Annual Technical Report for

ALTERNATE ACCESS for ELLs English Language Proficiency Test, Series 601, 2022-2023 Administration

Annual Technical Report No. 11

Prepared by

Psychometrics team

WIDA

December, 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.

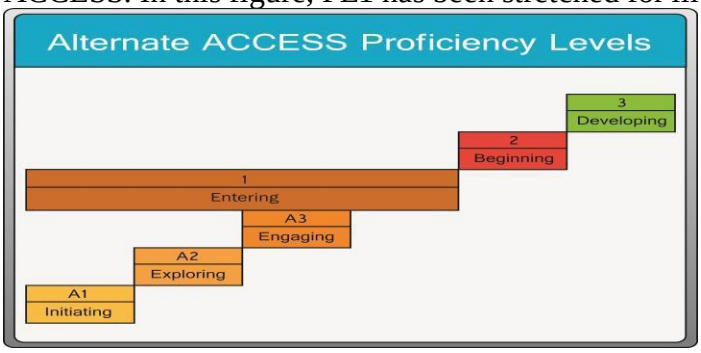


Figure 1.2.4A. Alternate ACCESS Proficiency Levels

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-level 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 include reading through the ACCESS for ELLs Test Administration Manual and the Alternate ACCESS Test Administration training materials (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). Indiscriminate 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 aspects 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.

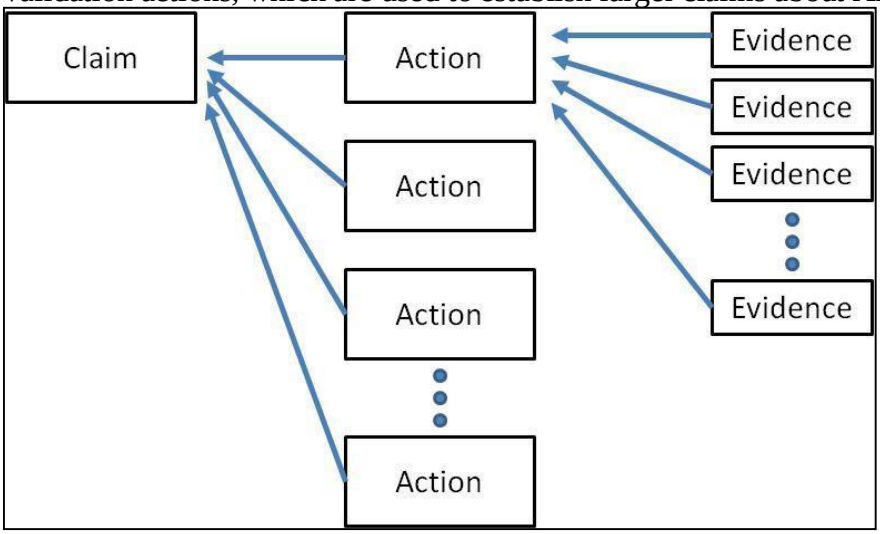


Figure 2A: General Argument Structure for Assessment Validation

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.

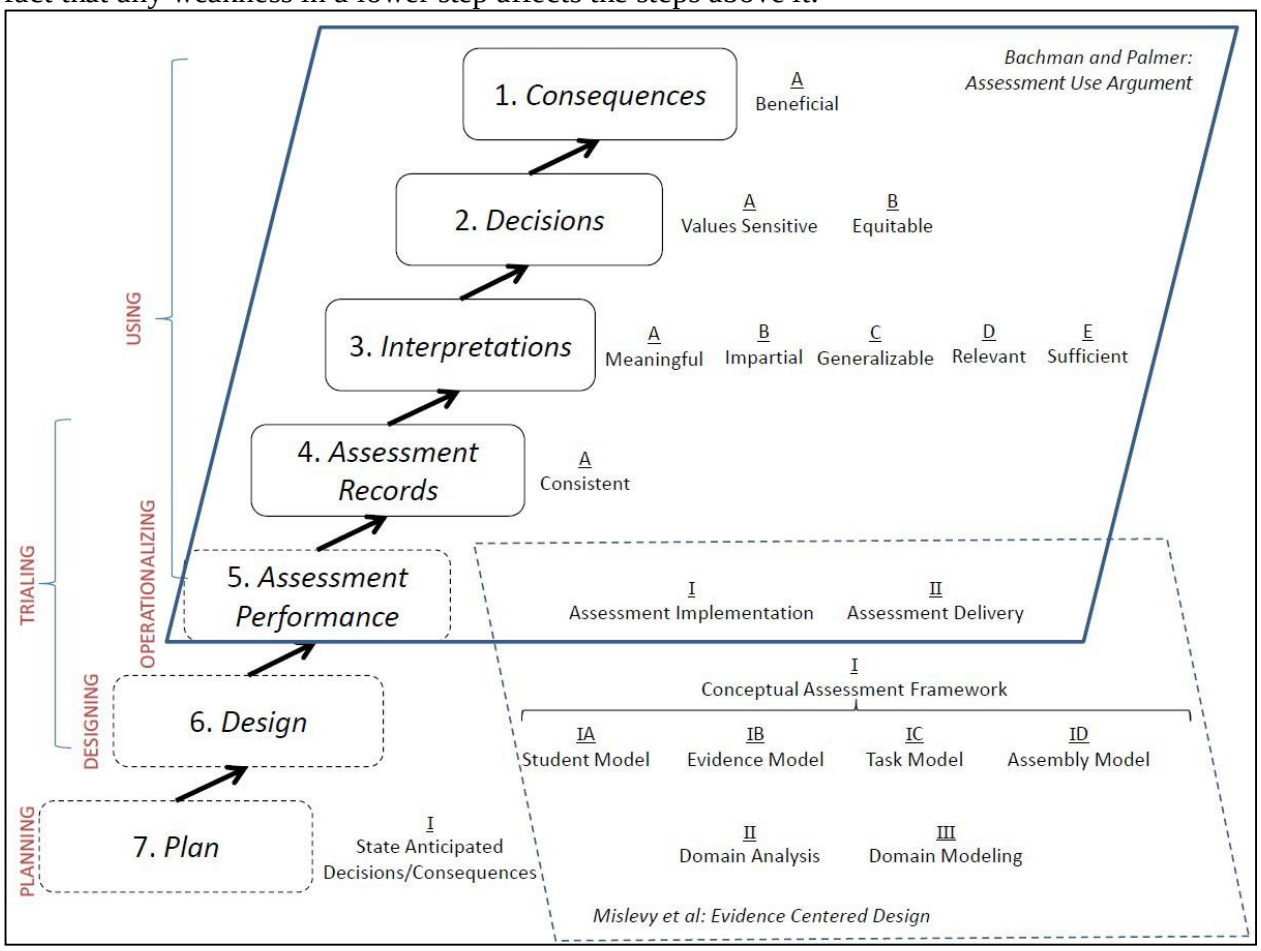


Figure 2.1A: CAL's Validation Framework (based on Bachman & Palmer, 2010; Mislevy, Almond, & Lukas, 2004)

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.

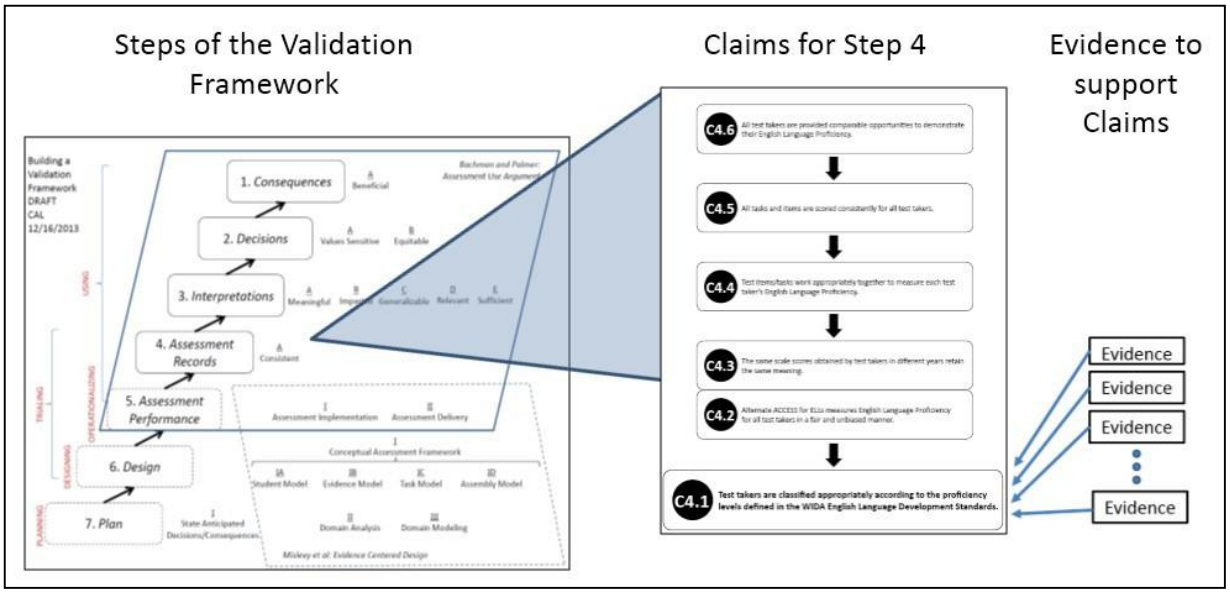


Figure 2.2A: Structure of the Argument-Based Approach Supporting Step 4 Contained in this Chapter

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.

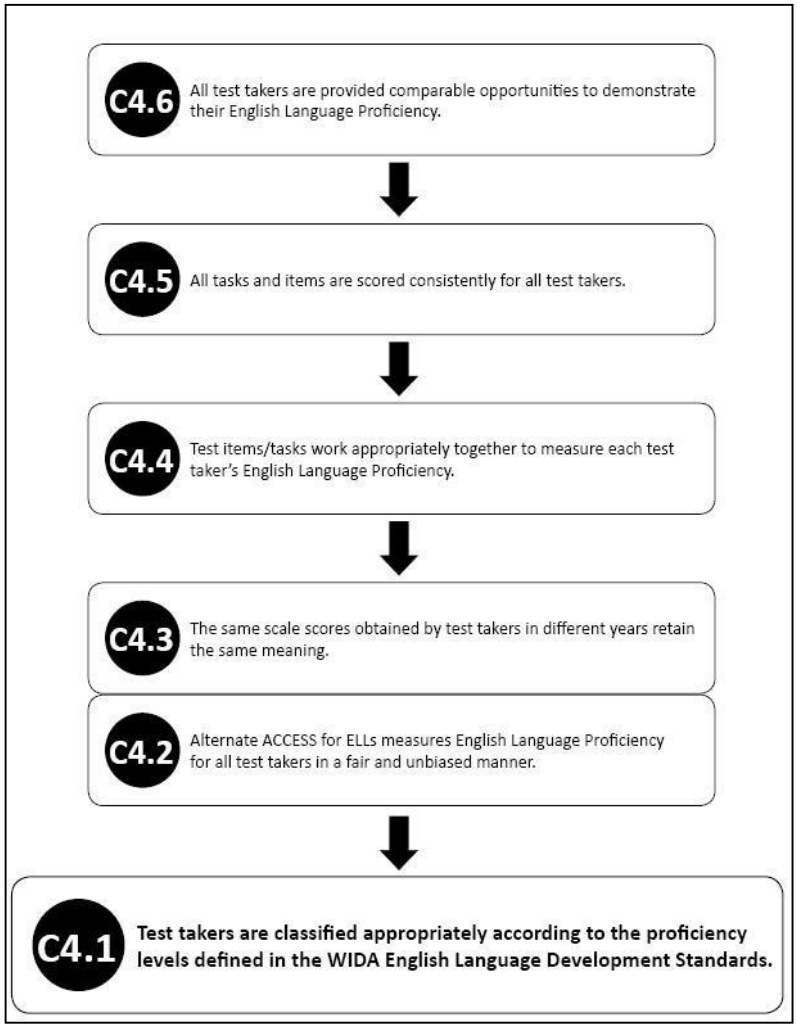


Figure 2.2.1A: Progression of Claims for Step 4: Assessment Records

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 ACCESS for ELLs 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. 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.

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 ACCESS for ELLs Test Administration Manual.

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 ACCESS for ELLs Test Administration Manual.

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-Level 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 contain 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-level cluster 9–12; two, select; three, select Reading under Domains. Information for 9–12 Reading is shown in section 6.5.2.3. Finally, look at the key that explains that reliability information is located in table F. The result is Table 6.5.2.3F. Click on 6.5.2.3 to go to the appropriate section, and then locate Table F.

