70	9.93
10	1538	910	949	936.58	9.88
11	1485	910	949	936.97	9.77
12	2333	910	949	936.49	10.16
Total	7069	910	949	936.42	9.97

Figure 6.4.8B

Table 6.4.8C

Proficiency Level Distribution: Over 9-12

Level	Grade 9 Count	Grade 9 Percent	Grade 10 Count	Grade 10 Percent	Grade 11 Count	Grade 11 Percent	Grade 12 Count	Grade 12 Percent	Total Count	Total Percent
A1	226	13.19	185	12.03	157	10.57	310	13.29	878	12.42
A2	181	10.57	146	9.49	136	9.16	201	8.62	664	9.39
A3	367	21.42	297	19.31	309	20.81	452	19.37	1425	20.16
P1	546	31.87	478	31.08	422	28.42	649	27.82	2095	29.64
P2	393	22.94	432	28.09	461	31.04	721	30.90	2007	28.39
Total	1713	100	1538	100	1485	100	2333	100	7069	100

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	106.546	0.942
Reading	0.35	118.279	0.945
Speaking	0.15	153.344	0.968
Writing	0.35	117.385	0.943
Overall Composite		99.420	0.981

*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**			**Consistency**
	Cut Point		**False Positives**	**False Negatives**	
	A1/A2	0.985	0.008	0.007	0.979
	A2/A3	0.969	0.019	0.011	0.956
	A3/P1	0.957	0.011	0.033	0.940
	P1/P2	0.683	0.305	0.012	0.778

Figure 6.4.8F CSEM for Overall Composite 9-12
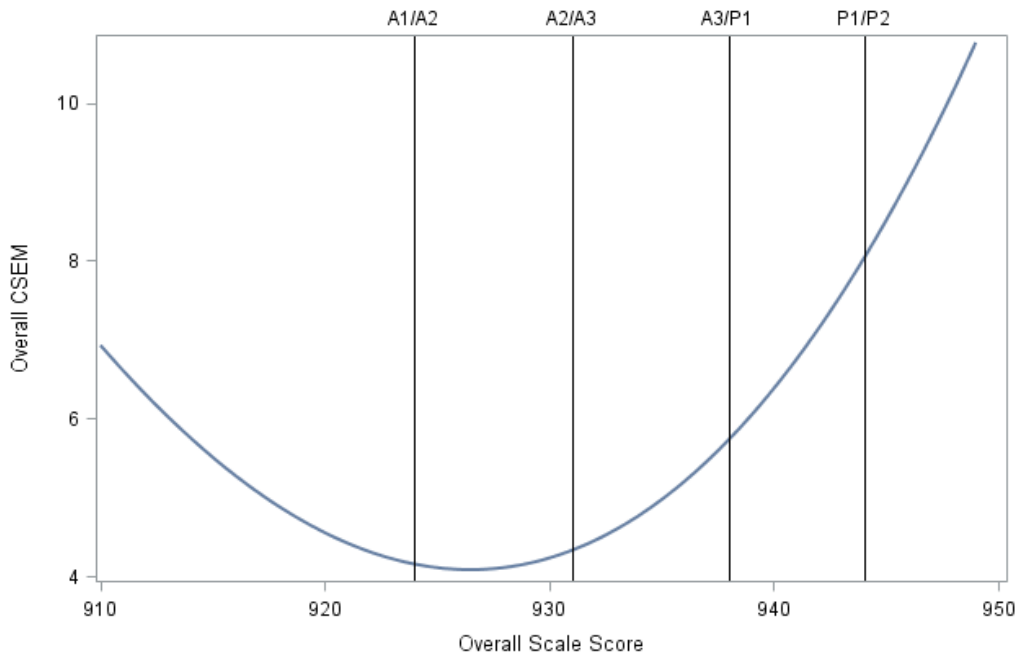
References

Allen, N.L., Carlson, J.E., & Zalanak, C.A. (1999). *The NAEP 1996 technical report.* Washington, DC: National Center for Education Statistics.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, D.C.: American Psychological Association.

Andrich D. (1978) A rating scale formulation for ordered response categories. *Psychometrika, 43,*561-573.

Bachman, L. F & A. B. Palmer (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. UK: Oxford University Press.

Brennan, R.L. (2004). BB-CLASS: a computer program that uses the beta-binomial model for classification consistency and accuracy. [Computer Software]. Iowa City, IA: CASMA.

Center for Applied Linguistics. (2012a). *Alternate ACCESS for ELLs™ Standard Setting Study: Technical Brief.* Available at www.wida.wisc.edu.

Center for Applied Linguistics (2013). *Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report*.

Chapelle, C. A., Enright, M. & Jamieson, J. (Eds.) (2008). Building a validity argument for the Test of English as a Foreign Language. London: Routledge.

Chapelle, C. A., Enright, M. K., & Jamieson, J. (2010). Does an Argument-Based Approach to Validity Make a Difference?. *Educational Measurement: Issues and Practice*, *29*(1), 3-13.

Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.

Cronbach, L.J., Schonemann, P., & McKie, D. (1965). Alpha coefficients for stratified-parallel tests. *Educational and Psychological Measurement, 25,* 291-312.

Every Student Succeeds (ESSA) Act of 2015, Pub. L. No. 114-95, § 114 Stat. 1177 (2015-2016). Individuals With Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Kamata, A., Turhan, A., & Darandari, E. (2003, April). *Estimating reliability for multidimensional composite scale scores.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Kane, M. (2002). Validating high-stakes testing programs. *Educational measurement: Issues and practice*, *21*(1), 31-41.

Kane, M., & Case, S.M. (2004). The reliability and validity of weighted composite scores. *Applied Measurement in Education, 17,* 221-240.

Kane, M. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*(1), 1-73.

Kenyon, D.M. (2006). *Development and Field Test of ACCESS for ELLs®.* (WIDA Consortium Technical Report No. 1).

Lee, W., Hanson, B.A., & Brennan, R.L. (2002). Estimating consistency and accuracy indices for multiple classifications. *Applied Psychological Measurement, 26*, 412-432.

Linacre, J.M. (2002, Autumn). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions, 16*, 878. Retrieved April 10, 2006, from www.rasch.org/rmt/rmt162f.htm)

Linacre, J.M. (2006). Winsteps (Version 3.60) [computer software]. Chicago, IL: Winsteps.com.

Livingston, S.A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement, 32,* 179-197.

Llosa, L. (2008). Building and Supporting a Validity Argument for a Standards-Based Classroom Assessment of English Proficiency Based on Teacher Judgments. *Educational Measurement: Issues and Practice*, *27*(3), 32-42.

Meyer, J.P. (2014). jMetrik. [Computer Software]. http://www.jmetrik.com.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2004). *A Brief Introduction to Evidence-Centered Design. CSE Report 632.* US Department of Education.

Rudner, L. (2001, Spring). Informed test component weighting. *Educational Measurement: Issues and Practice, 20:1, 16-19.*

WIDA Consortium. (2007). *English Language Proficiency Standards and Resource Guide, 2007Edition, 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.