2.5.1 Chapter 4 Visual Guide to Tables and Figures

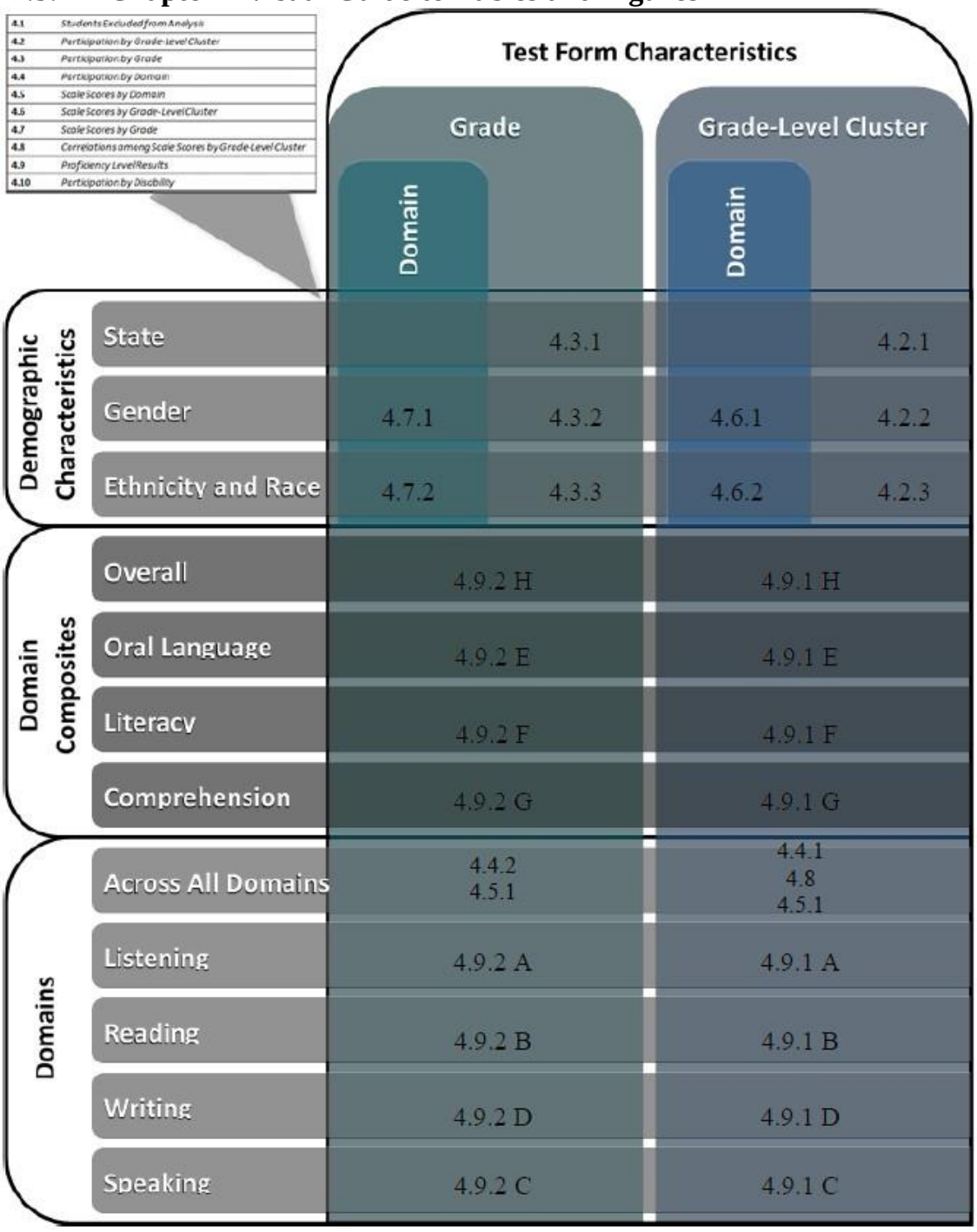


Figure 2.5.1 Chapter 4 Visual Guide to Tables and Figures

2.5.2 Chapter 6 Visual Guide to Tables and Figures

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

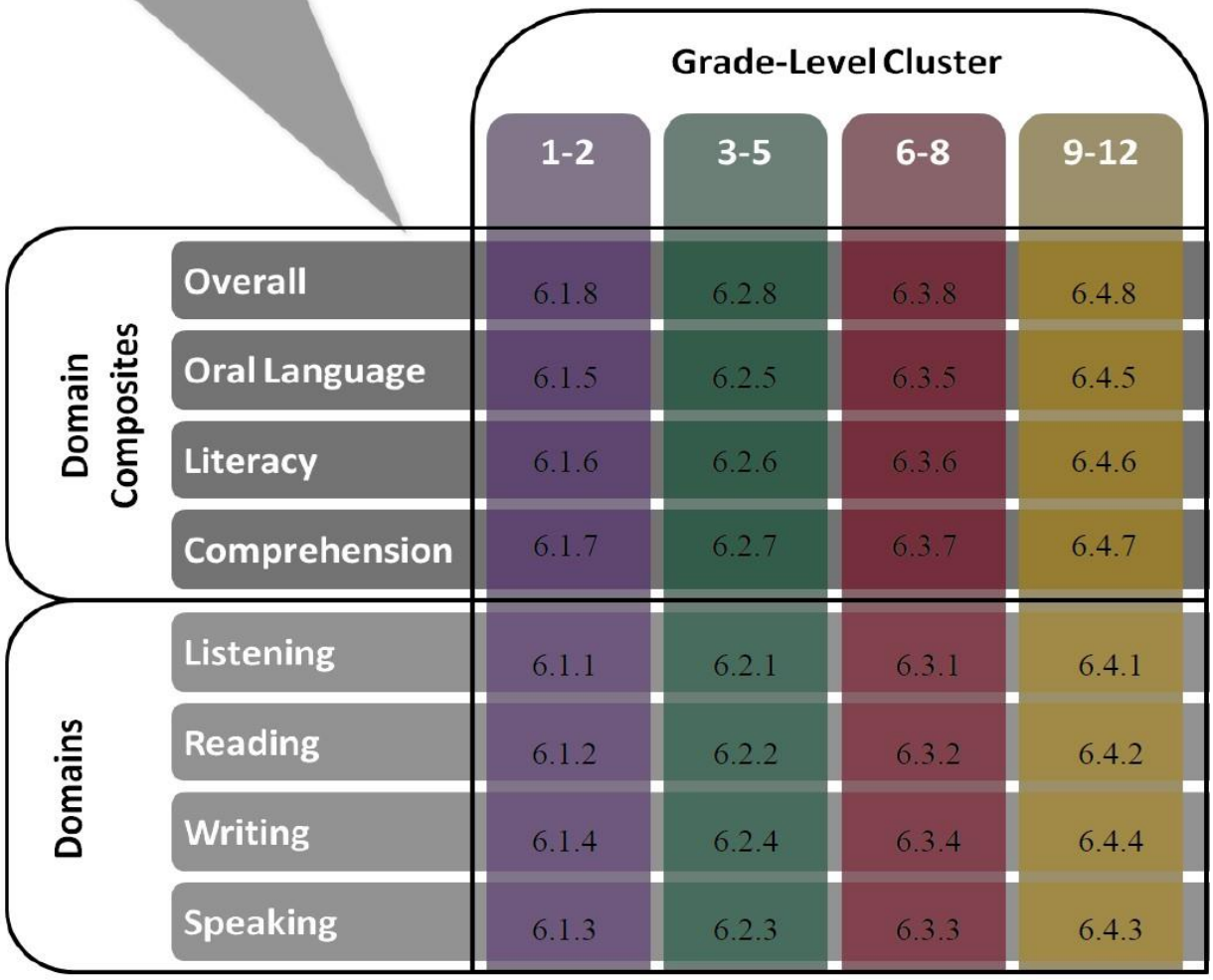


Figure 2.5.2 Chapter 6 Visual Guide to Tables and Figures

3. Descriptions of Student Results

Chapter 3 provides a description of the Chapter 4 tables summarizing students' participation, scale scores, and proficiency levels; results are further subdivided by grade, grade-level cluster, state, domain, domain and composite scores, gender, ethnicity/race, and disability. The 41 WIDA Consortium states/territories participated in the 2022-2023 Alternate ACCESS operational administration. The final number of students that have taken the 601 Alternate ACCESS tests is 29,301 as of December 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 2022-2023 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

	Cluster				
Grade	**1-2**	**3-5**	**6-8**	**9-12**	**Total**
1		1	0	0	1
2		8	0	0	8
3	3		0	0	3
4	1		0	0	1
5	0		3	0	3
6	0	4		0	4
7	0	0		0	0
8	0	2		4	6
9	0	0	4		4
10	0	0	0		0
11	0	0	0		0
12	0	0	0		0
Total	4	15	7	4	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				Total
	1-2	3-5	6-8	9-12	
AK	10	19	24	46	99
AL	57	119	69	67	312
BI	6	12	14	14	46
CO	139	268	216	211	834
DC	20	30	35	36	121
DD	3	8	4	4	19
DE	8	4	5	6	23
FL	282	403	166	151	1002
GA	251	446	317	312	1326
HI	51	66	61	86	264
ID	27	35	47	40	149
IL	1096	1373	1087	1453	5009
IN	231	318	302	489	1340
KY	106	119	88	104	417
MA	469	503	369	428	1769
MD	146	240	188	176	750
ME	16	25	5	33	79
MI	158	250	186	158	752
MN	325	416	256	298	1295
MO	53	65	45	45	208
MP	1	0	1	0	2
MT	5	8	8	1	22
NC	281	447	512	575	1815
ND	3	7	3	9	22
NH	8	8	8	14	38
NJ	216	197	100	69	582
NM	82	162	130	130	504
NV	113	256	276	382	1027
OK	151	263	226	169	809
PA	410	440	338	341	1529
RI	46	57	51	63	217
SC	126	142	82	112	462
SD	10	14	13	16	53
TN	113	165	108	99	485
UT	93	181	166	169	609
VA	602	647	483	737	2469
VI	0	1	0	0	1
VT	9	9	2	7	27
WA	539	637	432	625	2233
WI	56	139	146	211	552
WY	3	12	5	9	29
Total	6321	8511	6574	7895	29301

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	1521	24.06%	3815	60.35%	985	15.58%	6321
3-5	2335	27.44%	4904	57.62%	1272	14.95%	8511
6-8	1943	29.56%	3507	53.35%	1124	17.1%	6574
9-12	2254	28.55%	4132	52.34%	1509	19.11%	7895
Total	8053	27.48%	16358	55.83%	4890	16.69%	29301

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	3571	56.49%	2238	35.40%	512	8.09%	6321
3-5	5039	59.20%	2873	33.75%	599	7.03%	8511
6-8	4212	64.07%	1937	29.46%	425	6.46%	6574
9-12	5074	64.26%	2306	29.20%	515	6.52%	7895
Total	17896	61.07%	9354	31.92%	2051	6.99%	29301

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	7	3	4	6	9	12	4	8	19	7	12	8	99
AL	30	27	45	34	40	25	24	20	17	26	12	12	312
BI	2	4	5	5	2	5	2	7	7	1	2	4	46
CO	67	72	81	107	80	80	74	62	62	53	45	51	834
DC	6	14	8	13	9	17	8	10	7	7	5	17	121
DD	3	.	2	4	2	4	.	.	2	.	2	.	19
DE	3	5	1	1	2	3	.	2	1	.	2	3	23
FL	143	139	169	124	110	83	48	35	39	40	31	41	1002
GA	127	124	160	148	138	113	94	110	92	80	40	100	1326
HI	30	21	20	21	25	26	20	15	17	22	15	32	264
ID	9	18	8	15	12	17	17	13	17	11	4	8	149
IL	572	524	520	458	395	369	364	354	330	305	291	527	5009
IN	101	130	111	120	87	105	83	114	116	90	105	178	1340
KY	66	40	50	37	32	29	29	30	29	26	23	26	417
MA	243	226	180	201	122	123	124	122	124	117	79	108	1769
MD	72	74	69	90	81	70	71	47	56	47	34	39	750
ME	9	7	8	6	11	2	2	1	7	9	7	10	79
MI	78	80	87	86	77	60	64	62	35	41	48	34	752
MN	191	134	158	146	112	84	89	83	62	64	54	118	1295
MO	23	30	29	19	17	8	21	16	17	4	10	14	208
MP	.	1	.	.	.	1	.	.	.	.	.	.	2
MT	2	3	2	4	2	3	2	3	.	1	.	.	22
NC	138	143	137	159	151	159	187	166	137	125	106	207	1815
ND	2	1	2	2	3	1	2	.	3	.	3	3	22
NH	8	.	4	2	2	2	1	5	3	5	3	3	38
NJ	131	85	85	56	56	41	41	18	17	13	19	20	582
NM	37	45	51	63	48	47	43	40	42	30	21	37	504
NV	57	56	100	73	83	89	91	96	96	93	86	107	1027
OK	77	74	97	83	83	76	80	70	53	38	44	34	809
PA	222	188	166	134	140	107	116	115	74	70	69	128	1529
RI	24	22	16	24	17	18	19	14	18	14	13	18	217
SC	67	59	50	45	47	34	30	18	29	24	25	34	462
SD	6	4	5	3	6	4	6	3	1	4	1	10	53
TN	61	52	71	54	40	45	38	25	36	22	24	17	485
UT	41	52	63	50	68	61	54	51	47	40	40	42	609
VA	339	263	239	216	192	161	166	156	165	153	120	299	2469
VI	.	.	1	.	.	.	.	.	.	.	.	.	1
VT	5	4	5	3	1	1	1	.	1	2	2	2	27
WA	313	226	235	203	199	159	119	154	144	137	104	240	2233
WI	24	32	44	41	54	52	44	50	48	35	37	91	552
WY	2	1	4	3	5	3	1	1	1	2	4	2	29

State	Grade												
	1	2	3	4	5	6	7	8	9	10	11	12	Total
Total	3338	2983	3092	2859	2560	2299	2179	2096	1971	1758	1542	2624	29301

4.3.2 Participation by Grade by Gender

Table 4.3.2

Participation by Grade by Gender

Grade	Gender						Total
	Female		Male		Missing		
	Count	% within Grade	Count	% within Grade	Count	% within Grade	
1	787	23.58%	1998	59.86%	553	16.57%	3338
2	734	24.61%	1817	60.91%	432	14.48%	2983
3	853	27.59%	1795	58.05%	444	14.36%	3092
4	746	26.09%	1692	59.18%	421	14.73%	2859
5	736	28.75%	1417	55.35%	407	15.9%	2560
6	705	30.67%	1200	52.2%	394	17.14%	2299
7	628	28.82%	1187	54.47%	364	16.7%	2179
8	610	29.1%	1120	53.44%	366	17.46%	2096
9	581	29.48%	1043	52.92%	347	17.61%	1971
10	501	28.5%	946	53.81%	311	17.69%	1758
11	455	29.51%	824	53.44%	263	17.06%	1542
12	717	27.32%	1319	50.27%	588	22.41%	2624
Total	8053	27.48%	16358	55.83%	4890	16.69%	29301

4.3.3 Participation by Grade by Ethnicity

Table 4.3.3

Participation by Grade by Ethnicity

Grade	Hispanic/Non-Hispanic						Total
	Hispanic		Non-Hispanic		Missing		
	Count	% within Grade	Count	% within Grade	Count	% within Grade	
1	1861	55.75%	1184	35.47%	293	8.78%	3338
2	1710	57.32%	1054	35.33%	219	7.34%	2983
3	1783	57.66%	1085	35.09%	224	7.24%	3092
4	1678	58.69%	988	34.56%	193	6.75%	2859
5	1578	61.64%	800	31.25%	182	7.11%	2560
6	1444	62.81%	699	30.40%	156	6.79%	2299

Grade	Hispanic/Non-Hispanic						Total
	Hispanic		Non-Hispanic		Missing		
	Count	% within Grade	Count	% within Grade	Count	% within Grade	
7	1398	64.16%	645	29.60%	136	6.24%	2179
8	1370	65.36%	593	28.29%	133	6.35%	2096
9	1311	66.51%	521	26.43%	139	7.05%	1971
10	1165	66.27%	488	27.76%	105	5.97%	1758
11	992	64.33%	456	29.57%	94	6.09%	1542
12	1606	61.20%	841	32.05%	177	6.75%	2624
Total	17896	61.08%	9354	31.92%	2051	6.99%	29301

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	6268	6218	6153	6145
3-5	8455	8420	8369	8301
6-8	6524	6498	6428	6378
9-12	7850	7815	7727	7655
Total	29097	28951	28677	28479

4.4.2 Participation by Grade by Domain

Table 4.4.2

Participation by Grade by Domain

Grade	Domain			
	Listening	Reading	Speaking	Writing
1	3314	3292	3257	3250
2	2954	2926	2896	2895
3	3067	3050	3033	3011
4	2842	2834	2815	2794
5	2546	2536	2521	2496
6	2286	2273	2244	2224
7	2160	2150	2129	2114
8	2078	2075	2055	2040

Grade	Domain			
	Listening	Reading	Speaking	Writing
9	1960	1947	1929	1908
10	1748	1745	1718	1715
11	1534	1525	1515	1500
12	2608	2598	2565	2532
Total	29097	28951	28677	28479

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	929.85	11.55	3314	932.21	11.11	2954	930.96	11.41	6268
Reading	930.29	13.27	3292	933.12	13.17	2926	931.62	13.3	6218
Speaking	930.07	15.03	3257	932.64	14.59	2896	931.28	14.88	6153
Writing	926.01	11.19	3250	928.38	11.71	2895	927.13	11.5	6145
Oral	930.29	12.52	3247	932.83	12.04	2888	931.49	12.36	6135
Literacy	928.49	11.34	3229	931.13	11.56	2867	929.73	11.52	6096
Comprehension	930.25	12.36	3284	932.97	12.15	2918	931.53	12.34	6202
Overall	928.86	11.25	3198	931.46	11.23	2835	930.08	11.32	6033

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	934.33	10.96	3067	935.71	10.62	2842	936.97	10.44	2546	935.59	10.75	8455
Reading	933.46	11.35	3050	935.03	10.99	2834	936.17	10.78	2536	934.81	11.12	8420
Speaking	933.98	14.09	3033	934.97	13.63	2815	936.27	13.37	2521	935	13.75	8369
Writing	929.62	11.39	3011	931.04	11.58	2794	932.24	11.66	2496	930.88	11.59	8301
Oral	934.38	11.63	3025	935.54	11.23	2806	936.82	11.01	2509	935.5	11.35	8340
Literacy	931.84	10.65	2985	933.35	10.64	2787	934.6	10.51	2485	933.18	10.66	8257
Comprehension	933.79	10.86	3043	935.32	10.52	2821	936.47	10.36	2529	935.11	10.65	8393
Overall	932.42	10.57	2961	933.85	10.43	2765	935.12	10.29	2466	933.72	10.5	8192

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	936.25	10.64	2286	937.1	10.5	2160	937.66	10.36	2078	936.98	10.52	6524
Reading	936.48	11.92	2273	937.58	12.01	2150	938.16	11.89	2075	937.38	11.96	6498
Speaking	935.37	13.26	2244	936.07	13.35	2129	936.18	13.22	2055	935.86	13.28	6428
Writing	930.77	10.07	2224	932.01	10.55	2114	932.43	10.58	2040	931.71	10.42	6378
Oral	936.26	11.27	2235	937.06	11.32	2122	937.39	11.07	2045	936.89	11.23	6402
Literacy	933.95	10.23	2210	935.14	10.54	2099	935.63	10.48	2029	934.88	10.44	6338
Comprehension	936.52	11.1	2263	937.53	11.18	2141	938.1	11.06	2064	937.36	11.13	6468
Overall	934.5	10.15	2187	935.54	10.39	2081	936.01	10.24	2008	935.33	10.27	6276

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	937.79	10.2	1748	937.89	10.21	1534	937.87	10.14	2608	936.98	10.57	1960	937.63	10.28	7850
Reading	937.61	10.86	1745	937.89	10.72	1525	937.81	10.83	2598	936.82	11.05	1947	937.54	10.88	7815
Speaking	935.86	12.4	1718	936.24	12.13	1515	935.75	12.32	2565	935.08	12.77	1929	935.7	12.42	7727
Writing	932.79	10.4	1715	933.41	10.83	1500	933.33	10.78	2532	932.27	10.73	1908	932.96	10.7	7655
Oral	937.06	10.5	1715	937.31	10.36	1511	937.08	10.44	2559	936.3	10.8	1921	936.93	10.53	7706
Literacy	935.5	9.99	1709	935.92	10.08	1493	935.89	10.08	2523	934.82	10.19	1898	935.54	10.09	7623
Comprehension	937.79	10.35	1738	938.04	10.28	1519	937.95	10.39	2588	936.99	10.61	1943	937.69	10.42	7788
Overall	935.89	9.82	1690	936.22	9.83	1480	936.15	9.86	2499	935.19	10.04	1878	935.87	9.9	7547

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	930.86	11.33	1510	931.08	11.51	3777	930.68	11.11	981
Reading	930.93	12.67	1504	932.05	13.63	3747	931	12.89	967
Speaking	931.01	14.79	1485	931.62	14.84	3712	930.4	15.13	956
Writing	926.28	10.97	1482	927.64	11.76	3705	926.46	11.17	958
Oral	931.32	12.23	1481	931.7	12.44	3699	930.9	12.23	955
Literacy	928.97	10.96	1473	930.21	11.82	3674	929.08	11.08	949
Comprehension	931.02	11.87	1501	931.86	12.61	3735	931.04	11.94	966
Overall	929.52	10.88	1458	930.47	11.55	3637	929.45	10.98	938

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	936.06	10.46	2319	935.49	10.78	4867	935.12	11.11	1269
Reading	934.66	10.68	2315	935.02	11.24	4837	934.24	11.4	1268
Speaking	935.18	13.61	2300	935.1	13.72	4822	934.32	14.11	1247
Writing	930.47	11.21	2269	931.27	11.68	4782	930.15	11.84	1250
Oral	935.8	11.19	2291	935.51	11.35	4803	934.93	11.67	1246
Literacy	932.96	10.26	2260	933.47	10.77	4747	932.49	10.92	1250
Comprehension	935.16	10.27	2306	935.24	10.75	4822	934.54	10.95	1265
Overall	933.67	10.16	2240	933.91	10.59	4716	933.07	10.73	1236

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	937.47	10.36	1927	936.77	10.67	3479	936.8	10.31	1118
Reading	937.14	11.73	1919	937.54	12.12	3464	937.3	11.83	1115
Speaking	935.91	13.21	1907	935.99	13.24	3423	935.38	13.51	1098
Writing	931.67	10.32	1885	931.85	10.58	3402	931.34	10.09	1091
Oral	937.15	11.12	1896	936.85	11.31	3413	936.54	11.19	1093
Literacy	934.71	10.31	1873	935.04	10.58	3379	934.68	10.2	1086
Comprehension	937.37	10.91	1910	937.41	11.3	3445	937.19	11.01	1113
Overall	935.28	10.15	1857	935.43	10.4	3344	935.12	10.11	1075

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	937.38	10.47	2241	937.88	10.27	4102	937.32	10.01	1507
Reading	937.08	11.04	2228	937.85	10.87	4087	937.35	10.64	1500
Speaking	935.5	12.49	2206	935.99	12.3	4042	935.23	12.64	1479
Writing	932.23	10.84	2184	933.31	10.52	4013	933.09	10.93	1458
Oral	936.71	10.63	2203	937.19	10.47	4026	936.52	10.52	1477
Literacy	934.94	10.28	2172	935.87	9.98	3996	935.55	10.1	1455
Comprehension	937.31	10.61	2220	937.99	10.41	4069	937.44	10.17	1499
Overall	935.39	10.04	2149	936.17	9.81	3950	935.74	9.91	1448

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
Hispanic (of any Race)	Mean	931.19	931.64	931.15	926.71	931.55	929.54	931.61	929.97
	Std.	11.42	13.06	14.78	11.13	12.33	11.27	12.21	11.15
	N	3534	3503	3466	3461	3456	3428	3492	3395
Non-Hispanic American Indian	Mean	934.84	937.74	935.5	929.78	935.41	934.39	937.13	934.65
	Std.	10.36	11.36	13.38	11.13	10.96	10.2	10.45	9.8
	N	32	31	32	32	32	31	31	31
Non-Hispanic Asian	Mean	930.28	931.67	931.16	928.56	931.08	930.43	931.38	930.47
	Std.	11.56	14.49	15.13	12.49	12.51	12.52	13.19	12.03
	N	953	945	934	936	933	932	943	920
Non-Hispanic Black	Mean	929.66	931.01	931.06	927.17	930.68	929.48	930.75	929.71
	Std.	11.49	13.63	15.29	11.84	12.56	11.79	12.54	11.63
	N	602	598	591	589	587	586	597	579
Non-Hispanic Multiracial	Mean	928.46	928.78	926.95	925.05	928.36	927.41	928.87	927.45
	Std.	10.95	14.76	15.98	11.49	12.37	12.15	13.15	11.63
	N	46	45	44	44	44	44	45	44
Non-Hispanic Pacific Islander	Mean	930.08	928.45	928.16	924.06	929.4	926.57	928.94	927.16
	Std.	12.15	13.67	15.36	11.32	13.13	11.45	12.96	11.46
	N	50	49	50	50	50	49	49	49
Non-Hispanic White	Mean	931.03	931.56	931.32	926.5	931.55	929.42	931.48	929.84
	Std.	11.23	12.75	15.03	11.65	12.3	11.36	11.84	11.14
	N	541	538	529	535	527	529	537	520
Missing	Mean	932.18	932.31	933.08	928.27	932.9	930.59	932.37	931.07
	Std.	10.9	12.57	14.2	11.23	11.96	10.97	11.69	10.82
	N	510	509	507	498	506	497	508	495

Table 4.6.2 B

Mean Scale Scores by Ethnicity: 3-5

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	936.1	934.98	935.08	930.65	935.81	933.19	935.38	933.83
	Std.	10.71	10.94	13.78	11.32	11.33	10.44	10.53	10.33
	N	4997	4975	4943	4904	4922	4870	4957	4830
Non-Hispanic American Indian	Mean	937.67	938.05	936.69	933.34	937.45	936.23	938.12	936.47
	Std.	9.83	9.38	13	10.44	10.43	9.42	9.11	9.32
	N	58	57	59	59	58	57	57	57
Non-Hispanic Asian	Mean	934.59	934.79	934.79	931.79	934.89	933.58	934.76	933.78
	Std.	10.4	11.33	13.51	12.05	11.01	10.98	10.67	10.6
	N	1254	1253	1244	1241	1243	1239	1250	1231
Non-Hispanic Black	Mean	933.79	933.49	934.33	930.3	934.23	932.13	933.66	932.5
	Std.	11.14	11.7	14.08	12	11.78	11.15	11.16	11.01
	N	636	635	629	623	627	620	632	614
Non-Hispanic Multiracial	Mean	934.78	934.59	935.88	930.95	935.51	933.14	934.95	934.14
	Std.	11.91	12.65	13.77	13.02	12.56	12.27	12.09	12.08
	N	59	59	58	58	57	58	58	56
Non-Hispanic Pacific Islander	Mean	933.27	932.06	932.48	927.77	932.95	930.08	932.41	930.72
	Std.	12.43	12.21	14.5	12.5	12.7	11.77	11.95	11.73
	N	63	63	62	61	62	61	63	61
Non-Hispanic White	Mean	935.1	934.28	934.55	930.42	935	932.63	934.58	933.15
	Std.	10.73	11.17	13.87	11.83	11.43	10.83	10.7	10.64
	N	790	785	782	771	780	769	784	763
Missing	Mean	936.15	935.5	936.18	932.19	936.35	934.21	935.81	934.66
	Std.	10.8	11.1	13.42	11.7	11.34	10.72	10.66	10.52
	N	598	593	592	584	591	583	592	580

Table 4.6.2 C

Mean Scale Scores by Ethnicity: 6-8

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	937.59	937.95	936.29	931.87	937.42	935.25	937.93	935.74
	Std.	10.16	11.54	13.04	9.98	10.9	10	10.73	9.87
	N	4183	4160	4119	4086	4104	4064	4142	4028
Non-Hispanic American Indian	Mean	940.69	940.5	939.77	934.58	940.67	937.79	940.6	938.33
	Std.	9.66	10.16	13.02	9.12	10.6	8.73	9.83	8.92
	N	52	52	52	52	52	52	52	52
Non-Hispanic Asian	Mean	935.04	935.82	933.97	931.17	934.91	933.83	935.68	934.01
	Std.	11.37	12.95	14.05	11.24	12	11.36	12.1	11.18
	N	807	805	792	785	790	779	801	769
Non-Hispanic Black	Mean	935.76	936.17	934.77	931.11	935.77	934.04	936.11	934.42
	Std.	10.74	12.29	13.46	11.42	11.49	11.15	11.5	10.92
	N	435	433	429	426	428	422	432	419
Non-Hispanic Multiracial	Mean	938.28	939.34	937.58	931.69	938.31	935.8	939.09	936.29
	Std.	9.29	11.07	11.49	12.13	9.8	11.18	10.21	10.44
	N	36	35	36	36	36	35	35	35
Non-Hispanic Pacific Islander	Mean	935.47	935.1	936.26	931	936.09	933.48	935.79	934.59
	Std.	11.32	13.26	13.61	10.63	12.15	10.82	12.07	10.6
	N	57	59	58	58	57	58	57	56
Non-Hispanic White	Mean	935.04	935.45	934.41	930.23	935.22	933.15	935.51	933.7
	Std.	11.32	12.94	13.7	11.16	11.86	11.25	11.97	10.98
	N	534	532	525	522	520	518	529	510
Missing	Mean	938.04	938.25	937.46	933.33	938.14	936.04	938.2	936.44
	Std.	10.25	11.87	12.96	10.71	11.17	10.68	11.07	10.45
	N	420	422	417	413	415	410	420	407

Table 4.6.2 D

Mean Scale Scores by Ethnicity: 9-12

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	937.86	937.68	935.5	932.94	936.94	935.61	937.87	935.91
	Std.	10.27	10.86	12.6	10.65	10.6	10.05	10.41	9.9
	N	5044	5022	4970	4929	4955	4907	5005	4862
Non-Hispanic American Indian	Mean	939.87	938.78	938.88	934.1	939.71	936.67	939.24	937.59
	Std.	9.63	10.69	10.23	10.5	9.4	10.1	10.1	9.47
	N	55	54	52	51	52	51	54	49
Non-Hispanic Asian	Mean	936.58	937.05	935.05	932.8	936.1	935.22	937.02	935.42
	Std.	10.32	11.06	12.55	10.96	10.52	10.3	10.55	10.02
	N	990	985	970	964	969	961	985	951
Non-Hispanic Black	Mean	937.01	936.87	936	932.76	936.73	935.07	936.97	935.42
	Std.	10.6	11.16	12.37	10.85	10.77	10.4	10.76	10.2
	N	549	546	542	535	542	533	543	527
Non-Hispanic Multiracial	Mean	937.34	936.26	935.62	934.38	936.71	935.68	936.66	935.88
	Std.	9.21	11.27	10.99	11.95	9.44	11.05	10.38	10.45
	N	35	35	34	34	34	34	35	34
Non-Hispanic Pacific Islander	Mean	936.42	937.55	938.11	934.12	937.68	935.91	937.31	936.69
	Std.	10.16	11.49	10.81	9.75	9.7	10.1	10.9	9.46
	N	67	67	65	66	65	66	67	64
Non-Hispanic White	Mean	936.94	936.83	935.7	932.06	936.59	934.82	936.97	935.27
	Std.	10.53	11.12	12.36	11.38	10.67	10.53	10.69	10.25
	N	602	600	590	580	589	579	597	573
Missing	Mean	938.84	938.55	937.97	934.38	938.62	936.72	938.81	937.15
	Std.	9.64	9.9	10.67	9.65	9.49	9.11	9.47	8.86
	N	508	506	504	496	500	492	502	487

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	929.47	11.6	779	930.01	11.6	1985	929.79	11.29	550	929.85	11.55	3314
Reading	929.75	12.83	778	930.82	13.63	1970	929.11	12.42	544	930.29	13.27	3292
Speaking	929.9	15.01	768	930.35	15.03	1951	929.32	15.06	538	930.07	15.03	3257
Writing	925.39	10.81	772	926.41	11.44	1939	925.47	10.79	539	926.01	11.19	3250
Oral	930.01	12.53	766	930.51	12.56	1944	929.89	12.35	537	930.29	12.52	3247
Literacy	927.92	10.9	765	928.95	11.63	1932	927.63	10.8	532	928.49	11.34	3229
Comprehension	929.76	12.08	776	930.67	12.64	1965	929.46	11.68	543	930.25	12.36	3284
Overall	928.39	10.92	756	929.23	11.48	1918	928.21	10.85	524	928.86	11.25	3198

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	932.33	10.85	731	932.25	11.3	1792	931.8	10.8	431	932.21	11.11	2954
Reading	932.19	12.37	726	933.42	13.49	1777	933.44	13.08	423	933.12	13.17	2926
Speaking	932.21	14.46	717	933.01	14.51	1761	931.79	15.11	418	932.64	14.59	2896
Writing	927.25	11.08	710	928.99	11.95	1766	927.73	11.53	419	928.38	11.71	2895
Oral	932.72	11.76	715	933.02	12.17	1755	932.18	11.98	418	932.83	12.04	2888
Literacy	930.12	10.93	708	931.6	11.88	1742	930.92	11.17	417	931.13	11.56	2867
Comprehension	932.37	11.51	725	933.19	12.45	1770	933.06	11.98	423	932.97	12.15	2918
Overall	930.73	10.72	702	931.86	11.48	1719	931.02	10.96	414	931.46	11.23	2835

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	934.75	10.6	844	934.21	11	1780	934	11.46	443	934.33	10.96	3067
Reading	933.39	10.71	841	933.58	11.55	1767	933.13	11.75	442	933.46	11.35	3050
Speaking	934.31	14.01	840	933.93	14.1	1759	933.57	14.21	434	933.98	14.09	3033
Writing	929.13	10.8	830	930.08	11.53	1743	928.71	11.88	438	929.62	11.39	3011
Oral	934.7	11.44	836	934.31	11.63	1755	934.04	12	434	934.38	11.63	3025
Literacy	931.62	10.05	822	932.12	10.82	1725	931.18	11.08	438	931.84	10.65	2985
Comprehension	933.89	10.33	837	933.84	11.01	1765	933.43	11.27	441	933.79	10.86	3043
Overall	932.36	10.11	813	932.59	10.7	1717	931.89	10.93	431	932.42	10.57	2961

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	936.18	10.41	743	935.81	10.55	1679	934.51	11.22	420	935.71	10.62	2842
Reading	934.54	10.76	744	935.47	10.95	1671	934.12	11.48	419	935.03	10.99	2834
Speaking	934.86	13.49	737	935.32	13.5	1665	933.77	14.36	413	934.97	13.63	2815
Writing	930.1	11.34	730	931.66	11.67	1651	930.21	11.54	413	931.04	11.58	2794
Oral	935.71	11.12	735	935.78	11.13	1658	934.28	11.78	413	935.54	11.23	2806
Literacy	932.62	10.41	730	933.91	10.68	1644	932.44	10.76	413	933.35	10.64	2787
Comprehension	935.1	10.3	741	935.68	10.47	1662	934.26	11.06	418	935.32	10.52	2821
Overall	933.42	10.23	724	934.31	10.45	1631	932.77	10.64	410	933.85	10.43	2765

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	937.46	10.16	732	936.72	10.6	1408	936.97	10.39	406	936.97	10.44	2546
Reading	936.25	10.37	730	936.3	10.99	1399	935.56	10.8	407	936.17	10.78	2536
Speaking	936.51	13.17	723	936.31	13.39	1398	935.69	13.66	400	936.27	13.37	2521
Writing	932.42	11.32	709	932.32	11.76	1388	931.66	11.93	399	932.24	11.66	2496
Oral	937.16	10.82	720	936.71	11.1	1390	936.56	11.04	399	936.82	11.01	2509
Literacy	934.86	10.08	708	934.65	10.66	1378	933.97	10.75	399	934.6	10.51	2485
Comprehension	936.67	9.98	728	936.48	10.57	1395	936.04	10.33	406	936.47	10.36	2529
Overall	935.44	9.89	703	935.09	10.45	1368	934.66	10.42	395	935.12	10.29	2466

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	10.44	700	936	10.85	1194	935.72	10.31	392	936.25	10.64	2286
Reading	936.38	11.74	697	936.78	12.01	1185	935.76	11.97	391	936.48	11.92	2273
Speaking	935.35	13.18	693	935.86	13.05	1166	933.94	13.93	385	935.37	13.26	2244
Writing	930.5	10.01	686	931	10.23	1161	930.54	9.71	377	930.77	10.07	2224
Oral	936.62	11.13	689	936.4	11.31	1163	935.19	11.38	383	936.26	11.27	2235
Literacy	933.7	10.19	682	934.22	10.34	1152	933.58	9.92	376	933.95	10.23	2210
Comprehension	936.72	10.86	693	936.64	11.27	1180	935.81	11.05	390	936.52	11.1	2263
Overall	934.41	10.11	675	934.75	10.24	1138	933.93	9.94	374	934.5	10.15	2187

Table 4.7.1 G

Mean Scale Scores by Gender: Grade 7

	Female			Male			Missing			Total		
	Mean	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**
Listening	937.44	10.26	624	936.93	10.57	1174	937.1	10.71	362	937.1	10.5	2160
Reading	937.28	11.72	618	937.6	12.17	1171	938.02	11.99	361	937.58	12.01	2150
Speaking	936.08	13.36	615	935.94	13.38	1157	936.5	13.24	357	936.07	13.35	2129
Writing	932.33	10.61	605	931.96	10.57	1151	931.6	10.37	358	932.01	10.55	2114
Oral	937.24	11.2	613	936.9	11.37	1153	937.27	11.37	356	937.06	11.32	2122
Literacy	935.16	10.42	601	935.13	10.62	1142	935.12	10.55	356	935.14	10.54	2099
Comprehension	937.42	10.93	617	937.51	11.29	1163	937.78	11.3	361	937.53	11.18	2141
Overall	935.59	10.25	598	935.48	10.44	1130	935.65	10.44	353	935.54	10.39	2081

Table 4.7.1 H

Mean Scale Scores by Gender: Grade 8

	Female			Male			Missing			Total		
	Mean	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**
Listening	938.1	10.34	603	937.42	10.54	1111	937.65	9.83	364	937.66	10.36	2078
Reading	937.89	11.7	604	938.29	12.16	1108	938.25	11.38	363	938.16	11.89	2075
Speaking	936.38	13.1	599	936.19	13.3	1100	935.81	13.21	356	936.18	13.22	2055
Writing	932.34	10.28	594	932.63	10.88	1090	931.93	10.17	356	932.43	10.58	2040
Oral	937.66	11.03	594	937.28	11.22	1097	937.27	10.7	354	937.39	11.07	2045
Literacy	935.43	10.26	590	935.82	10.73	1085	935.41	10.07	354	935.63	10.48	2029
Comprehension	938.06	10.92	600	938.13	11.3	1102	938.1	10.56	362	938.1	11.06	2064
Overall	935.97	10.03	584	936.09	10.48	1076	935.85	9.85	348	936.01	10.24	2008

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	936.99	10.63	578	937.29	10.66	1037	936.03	10.16	345	936.98	10.57	1960
Reading	936.41	11.23	573	937.21	11.09	1030	936.35	10.59	344	936.82	11.05	1947
Speaking	935.12	12.66	565	935.31	12.76	1024	934.31	12.97	340	935.08	12.77	1929
Writing	931.55	10.93	562	932.75	10.72	1015	932.03	10.35	331	932.27	10.73	1908
Oral	936.39	10.79	564	936.56	10.83	1019	935.39	10.71	338	936.3	10.8	1921
Literacy	934.23	10.45	559	935.25	10.16	1009	934.52	9.78	330	934.82	10.19	1898
Comprehension	936.73	10.79	572	937.34	10.67	1028	936.35	10.12	343	936.99	10.61	1943
Overall	934.86	10.16	551	935.56	10.07	999	934.63	9.7	328	935.19	10.04	1878

Table 4.7.1 J

Mean Scale Scores by Gender: Grade 10

	Female			Male			Missing			Total		
	Mean	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**
Listening	937.27	10.24	498	938.42	10.1	939	936.72	10.35	311	937.79	10.2	1748
Reading	936.95	10.84	497	938.3	10.76	937	936.59	11.12	311	937.61	10.86	1745
Speaking	935.12	12.78	490	936.61	11.96	924	934.78	12.99	304	935.86	12.4	1718
Writing	931.99	10.22	492	933.53	10.17	925	931.83	11.24	298	932.79	10.4	1715
Oral	936.41	10.62	490	937.77	10.24	921	935.97	10.92	304	937.06	10.5	1715
Literacy	934.74	9.87	489	936.21	9.81	922	934.53	10.58	298	935.5	9.99	1709
Comprehension	937.19	10.33	495	938.47	10.25	932	936.71	10.58	311	937.79	10.35	1738
Overall	935.1	9.75	485	936.62	9.61	908	934.94	10.38	297	935.89	9.82	1690

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	937.47	10.59	455	937.93	10.24	816	938.47	9.44	263	937.89	10.21	1534
Reading	937.56	10.84	448	937.79	10.99	818	938.8	9.59	259	937.89	10.72	1525
Speaking	936.47	11.95	446	936.08	12.37	811	936.33	11.72	258	936.24	12.13	1515
Writing	933.19	10.92	439	933.15	10.79	808	934.64	10.74	253	933.41	10.83	1500
Oral	937.26	10.42	446	937.23	10.53	807	937.67	9.73	258	937.31	10.36	1511
Literacy	935.63	10.26	437	935.73	10.16	804	937.05	9.47	252	935.92	10.08	1493
Comprehension	937.72	10.48	448	937.96	10.49	812	938.85	9.24	259	938.04	10.28	1519
Overall	936.02	9.96	433	936.06	9.93	795	937.08	9.25	252	936.22	9.83	1480

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	937.73	10.42	710	937.94	10.09	1310	937.88	9.91	588	937.87	10.14	2608
Reading	937.42	11.14	710	938.07	10.68	1302	937.69	10.78	586	937.81	10.83	2598
Speaking	935.46	12.48	705	936.01	12.09	1283	935.51	12.64	577	935.75	12.32	2565
Writing	932.35	11.13	691	933.71	10.43	1265	933.67	11.07	576	933.33	10.78	2532
Oral	936.83	10.64	703	937.26	10.3	1279	936.96	10.49	577	937.08	10.44	2559
Literacy	935.22	10.43	687	936.19	9.81	1261	936.01	10.2	575	935.89	10.08	2523
Comprehension	937.61	10.74	705	938.18	10.24	1297	937.84	10.29	586	937.95	10.39	2588
Overall	935.61	10.18	680	936.42	9.63	1248	936.21	9.98	571	936.15	9.86	2499

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
Hispanic (of any Race)	Mean	930.16	930.41	930	925.61	930.43	928.35	930.43	928.82
	Std. Dev.	11.5	12.9	14.81	10.72	12.36	10.96	12.11	10.95
	N	1848	1835	1811	1807	1808	1797	1830	1780
Non-Hispanic American Indian	Mean	929.85	934.67	931.31	925.31	930.77	930.83	933.67	931.17
	Std. Dev.	14.02	13.9	15.63	11.81	13.74	11.14	13.43	11.34
	N	13	12	13	13	13	12	12	12
Non-Hispanic Asian	Mean	928.71	929.16	929.29	926.89	929.32	928.3	929.12	928.46
	Std. Dev.	11.51	14.24	15.34	12.07	12.56	12.15	12.96	11.81
	N	508	504	501	500	500	497	503	493
Non-Hispanic Black	Mean	928.98	930.32	930.13	926.3	929.87	928.73	930.09	928.96
	Std. Dev.	11.76	13.87	15.89	11.68	13.07	11.87	12.8	11.88
	N	343	342	338	337	334	335	341	330
Non-Hispanic Multiracial	Mean	925.76	926.57	926.3	926.25	926.65	927	926.38	926.75
	Std. Dev.	11.41	15.55	15.91	12.1	12.48	12.78	13.93	12.1
	N	21	21	20	20	20	20	21	20
Non-Hispanic Pacific Islander	Mean	926.68	925.89	925.89	923.14	926.54	924.71	926.18	925
	Std. Dev.	12.69	13.99	16.32	13.03	13.97	12.52	13.44	12.34
	N	28	28	28	28	28	28	28	28
Non-Hispanic White	Mean	929.92	930.78	930.52	925.76	930.59	928.72	930.61	929.04
	Std. Dev.	11.9	13.26	15.28	11.7	12.88	11.68	12.45	11.58
	N	260	258	255	259	253	254	257	250
Missing	Mean	931.39	931.47	932.02	927.16	931.97	929.54	931.49	930.02
	Std. Dev.	10.89	12.72	14.32	11.13	12.05	11.01	11.83	10.92
	N	293	292	291	286	291	286	292	285

Table 4.7.2 B

Mean Scale Scores by Ethnicity: Grade 2

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	932.31	932.99	932.4	927.91	932.78	930.85	932.91	931.24
	Std.	11.23	13.11	14.66	11.45	12.18	11.46	12.19	11.23
	N	1686	1668	1655	1654	1648	1631	1662	1615
Non-Hispanic American Indian	Mean	938.26	939.68	938.37	932.84	938.58	936.63	939.32	936.84
	Std.	4.84	9.32	11.13	9.79	7.41	9.15	7.67	8.26
	N	19	19	19	19	19	19	19	19
Non-Hispanic Asian	Mean	932.07	934.54	933.31	930.48	933.11	932.87	933.97	932.79
	Std.	11.36	14.25	14.61	12.7	12.15	12.5	12.99	11.88
	N	445	441	433	436	433	435	440	427
Non-Hispanic Black	Mean	930.58	931.94	932.3	928.33	931.75	930.47	931.62	930.71
	Std.	11.08	13.27	14.39	11.96	11.8	11.63	12.15	11.24
	N	259	256	253	252	253	251	256	249
Non-Hispanic Multiracial	Mean	930.72	930.71	927.5	924.04	929.79	927.75	931.04	928.04
	Std.	10.23	14.08	16.37	11.11	12.36	11.87	12.31	11.46
	N	25	24	24	24	24	24	24	24
Non-Hispanic Pacific Islander	Mean	934.41	931.86	931.05	925.23	933.05	929.05	932.62	930.05
	Std.	10.1	12.76	13.88	8.83	11.25	9.57	11.58	9.72
	N	22	21	22	22	22	21	21	21
Non-Hispanic White	Mean	932.06	932.29	932.06	927.2	932.43	930.05	932.28	930.59
	Std.	10.48	12.23	14.78	11.57	11.7	11.03	11.21	10.68
	N	281	280	274	276	274	275	280	270
Missing	Mean	933.24	933.45	934.5	929.75	934.16	932	933.56	932.5
	Std.	10.84	12.31	13.95	11.21	11.75	10.78	11.42	10.53
	N	217	217	216	212	215	211	216	210

Table 4.7.2 C

Mean Scale Scores by Ethnicity: Grade 3

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	934.82	933.58	934.19	929.35	934.75	931.81	934.01	932.53
	Std.	11.04	11.27	14.13	11	11.67	10.43	10.84	10.43
	N	1762	1753	1742	1731	1735	1710	1748	1695
Non-Hispanic American Indian	Mean	936.67	937.41	936.67	932.61	936.78	935.41	937.71	935.94
	Std.	10.23	8.56	11.82	7.01	10.02	7.31	8.33	7.34
	N	18	17	18	18	18	17	17	17
Non-Hispanic Asian	Mean	933.58	933.27	933.47	930.41	933.73	932.11	933.41	932.41
	Std.	10.45	11.44	13.68	11.9	11.12	10.9	10.75	10.56
	N	461	460	457	459	456	457	458	452
Non-Hispanic Black	Mean	932.57	932.59	933.33	929.82	933.11	931.33	932.68	931.6
	Std.	11.22	11.85	14.48	12.19	12.06	11.25	11.28	11.16
	N	261	260	258	256	258	255	260	253
Non-Hispanic Multiracial	Mean	933.5	933.72	935.53	928.39	934.59	931.22	933.67	932.41
	Std. Dev.	13.24	13.06	14.49	13.63	13.74	12.84	12.93	13.19
	N	18	18	17	18	17	18	18	17
Non-Hispanic Pacific Islander	Mean	932.7	931.39	929.17	927	931.17	929.18	931.74	929.36
	Std.	12.78	13.33	16.09	13	13.73	12.64	12.75	12.59
	N	23	23	23	22	23	22	23	22
Non-Hispanic White	Mean	934.15	933.48	934.16	929.11	934.35	931.56	933.76	932.2
	Std.	10.79	11.29	14.13	11.7	11.58	10.78	10.78	10.7
	N	300	297	296	290	296	289	297	288
Missing	Mean	934.38	933.89	934.1	930.59	934.45	932.57	934.1	932.92
	Std.	10.82	11.15	14.07	11.91	11.59	10.8	10.7	10.65
	N	224	222	222	217	222	217	222	217

Table 4.7.2 D

Mean Scale Scores by Ethnicity: Grade 4

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	936.23	935.1	934.78	930.57	935.72	933.17	935.54	933.81
	Std.	10.48	10.74	13.75	11.19	11.22	10.3	10.3	10.2
	N	1665	1661	1651	1640	1643	1635	1652	1622
Non-Hispanic American Indian	Mean	936.74	936.7	934.96	931.57	935.96	934.48	936.74	934.74
	Std.	11.64	11.32	14.11	11.87	11.85	11.28	11.28	11.22
	N	23	23	23	23	23	23	23	23
Non-Hispanic Asian	Mean	934.91	935.49	935.22	932.4	935.29	934.24	935.35	934.34
	Std.	10.39	11.26	13.62	12.36	11.05	11.24	10.66	10.79
	N	453	453	448	447	448	447	452	444
Non-Hispanic Black	Mean	934.02	934.06	934.8	930.48	934.58	932.54	934.04	932.87
	Std.	10.78	11.63	13.43	12.18	11.23	11.26	11.01	10.91
	N	214	215	213	211	213	210	213	208
Non-Hispanic Multiracial	Mean	931.75	932.7	933.26	930.04	932.65	931.57	932.48	931.78
	Std. Dev.	12.94	13.86	15.8	14.79	14.19	13.71	13.44	13.62
	N	24	23	23	23	23	23	23	23
Non-Hispanic Pacific Islander	Mean	930.3	930.6	930.16	926	930	928.47	930.55	928.79
	Std.	13.5	12.33	13.87	13.32	12.75	11.99	12.39	12.02
	N	20	20	19	19	19	19	20	19
Non-Hispanic White	Mean	935.12	934.36	934.67	931.26	935.06	933.14	934.66	933.53
	Std.	11.02	11.27	13.6	11.86	11.42	10.88	10.86	10.65
	N	251	248	247	242	247	242	248	240
Missing	Mean	936.71	935.82	937.28	932.78	937.14	934.63	936.28	935.24
	Std.	10.5	10.82	12.4	11.03	10.73	10.22	10.33	9.91
	N	192	191	191	189	190	188	190	186

Table 4.7.2 E

Mean Scale Scores by Ethnicity: Grade 5

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	937.39	936.42	936.39	932.22	937.09	934.75	936.76	935.32
	Std. Dev.	10.4	10.59	13.32	11.63	10.93	10.39	10.24	10.17
	N	1570	1561	1550	1533	1544	1525	1557	1513
Non-Hispanic American Indian	Mean	940	940.53	938.94	936.33	940.18	939.41	940.41	939.35
	Std. Dev.	6.2	6.98	13.03	11.24	8.74	8.16	6.15	7.98
	N	17	17	18	18	17	17	17	17
Non-Hispanic Asian	Mean	935.54	935.9	936.02	932.87	935.92	934.71	935.81	934.88
	Std. Dev.	10.25	11.09	13.02	11.69	10.7	10.55	10.43	10.23
	N	340	340	339	335	339	335	340	335
Non-Hispanic Black	Mean	935.45	934.18	935.32	930.84	935.61	932.88	934.72	933.48
	Std. Dev.	11.33	11.52	14.28	11.47	11.96	10.81	11.11	10.85
	N	161	160	158	156	156	155	159	153
Non-Hispanic Multiracial	Mean	940.41	937.89	939.56	934.88	940.29	937.29	939.65	939.38
	Std. Dev.	6.19	10.45	9.53	9	7.02	8.78	7.79	6.02
	N	17	18	18	17	17	17	17	16
Non-Hispanic Pacific Islander	Mean	936.9	934.3	938.5	930.3	937.8	932.6	935.05	934.05
	Std. Dev.	10.45	10.96	11.69	11.31	10.42	10.7	10.62	10.25
	N	20	20	20	20	20	20	20	20
Non-Hispanic White	Mean	936.27	935.19	934.9	931.16	935.73	933.39	935.51	933.91
	Std. Dev.	10.27	10.89	13.87	11.86	11.24	10.77	10.4	10.52
	N	239	240	239	239	237	238	239	235
Missing	Mean	937.73	937.15	937.57	933.52	937.87	935.75	937.41	936.16
	Std. Dev.	10.82	11.12	13.4	11.98	11.39	10.92	10.72	10.75
	N	182	180	179	178	179	178	180	177

Table 4.7.2 F

Mean Scale Scores by Ethnicity: Grade 6

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	937.03	937.29	936.17	931.12	937.09	934.53	937.34	935.16
	Std. Dev.	10.18	11.42	12.77	9.69	10.77	9.76	10.6	9.66
	N	1440	1425	1410	1394	1406	1386	1421	1374
Non-Hispanic American Indian	Mean	939.41	939.41	937.91	933.77	939.09	936.91	939.45	937.18
	Std. Dev.	10.14	9.99	14.02	8.8	11.73	8.45	9.83	9.1
	N	22	22	22	22	22	22	22	22
Non-Hispanic Asian	Mean	934.07	934.3	932.62	930.04	933.6	932.51	934.23	932.66
	Std. Dev.	11.73	13.18	14.56	11.15	12.51	11.46	12.39	11.44
	N	292	293	289	285	287	282	291	279
Non-Hispanic Black	Mean	934.62	935.24	934.2	930.37	934.95	933.15	935.12	933.56
	Std. Dev.	11.29	12.76	13.76	10.84	11.84	11.23	11.99	11.03
	N	152	152	149	149	149	148	151	146
Non-Hispanic Multiracial	Mean	937.25	936.83	937.33	931	937.67	934.25	937	935
	Std. Dev.	8	11.6	10.48	12.39	8.85	11.76	9.99	10.4
	N	12	12	12	12	12	12	12	12
Non-Hispanic Pacific Islander	Mean	933.11	931.74	935.53	928.44	934.28	930.89	932.94	932.59
	Std. Dev.	11.18	13.04	13.58	9.06	11.95	8.24	11.7	8.22
	N	18	19	19	18	18	18	18	17
Non-Hispanic White	Mean	934.35	934.19	933.12	928.6	934.17	931.65	934.43	932.28
	Std. Dev.	11.25	12.74	13.91	10.82	11.69	10.87	11.67	10.74
	N	195	194	190	191	188	189	193	186
Missing	Mean	936.91	937.44	936.56	931.84	937.16	934.87	937.25	935.32
	Std. Dev.	10.5	11.25	13.01	9.45	11.25	9.8	10.65	9.86
	N	155	156	153	153	153	153	155	151

Table 4.7.2 G

Mean Scale Scores by Ethnicity: Grade 7

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	937.72	938.16	936.57	932.09	937.61	935.45	938.1	935.89
	Std. Dev.	10.06	11.42	13.04	10.04	10.91	10.04	10.63	9.92
	N	1385	1380	1366	1361	1362	1352	1374	1341
Non-Hispanic American Indian	Mean	940.8	938.6	938.4	931.1	940	935	939.3	936.2
	Std. Dev.	11.08	11.47	15.2	10.37	12.46	10.64	11.16	10.82
	N	10	10	10	10	10	10	10	10
Non-Hispanic Asian	Mean	934.67	935.48	933.84	931.11	934.67	933.65	935.44	933.88
	Std. Dev.	11.64	13.63	14.28	11.3	12.39	11.63	12.67	11.53
	N	268	266	262	261	262	259	264	255
Non-Hispanic Black	Mean	936.84	937.34	934.99	932.22	936.5	935.26	937.24	935.51
	Std. Dev.	10.03	11.75	13.46	11.61	11.07	10.83	10.91	10.54
	N	144	143	142	138	141	137	143	137
Non-Hispanic Multiracial	Mean	937.67	939.25	935.08	929.33	936.75	934.5	938.92	934.92
	Std. Dev.	10.25	10.13	13.41	13.83	10.96	11.39	9.9	11.01
	N	12	12	12	12	12	12	12	12
Non-Hispanic Pacific Islander	Mean	935.57	937.22	937.22	932.91	936.74	935.26	936.7	935.48
	Std. Dev.	12.07	13.46	13.77	11.6	12.76	12.14	12.79	12.06
	N	23	23	23	23	23	23	23	23
Non-Hispanic White	Mean	935.9	936.62	935.54	931.56	936.38	934.47	936.62	935.02
	Std. Dev.	11.24	13.05	13.53	11.3	11.81	11.32	12.09	10.95
	N	183	181	179	177	177	175	180	172
Missing	Mean	937.44	937.11	936.9	933.46	937.55	935.68	937.2	935.97
	Std. Dev.	11.12	12.93	13.78	11.45	12.02	11.55	12.11	11.31
	N	135	135	135	132	135	131	135	131

Table 4.7.2 H

Mean Scale Scores by Ethnicity: Grade 8

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	938.03	938.42	936.12	932.44	937.57	935.8	938.39	936.18
	Std. Dev.	10.23	11.77	13.33	10.18	11.04	10.16	10.94	10.01
	N	1358	1355	1343	1331	1336	1326	1347	1313
Non-Hispanic American Indian	Mean	942.05	942.65	942.5	937.2	942.75	940.15	942.5	940.65
	Std. Dev.	8.65	9.82	10.75	8.51	8.26	7.84	9.36	7.55
	N	20	20	20	20	20	20	20	20
Non-Hispanic Asian	Mean	936.6	937.99	935.75	932.59	936.73	935.6	937.64	935.77
	Std. Dev.	10.5	11.61	13	11.17	10.72	10.74	10.86	10.26
	N	247	246	241	239	241	238	246	235
Non-Hispanic Black	Mean	935.88	935.99	935.17	930.81	935.91	933.79	936.03	934.23
	Std. Dev.	10.8	12.29	13.21	11.82	11.55	11.37	11.53	11.17
	N	139	138	138	139	138	137	138	136
Non-Hispanic Multiracial	Mean	939.92	942.18	940.33	934.75	940.5	938.91	941.55	939.18
	Std. Dev.	10.06	11.81	10.74	10.31	9.93	10.7	11.17	10.25
	N	12	11	12	12	12	11	11	11
Non-Hispanic Pacific Islander	Mean	938	936	935.75	931.12	937.19	933.82	937.69	935.44
	Std. Dev.	10.46	13.27	14.21	10.83	12	11.41	11.58	10.94
	N	16	17	16	17	16	17	16	16
Non-Hispanic White	Mean	934.89	935.68	934.67	930.71	935.16	933.49	935.56	933.95
	Std. Dev.	11.51	13.01	13.59	11.24	12.06	11.48	12.16	11.17
	N	156	157	156	154	155	154	156	152
Missing	Mean	940.02	940.37	939.12	934.96	939.96	937.83	940.36	938.27
	Std. Dev.	8.69	11.2	11.92	11.16	9.94	10.61	10.1	10.05
	N	130	131	129	128	127	126	130	125

Table 4.7.2 I

Mean Scale Scores by Ethnicity: Grade 9

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	937.16	936.99	934.95	932.34	936.31	934.97	937.16	935.31
	Std. Dev.	10.54	11.02	12.87	10.67	10.84	10.11	10.59	9.98
	N	1301	1293	1284	1272	1277	1263	1289	1249
Non-Hispanic American Indian	Mean	944.37	941.84	941.68	937.42	943.16	940	942.68	940.68
	Std. Dev.	4.18	7.81	5.75	6.65	4.71	6.64	6.4	5.93
	N	19	19	19	19	19	19	19	19
Non-Hispanic Asian	Mean	935.65	936.12	933.92	931.54	935.1	934.1	936.08	934.39
	Std. Dev.	10.72	11.51	13.45	10.8	10.88	10.35	10.95	10.08
	N	198	197	192	191	192	191	197	188
Non-Hispanic Black	Mean	935.84	936.27	934.78	931.54	935.73	934.2	936.31	934.65
	Std. Dev.	10.94	10.6	13.22	10.25	11.02	9.83	10.39	9.92
	N	122	120	117	115	117	115	120	113
Non-Hispanic Multiracial	Mean	940.62	940.62	939.15	935.77	940.08	938.31	940.69	938.69
	Std. Dev.	5.65	6.98	6.53	9.46	5.85	7.96	6.68	7.1
	N	13	13	13	13	13	13	13	13
Non-Hispanic Pacific Islander	Mean	937.09	936.55	940.45	934.45	938.91	935.64	936.82	936.55
	Std. Dev.	7.3	9.64	8.55	5.22	7.25	6.73	8.94	6.88
	N	11	11	11	11	11	11	11	11
Non-Hispanic White	Mean	936.2	935.62	934.85	931.01	935.74	933.56	935.85	934.03
	Std. Dev.	11.15	12.01	12.62	12.36	11.32	11.69	11.45	11.36
	N	158	158	156	156	156	156	158	155
Missing	Mean	937.66	937.09	936.77	933.52	937.5	935.35	937.43	935.77
	Std. Dev.	10.4	10.48	11.68	10.17	10.36	9.71	10.18	9.69
	N	138	136	137	131	136	130	136	130

Table 4.7.2 J

Mean Scale Scores by Ethnicity: Grade 10

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	937.97	937.68	935.65	932.75	937.04	935.52	937.89	935.85
	Std. Dev.	10.4	10.91	12.54	10.31	10.66	9.93	10.47	9.85
	N	1157	1155	1137	1136	1135	1131	1150	1119
Non-Hispanic American Indian	Mean	938.3	938	941.22	932.8	940.56	935.6	938.2	938.78
	Std. Dev.	9.67	12.1	8.23	11.19	8.47	10.18	11.17	7.41
	N	10	10	9	10	9	10	10	9
Non-Hispanic Asian	Mean	936.66	936.92	935.36	932.87	936.36	935.21	936.99	935.66
	Std. Dev.	10.24	11.4	12.6	10.84	10.47	10.54	10.62	10.07
	N	215	213	210	208	210	208	213	205
Non-Hispanic Black	Mean	937.27	937.21	936.21	932.21	936.85	934.91	937.34	935.33
	Std. Dev.	10.97	11.93	12.68	11.26	11.39	11.16	11.47	11.08
	N	125	124	124	122	124	121	124	121
Non-Hispanic Multiracial	Mean	927.57	928.14	927.71	924.14	927.86	926.43	928	926.71
	Std. Dev.	11.75	12.69	11.64	11.55	10.24	11.33	11.96	10.78
	N	7	7	7	7	7	7	7	7
Non-Hispanic Pacific Islander	Mean	939.88	943.65	942	937.53	941.12	940.71	942.59	940.82
	Std. Dev.	4.26	4.95	2.81	7.15	3.04	5.02	4.46	4.39
	N	17	17	17	17	17	17	17	17
Non-Hispanic White	Mean	937.62	937.06	935.7	931.38	936.91	934.56	937.31	935.16
	Std. Dev.	9.08	9.64	12.57	10.28	9.73	9.47	9.26	9.26
	N	113	114	110	111	109	111	113	109
Missing	Mean	939.23	938.9	937.98	935.07	938.75	937.23	939.26	937.74
	Std. Dev.	8.11	9.1	10.68	9.42	8.84	8.61	8.17	8.09
	N	104	105	104	104	104	104	104	103

Table 4.7.2 K

Mean Scale Scores by Ethnicity: Grade 11

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	938.55	938.45	936.38	933.57	937.75	936.27	938.65	936.58
	Std. Dev.	9.76	10.37	12.03	10.56	10	9.78	9.91	9.54
	N	986	980	972	968	969	963	976	955
Non-Hispanic American Indian	Mean	933.78	932.22	929.33	925.43	931.67	927.57	932.67	927.57
	Std. Dev.	14.02	14.01	15.91	14.74	13.91	14.5	13.83	13.91
	N	9	9	9	7	9	7	9	7
Non-Hispanic Asian	Mean	936.46	937.02	935.31	933.24	936.08	935.43	936.98	935.46
	Std. Dev.	10.01	10.29	12.19	11.2	10.25	9.91	9.85	9.71
	N	192	191	192	190	192	189	191	189
Non-Hispanic Black	Mean	936.5	936.43	936.25	932.87	936.59	934.81	936.45	935.22
	Std. Dev.	10.93	11.48	12.27	11.41	11.09	11.02	11.03	10.64
	N	111	112	110	111	110	111	111	109
Non-Hispanic Multiracial	Mean	944	942.8	943.8	943.4	944	943.4	943.4	943.4
	Std. Dev.	4.58	4.82	2.68	6.43	3.39	3.36	3.91	3.05
	N	5	5	5	5	5	5	5	5
Non-Hispanic Pacific Islander	Mean	931.63	932.68	931.5	931.06	931.78	931.56	932.42	931.53
	Std. Dev.	13.9	15.27	15.83	13.5	14.27	13.94	14.5	13.99
	N	19	19	18	18	18	18	19	17
Non-Hispanic White	Mean	936.3	936.93	935.8	932.59	936.31	935.18	936.89	935.61
	Std. Dev.	11.44	12.33	13.14	12.01	11.79	11.31	11.88	10.98
	N	119	117	116	110	116	110	117	109
Missing	Mean	938.85	938.2	938.33	934.27	938.71	936.68	938.53	937.1
	Std. Dev.	10.77	10.44	10.21	9.61	9.85	9.34	10.15	9.11
	N	93	92	93	91	92	90	91	89

Table 4.7.2 L

Mean Scale Scores by Ethnicity: Grade 12

Ethnicity		Listening	Reading	Speaking	Writing	Oral	Literacy	Comprehension	Overall
Hispanic (of any Race)	Mean	937.93	937.77	935.3	933.18	936.89	935.77	937.95	936.04
	Std. Dev.	10.23	10.96	12.73	10.9	10.69	10.22	10.48	10.06
	N	1600	1594	1577	1553	1574	1550	1590	1539
Non-Hispanic American Indian	Mean	939	939.31	939.67	934.8	939.67	937.4	939.5	937.64
	Std. Dev.	9.75	9.98	9.06	10.42	9.13	9.66	9.61	9.49
	N	17	16	15	15	15	15	16	14
Non-Hispanic Asian	Mean	937.07	937.62	935.33	933.18	936.47	935.68	937.55	935.8
	Std. Dev.	10.31	11.01	12.23	10.98	10.49	10.33	10.64	10.11
	N	385	384	376	375	375	373	384	369
Non-Hispanic Black	Mean	937.87	937.28	936.47	933.79	937.34	935.87	937.47	936.08
	Std. Dev.	9.92	10.86	11.74	10.6	10.04	9.87	10.39	9.51
	N	191	190	191	187	191	186	188	184
Non-Hispanic Multiracial	Mean	936.6	933	932.11	935.33	934.67	934.78	934.1	934.78
	Std. Dev.	7.86	13.52	14.08	13.7	10.36	13.51	11.45	12.66
	N	10	10	9	9	9	9	10	9
Non-Hispanic Pacific Islander	Mean	937.65	937.55	939.53	933.8	939.47	935.9	937.75	937.68
	Std. Dev.	9.85	10.51	8.24	9.23	7.37	9.55	10.25	7.31
	N	20	20	19	20	19	20	20	19
Non-Hispanic White	Mean	937.47	937.55	936.29	932.96	937.21	935.73	937.67	936.12
	Std. Dev.	10.24	10.44	11.62	10.8	10	9.63	10.1	9.39
	N	212	211	208	203	208	202	209	200
Missing	Mean	939.54	939.67	938.73	934.68	939.39	937.5	939.78	937.9
	Std. Dev.	9.21	9.5	10.06	9.44	8.93	8.74	9.19	8.43
	N	173	173	170	170	168	168	171	165

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	0.841	0.752	0.673
	N	6268	6202	6135	6095
Reading	Pearson Correlation		1	0.742	0.715
	N		6218	6122	6096
Speaking	Pearson Correlation			1	0.705
	N			6153	6060
Writing	Pearson Correlation				1
	N				6145

**. 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	0.868	0.748	0.688
	N	8455	8393	8340	8247
Reading	Pearson Correlation		1	0.762	0.762
	N		8420	8333	8257
Speaking	Pearson Correlation			1	0.72
	N			8369	8230
Writing	Pearson Correlation				1
	N				8301

**. 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	0.858	0.758	0.7
	N	6524	6468	6402	6330
Reading	Pearson Correlation		1	0.775	0.744
	N		6498	6401	6338
Speaking	Pearson Correlation			1	0.721
	N			6428	6311
Writing	Pearson Correlation				1
	N				6378

**. 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	0.874	0.746	0.719
	N	7850	7788	7706	7611
Reading	Pearson Correlation		1	0.77	0.762
	N		7815	7695	7623
Speaking	Pearson Correlation			1	0.721
	N			7727	7576
Writing	Pearson Correlation				1
	N				7655

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

4.9 Proficiency Levels

4.9.1 Proficiency Level by Grade-level Cluster

Table 4.9.1 A

Proficiency Level by Cluster: Listening

	Listening Proficiency Range										
	A1		A2		A3		P1		P2		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1617	25.8	898	14.33	1201	19.16	1152	18.38%	1400	22.34	6268
3-5	1371	16.22	888	10.5%	1262	14.93	1962	23.21%	2972	35.15	8455
6-8	841	12.89	656	10.06	886	13.58	931	14.27%	3210	49.2	6524
9-12	898	11.44	674	8.59%	1145	14.59	1662	21.17%	3471	44.22	7850
Total	4727	16.25	3116	10.71	4494	15.44	5707	19.61%	11053	37.99	29097

Table 4.9.1 B

Proficiency Level by Cluster: Reading

	Reading Proficiency Range										
	A1		A2		A3		P1		P2		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1694	27.24	989	15.91	1084	17.43	1157	18.61%	1294	20.81	6218
3-5	1398	16.6	1196	14.2%	1250	14.85	1887	22.41%	2689	31.94	8420
6-8	949	14.6	514	7.91%	745	11.47	1154	17.76%	3136	48.26	6498
9-12	930	11.9	799	10.22	891	11.4%	1468	18.78%	3727	47.69	7815
Total	4971	17.17	3498	12.08	3970	13.71	5666	19.57%	10846	37.46	28951

Table 4.9.1 C

Proficiency Level by Cluster: Speaking

	Speaking Proficiency Range										
	A1		A2		A3		P1		P2		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	2060	33.48%	280	4.55%	862	14.01%	1828	29.71%	1123	18.25%	6153
3-5	2002	23.92%	444	5.31%	730	8.72%	2520	30.11%	2673	31.94%	8369
6-8	1353	21.05%	249	3.87%	764	11.89%	1750	27.22%	2312	35.97%	6428
9-12	1502	19.44%	288	3.73%	845	10.94%	2014	26.06%	3078	39.83%	7727
Total	6917	24.12%	1261	4.4%	3201	11.16%	8112	28.29%	9186	32.03%	28677

Table 4.9.1 D

Proficiency Level by Cluster: Writing

	Writing Proficiency Range												
	A1		A2		A3		P1		P2		P3		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	2185	35.56%	1293	21.04%	1539	25.04%	955	15.54%	121	1.97%	52	0.85%	6145
3-5	1949	23.48%	1850	22.29%	2069	24.92%	1559	18.78%	705	8.49%	169	2.04%	8301
6-8	1222	19.16%	1802	28.25%	1188	18.63%	1884	29.54%	105	1.65%	177	2.78%	6378
9-12	1276	16.67%	1951	25.49%	1349	17.62%	2655	34.68%	137	1.79%	287	3.75%	7655
Total	6632	23.29%	6896	24.21%	6145	21.58%	7053	24.77%	1068	3.75%	685	2.41%	28479

Table 4.9.1 E

Proficiency Level by Cluster: Oral

	Oral Proficiency Range										
	A1		A2		A3		P1		P2		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1898	30.94%	544	8.87%	943	15.37%	1542	25.13%	1208	19.69%	6135
3-5	1713	20.54%	592	7.1%	1163	13.94%	2252	27.0%	2620	31.41%	8340
6-8	1120	17.49%	457	7.14%	857	13.39%	1318	20.59%	2650	41.39%	6402
9-12	1227	15.92%	494	6.41%	1037	13.46%	2088	27.1%	2860	37.11%	7706
Total	5958	20.84%	2087	7.3%	4000	13.99%	7200	25.19%	9338	32.67%	28583

Table 4.9.1 F

Proficiency Level by Cluster: Literacy

	Literacy Proficiency Range										
	A1		A2		A3		P1		P2		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1860	30.51%	1173	19.24%	1466	24.05%	998	16.37%	599	9.83%	6096
3-5	1609	19.49%	1520	18.41%	1862	22.55%	1905	23.07%	1361	16.48%	8257
6-8	974	15.37%	864	13.63%	1453	22.93%	1987	31.35%	1060	16.72%	6338
9-12	1026	13.46%	1025	13.45%	1683	22.08%	2382	31.25%	1507	19.77%	7623
Total	5469	19.32%	4582	16.18%	6464	22.83%	7272	25.68%	4527	15.99%	28314

Table 4.9.1 G

Proficiency Level by Cluster: Comprehension

	Comprehension Proficiency Range										
	A1		A2		A3		P1		P2		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1631	26.3%	969	15.62%	1049	16.91%	1375	22.17%	1178	18.99%	6202
3-5	1341	15.98%	1088	12.96%	1262	15.04%	1735	20.67%	2967	35.35%	8393
6-8	905	13.99%	530	8.19%	728	11.26%	1260	19.48%	3045	47.08%	6468
9-12	883	11.34%	739	9.49%	928	11.92%	1603	20.58%	3635	46.67%	7788
Total	4760	16.5%	3326	11.53%	3967	13.75%	5973	20.7%	10825	37.52%	28851

Table 4.9.1 H

Proficiency Level by Cluster: Overall

	Overall Proficiency Range										
	A1		A2		A3		P1		P2		
Cluster	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Total
1-2	1771	29.36%	922	15.28%	1495	24.78%	1189	19.71%	656	10.87%	6033
3-5	1515	18.49%	1093	13.34%	1900	23.19%	2120	25.88%	1564	19.09%	8192
6-8	939	14.96%	629	10.02%	1357	21.62%	1971	31.41%	1380	21.99%	6276
9-12	998	13.22%	755	10.0%	1659	21.98%	2307	30.57%	1828	24.22%	7547
Total	5223	18.62%	3399	12.12%	6411	22.86%	7587	27.05%	5428	19.35%	28048

4.9.2 Proficiency Level by Grade

Table 4.9.2 A

Proficiency Level by Grade: Listening

Grade	Listening Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	953	28.76%	532	16.05%	631	19.04%	596	17.98%	602	18.17%	3314
2	664	22.48%	366	12.39%	570	19.3%	556	18.82%	798	27.01%	2954
3	582	18.98%	368	12.0%	493	16.07%	713	23.25%	911	29.7%	3067
4	444	15.62%	301	10.59%	437	15.38%	666	23.43%	994	34.98%	2842
5	345	13.55%	219	8.6%	332	13.04%	583	22.9%	1067	41.91%	2546
6	334	14.61%	252	11.02%	326	14.26%	325	14.22%	1049	45.89%	2286
7	267	12.36%	214	9.91%	288	13.33%	327	15.14%	1064	49.26%	2160
8	240	11.55%	190	9.14%	272	13.09%	279	13.43%	1097	52.79%	2078
9	249	12.7%	189	9.64%	289	14.74%	418	21.33%	815	41.58%	1960
10	201	11.5%	146	8.35%	236	13.5%	374	21.4%	791	45.25%	1748
11	169	11.02%	119	7.76%	242	15.78%	307	20.01%	697	45.44%	1534
12	279	10.7%	220	8.44%	378	14.49%	563	21.59%	1168	44.79%	2608
Total	4727	16.25%	3116	10.71%	4494	15.44%	5707	19.61%	11053	37.99%	29097

Table 4.9.2 B

Proficiency Level by Grade: Reading

Grade	Reading Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	1008	30.62%	556	16.89%	565	17.16%	577	17.53%	586	17.8%	3292
2	686	23.44%	433	14.8%	519	17.74%	580	19.82%	708	24.2%	2926
3	602	19.74%	466	15.28%	475	15.57%	690	22.62%	817	26.79%	3050
4	455	16.06%	397	14.01%	437	15.42%	615	21.7%	930	32.82%	2834
5	341	13.45%	333	13.13%	338	13.33%	582	22.95%	942	37.15%	2536
6	361	15.88%	203	8.93%	289	12.71%	429	18.87%	991	43.6%	2273
7	304	14.14%	172	8.0%	234	10.88%	379	17.63%	1061	49.35%	2150
8	284	13.69%	139	6.7%	222	10.7%	346	16.67%	1084	52.24%	2075
9	249	12.79%	227	11.66%	242	12.43%	358	18.39%	871	44.74%	1947
10	203	11.63%	176	10.09%	195	11.17%	333	19.08%	838	48.02%	1745
11	173	11.34%	136	8.92%	180	11.8%	298	19.54%	738	48.39%	1525
12	305	11.74%	260	10.01%	274	10.55%	479	18.44%	1280	49.27%	2598
Total	4971	17.17%	3498	12.08%	3970	13.71%	5666	19.57%	10846	37.46%	28951

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	1207	37.06%	150	4.61%	438	13.45%	957	29.38%	505	15.51%	3257
2	853	29.45%	130	4.49%	424	14.64%	871	30.08%	618	21.34%	2896
3	812	26.77%	148	4.88%	273	9.0%	943	31.09%	857	28.26%	3033
4	668	23.73%	164	5.83%	254	9.02%	846	30.05%	883	31.37%	2815
5	522	20.71%	132	5.24%	203	8.05%	731	29.0%	933	37.01%	2521
6	491	21.88%	82	3.65%	297	13.24%	637	28.39%	737	32.84%	2244
7	434	20.39%	87	4.09%	242	11.37%	576	27.05%	790	37.11%	2129
8	428	20.83%	80	3.89%	225	10.95%	537	26.13%	785	38.2%	2055
9	413	21.41%	73	3.78%	197	10.21%	515	26.7%	731	37.9%	1929
10	337	19.62%	66	3.84%	173	10.07%	430	25.03%	712	41.44%	1718
11	268	17.69%	55	3.63%	176	11.62%	374	24.69%	642	42.38%	1515
12	484	18.87%	94	3.66%	299	11.66%	695	27.1%	993	38.71%	2565
Total	6917	24.12%	1261	4.4%	3201	11.16%	8112	28.29%	9186	32.03%	28677

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	1264	38.89%	718	22.09%	787	24.22%	417	12.83%	44	1.35%	20	0.62%	3250
2	921	31.81%	575	19.86%	752	25.98%	538	18.58%	77	2.66%	32	1.11%	2895
3	794	26.37%	736	24.44%	726	24.11%	523	17.37%	188	6.24%	44	1.46%	3011
4	646	23.12%	605	21.65%	704	25.2%	538	19.26%	241	8.63%	60	2.15%	2794
5	509	20.39%	509	20.39%	639	25.6%	498	19.95%	276	11.06%	65	2.6%	2496
6	462	20.77%	667	29.99%	455	20.46%	573	25.76%	30	1.35%	37	1.66%	2224
7	387	18.31%	597	28.24%	371	17.55%	652	30.84%	39	1.84%	68	3.22%	2114
8	373	18.28%	538	26.37%	362	17.75%	659	32.3%	36	1.76%	72	3.53%	2040
9	346	18.13%	500	26.21%	368	19.29%	599	31.39%	31	1.62%	64	3.35%	1908
10	278	16.21%	460	26.82%	294	17.14%	604	35.22%	33	1.92%	46	2.68%	1715
11	246	16.4%	351	23.4%	267	17.8%	541	36.07%	28	1.87%	67	4.47%	1500
12	406	16.03%	640	25.28%	420	16.59%	911	35.98%	45	1.78%	110	4.34%	2532
Total	6632	23.29%	6896	24.21%	6145	21.58%	7053	24.77%	1068	3.75%	685	2.41%	28479

Table 4.9.2 E

Proficiency Level by Grade: Oral

Grade	Oral Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	1114	34.31%	310	9.55%	503	15.49%	782	24.08%	538	16.57%	3247
2	784	27.15%	234	8.1%	440	15.24%	760	26.32%	670	23.2%	2888
3	726	24.0%	202	6.68%	435	14.38%	856	28.3%	806	26.64%	3025
4	561	19.99%	213	7.59%	400	14.26%	770	27.44%	862	30.72%	2806
5	426	16.98%	177	7.05%	328	13.07%	626	24.95%	952	37.94%	2509
6	415	18.57%	174	7.79%	321	14.36%	479	21.43%	846	37.85%	2235
7	373	17.58%	131	6.17%	276	13.01%	447	21.07%	895	42.18%	2122
8	332	16.23%	152	7.43%	260	12.71%	392	19.17%	909	44.45%	2045
9	340	17.7%	135	7.03%	256	13.33%	533	27.75%	657	34.2%	1921
10	262	15.28%	123	7.17%	213	12.42%	452	26.36%	665	38.78%	1715
11	231	15.29%	86	5.69%	213	14.1%	391	25.88%	590	39.05%	1511
12	394	15.4%	150	5.86%	355	13.87%	712	27.82%	948	37.05%	2559
Total	5958	20.84%	2087	7.3%	4000	13.99%	7200	25.19%	9338	32.67%	28583

Table 4.9.2 F

Proficiency Level by Grade: Literacy

Grade	Literacy Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	1095	33.91%	669	20.72%	754	23.35%	471	14.59%	240	7.43%	3229
2	765	26.68%	504	17.58%	712	24.83%	527	18.38%	359	12.52%	2867
3	667	22.35%	616	20.64%	675	22.61%	662	22.18%	365	12.23%	2985
4	539	19.34%	522	18.73%	598	21.46%	655	23.5%	473	16.97%	2787
5	403	16.22%	382	15.37%	589	23.7%	588	23.66%	523	21.05%	2485
6	363	16.43%	355	16.06%	535	24.21%	673	30.45%	284	12.85%	2210
7	319	15.2%	258	12.29%	480	22.87%	676	32.21%	366	17.44%	2099
8	292	14.39%	251	12.37%	438	21.59%	638	31.44%	410	20.21%	2029
9	286	15.07%	270	14.23%	429	22.6%	605	31.88%	308	16.23%	1898
10	227	13.28%	220	12.87%	395	23.11%	535	31.3%	332	19.43%	1709
11	189	12.66%	193	12.93%	334	22.37%	459	30.74%	318	21.3%	1493
12	324	12.84%	342	13.56%	525	20.81%	783	31.03%	549	21.76%	2523
Total	5469	19.32%	4582	16.18%	6464	22.83%	7272	25.68%	4527	15.99%	28314

Table 4.9.2 G

Proficiency Level by Grade: Comprehension

Grade	Comprehension Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	979	29.81%	549	16.72%	549	16.72%	678	20.65%	529	16.11%	3284
2	652	22.34%	420	14.39%	500	17.14%	697	23.89%	649	22.24%	2918
3	582	19.13%	431	14.16%	474	15.58%	661	21.72%	895	29.41%	3043
4	433	15.35%	370	13.12%	435	15.42%	552	19.57%	1031	36.55%	2821
5	326	12.89%	287	11.35%	353	13.96%	522	20.64%	1041	41.16%	2529
6	337	14.89%	219	9.68%	286	12.64%	463	20.46%	958	42.33%	2263
7	296	13.83%	169	7.89%	226	10.56%	408	19.06%	1042	48.67%	2141
8	272	13.18%	142	6.88%	216	10.47%	389	18.85%	1045	50.63%	2064
9	239	12.3%	215	11.07%	241	12.4%	397	20.43%	851	43.8%	1943
10	198	11.39%	156	8.98%	206	11.85%	361	20.77%	817	47.01%	1738
11	158	10.4%	134	8.82%	185	12.18%	311	20.47%	731	48.12%	1519
12	288	11.13%	234	9.04%	296	11.44%	534	20.63%	1236	47.76%	2588
Total	4760	16.5%	3326	11.53%	3967	13.75%	5973	20.7%	10825	37.52%	28851

Table 4.9.2 H

Proficiency Level by Grade: Overall

Grade	Overall Proficiency Range										Total
	A1		A2		A3		P1		P2		
	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	Count	% within PL	
1	1048	32.77%	529	16.54%	784	24.52%	578	18.07%	259	8.1%	3198
2	723	25.5%	393	13.86%	711	25.08%	611	21.55%	397	14.0%	2835
3	648	21.88%	420	14.18%	705	23.81%	764	25.8%	424	14.32%	2961
4	489	17.69%	387	14.0%	642	23.22%	701	25.35%	546	19.75%	2765
5	378	15.33%	286	11.6%	553	22.42%	655	26.56%	594	24.09%	2466
6	348	15.91%	256	11.71%	516	23.59%	682	31.18%	385	17.6%	2187
7	313	15.04%	182	8.75%	435	20.9%	675	32.44%	476	22.87%	2081
8	278	13.84%	191	9.51%	406	20.22%	614	30.58%	519	25.85%	2008
9	274	14.59%	211	11.24%	408	21.73%	589	31.36%	396	21.09%	1878
10	221	13.08%	169	10.0%	372	22.01%	526	31.12%	402	23.79%	1690
11	187	12.64%	132	8.92%	326	22.03%	449	30.34%	386	26.08%	1480
12	316	12.65%	243	9.72%	553	22.13%	743	29.73%	644	25.77%	2499
Total	5223	18.62%	3399	12.12%	6411	22.86%	7587	27.05%	5428	19.35%	28048

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	16	2	93	7	32	749	89	138	7	60	1936	3	21	7190	10343
	DB	2	0	1	0	0	3	1	1	1	0	3	0	0	8	20
	DD	48	3	4	2	25	58	36	34	13	25	438	1	13	1228	1928
	ED	1	0	0	0	0	8	1	3	0	3	6	0	0	49	71
	HI	6	1	3	0	1	22	7	5	2	7	19	0	0	48	121
	ID	347	8	29	23	126	22	120	415	112	60	1857	5	76	5994	9194
	MD	93	12	14	2	45	219	120	75	49	16	343	5	67	1457	2517
	OHI	27	3	20	2	23	100	19	4	15	16	232	0	21	824	1306
	OI	1	0	2	1	2	14	3	8	0	1	29	0	2	72	135
	SLD	2	0	4	2	2	10	4	5	1	0	79	0	0	295	404
	SLI	6	0	8	1	3	8	4	11	1	8	3	0	1	171	225
	TBI	3	0	1	0	0	11	5	1	2	4	14	1	10	120	172
	VI	3	0	1	0	1	13	4	4	1	1	2	0	0	22	52
	NPD	6	0	3	0	0	8	4	1	0	1	5	0	1	2784	2813
Total		561	29	183	40	260	1245	417	705	204	202	4966	15	212	20262	29301

Acronyms for Table 4.10.1

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

5. Analyses of Test Forms: Overview

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

5.1 Background

5.1.1 Measurement Models Used

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

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

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

where

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

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

B_n = ability of person "n"

D_i = difficulty of task "i"

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

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

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

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

According to Linacre (2002):

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

values between 1.5 and 2.0 are "unproductive for construction of measurement, but not degrading" values between 0.5 and 1.5 should be considered "productive for measurement"

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

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

5.1.2 Sampling

The results presented in most of the tables in Chapter 6 are based on the full data set of all students who were administered operational Series 601 of Alternate ACCESS for ELLs in the academic year 2022-2023. The item analysis summary tables (Table F), the complete item analysis tables (Table G), and the raw score to scale score conversion tables (Table H) use item difficulties from this calibration.

5.1.3 Scaling

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

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

Table 5.1.3A

Scaling Equation for each Domain

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

5.1.4 DIF Analyses

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

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

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

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

5.1.5 Reliability of Composites

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

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

Where

k = number of components j

w_j = domain weight of component j

σ_j^2 = variance of component j

σ_c^2 = variance of composite

ρ_j = reliability coefficient of component j.

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

5.1.6 Accuracy and Consistency of Classification

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

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

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

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

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

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

5.2 Descriptions

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

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

5.2.1 Raw Score Information (Figure A and Table A)

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

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

- The number of students in the analyses (the number of students who were not absent, invalid, refused, exempt, or in the wrong cluster)
- The minimum observed raw score
- The maximum observed raw score

- The mean (average) raw score
- The standard deviation (std. dev.) of the raw scores

5.2.2 Scale Score Information (Figure B and Table B)

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

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

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

5.2.3 Proficiency Level Information (Figure C and Table C)

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

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

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

5.2.4 Equating Summary Table (Table D)

[4] In Series 601, 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 601 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., 601).

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-level 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 error 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 the 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

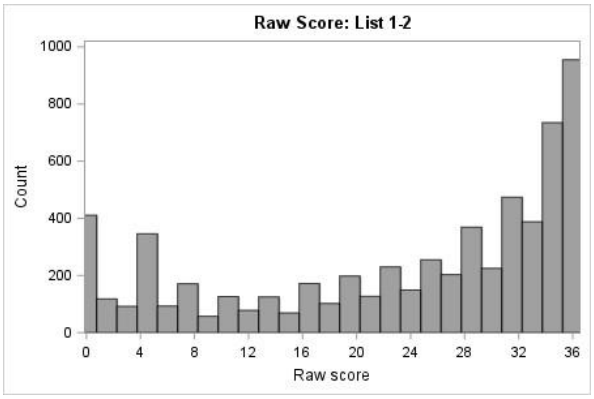


Table 6.1.1A

Raw Score Descriptive Statistics: List 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3314	0	36	22.1	12.33
2	2954	0	36	24.66	11.84
Total	6268	0	36	23.31	12.17

Figure 6.1.1B

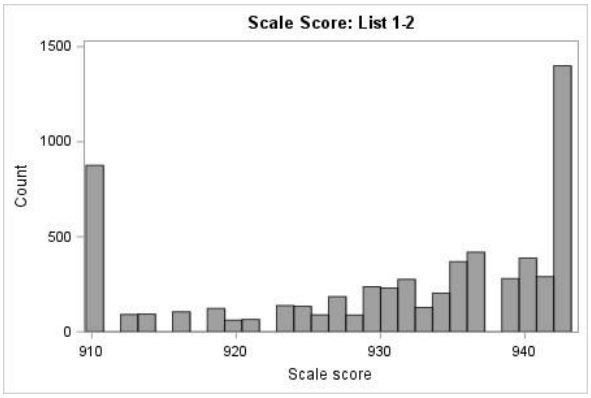


Table 6.1.1B

Scale Score Descriptive Statistics: List 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3314	910	943	929.85	11.55
2	2954	910	943	932.21	11.11
Total	6268	910	943	930.96	11.41

Figure 6.1.1C

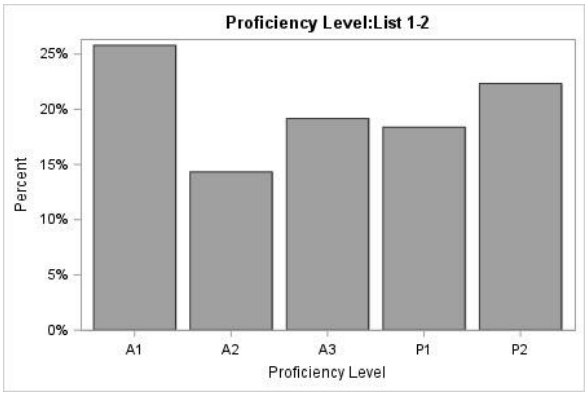


Table 6.1.1C

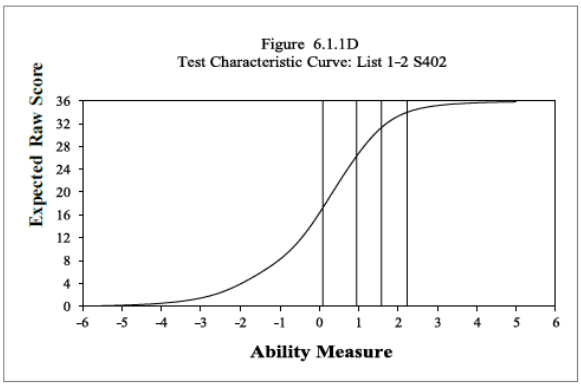
Proficiency Level Distribution: List 1-2

	Grade 1		Grade 2		Total	
Level	Count	Percent	Count	Percent	Count	Percent
A1	953	28.76%	664	22.48%	1617	25.8%
A2	532	16.05%	366	12.39%	898	14.33%
A3	631	19.04%	570	19.3%	1201	19.16%
P1	596	17.98%	556	18.82%	1152	18.38%
P2	602	18.17%	798	27.01%	1400	22.34%
Total	3314	100.0%	2954	100.0%	6268	100.0%

Table 6.1.1D

Equating Summary: List 1-2

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 102. Thus, the results from the S102 of the Alternate ACCESS were used to determine raw-to-scale score conversions.



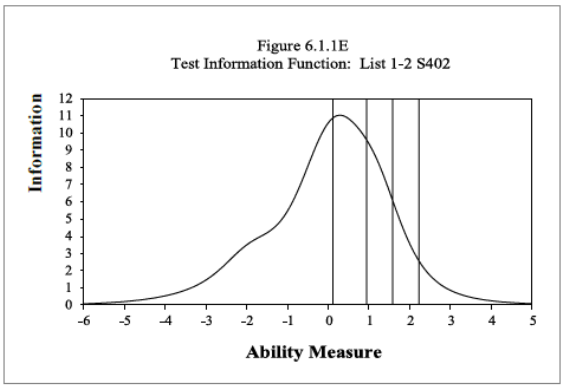


Table 6.1.1E

Reliability: List 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
6268	9	0.941	2.9557

Table 6.1.1F

Item Analysis Summary: List 1-2

Note: The contents of this table have been redacted in this version of the document.

Table 6.1.1G

Complete Item Analysis: List 1-2

Note: The contents of this table have been redacted in this version of the document.

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	7.66	7.66	A1	5.31	5.31
1	A1	0.45	8.12	A1	0.51	5.82
2	A1	1.57	9.69	A1	1.22	7.04
3	A1	1.63	11.32	A1	1.29	8.33
4	A1	4.44	15.75	A1	3.66	11.98
5	A1	1.6	17.35	A1	1.29	13.27
6	A1	1.78	19.13	A1	1.15	14.42
7	A1	1.87	21	A1	1.46	15.88
8	A1	1.33	22.33	A1	0.74	16.62
9	A1	0.94	23.26	A1	0.88	17.5
10	A1	1	24.26	A1	0.95	18.45
11	A1	1.03	25.29	A1	1.05	19.5
12	A1	1.48	26.77	A1	0.98	20.48
13	A1	0.91	27.67	A1	1.02	21.5
14	A1	1.09	28.76	A1	0.98	22.48
15	A2	1.27	30.02	A2	0.91	23.39
16	A2	1.81	31.83	A2	0.98	24.37
17	A2	1.63	33.46	A2	0.98	25.36
18	A2	1.84	35.3	A2	1.39	26.74
19	A2	1.57	36.87	A2	1.22	27.96
20	A2	1.81	38.68	A2	1.69	29.65
21	A2	2.14	40.83	A2	1.9	31.55
22	A2	1.87	42.7	A2	1.46	33.01
23	A2	2.11	44.81	A2	1.86	34.87
24	A3	2.41	47.22	A3	2.34	37.2
25	A3	1.81	49.03	A3	2.27	39.47
26	A3	1.63	50.66	A3	2.51	41.98
27	A3	3.2	53.86	A3	3.28	45.26
28	A3	3.32	57.18	A3	2.37	47.63
29	A3	3.08	60.26	A3	2.95	50.58
30	A3	3.59	63.85	A3	3.59	54.16
31	P1	2.96	66.81	P1	3.25	57.41
32	P1	4.38	71.18	P1	4.57	61.98
33	P1	6.1	77.28	P1	6.3	68.28
34	P1	4.56	81.83	P1	4.71	72.99
35	P2	6.16	87.99	P2	8.16	81.14
36	P2	12.01	100	P2	18.86	100

Table 6.1.1J

Accuracy and Consistency of Classification Indices: List 1-2

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.680	0.571	0.448
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.881	0.148
	A2	0.620	0.274
	A3	0.594	0.208
	P1	0.333	0.228
	P2	0.761	0.704

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.959	0.026	0.015	0.940
	A2/A3	0.931	0.033	0.036	0.909
	A3/P1	0.917	0.016	0.067	0.885
	P1/P2	0.852	0.057	0.091	0.778

6.1.2 Reading 1-2

Figure 6.1.2A

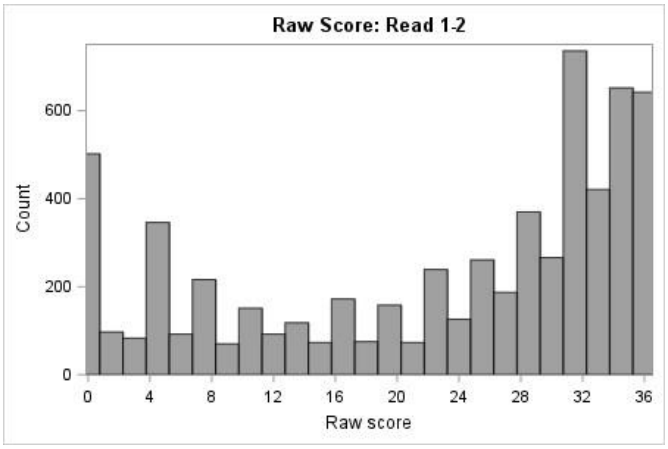


Table 6.1.2A

Raw Score Descriptive Statistics: Read 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3292	0	36	21.37	12.49
2	2926	0	36	23.92	11.92
Total	6218	0	36	22.57	12.29

Figure 6.1.2B

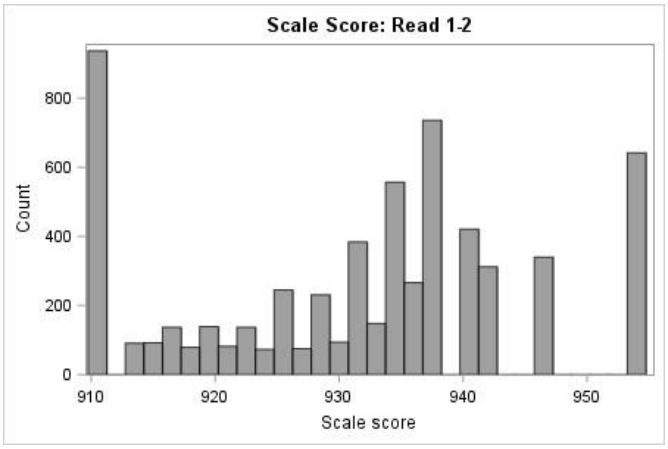


Table 6.1.2B

Scale Score Descriptive Statistics: Read 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3292	910	954	930.29	13.27
2	2926	910	954	933.12	13.17
Total	6218	910	954	931.62	13.3

Figure 6.1.2C

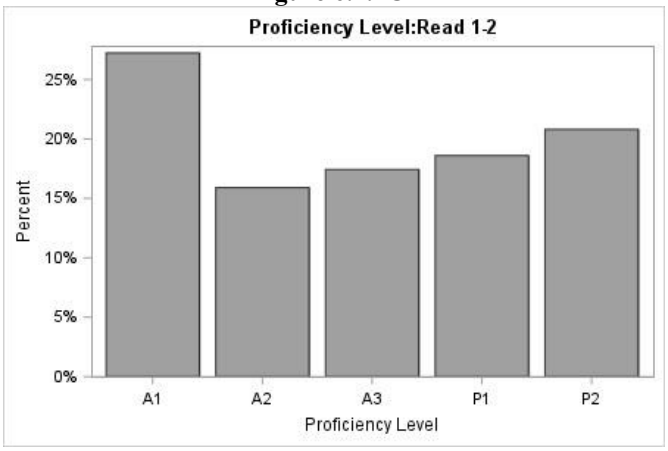


Table 6.1.2C

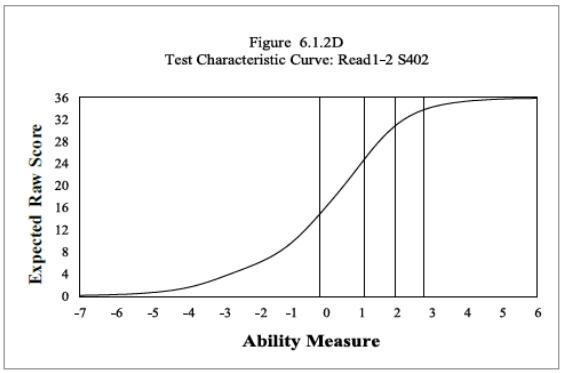
Proficiency Level Distribution: Read 1-2

	Grade 1		Grade 2		Total	
Level	Count	Percent	Count	Percent	Count	Percent
A1	1008	30.62%	686	23.44%	1694	27.24%
A2	556	16.89%	433	14.8%	989	15.91%
A3	565	17.16%	519	17.74%	1084	17.43%
P1	577	17.53%	580	19.82%	1157	18.61%
P2	586	17.8%	708	24.2%	1294	20.81%
Total	3292	100.0%	2926	100.0%	6218	100.0%

Table 6.1.2D

Equating Summary: Read 1-2

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).



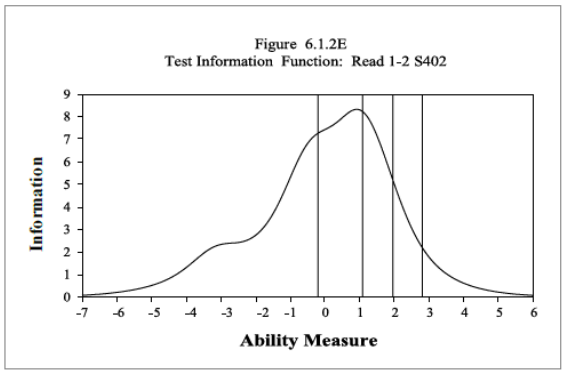


Table 6.1.2E

Reliability: Read 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
6218	9	0.9477	2.8121

Table 6.1.2F

Item Analysis Summary: Read 1-2

Note: The contents of this table have been redacted in this version of the document.

Table 6.1.2G

Complete Item Analysis: Read 1-2

Note: The contents of this table have been redacted in this version of the document.

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	9.3	9.3	A1	6.7	6.7
1	A1	0.67	9.96	A1	0.31	7.01
2	A1	1.09	11.06	A1	1.03	8.03
3	A1	1.52	12.58	A1	1.13	9.16
4	A1	4.56	17.13	A1	3.59	12.75
5	A1	1.58	18.71	A1	1.33	14.08
6	A1	1.88	20.6	A1	1.03	15.11
7	A1	2.46	23.06	A1	1.91	17.02
8	A1	1.43	24.48	A1	1.09	18.11
9	A1	1.28	25.76	A1	0.96	19.07
10	A1	1.12	26.88	A1	1.09	20.16
11	A1	1.43	28.31	A1	1.2	21.36
12	A1	1.67	29.98	A1	1.26	22.62
13	A1	0.64	30.62	A1	0.82	23.44
14	A2	1.22	31.83	A2	1.13	24.57
15	A2	1.31	33.14	A2	1.03	25.6
16	A2	1.46	34.6	A2	1.33	26.93
17	A2	1.46	36.06	A2	1.26	28.2
18	A2	1.31	37.36	A2	1.09	29.29
19	A2	1.46	38.82	A2	1.13	30.42
20	A2	1.58	40.4	A2	0.85	31.27
21	A2	1.28	41.68	A2	1.06	32.33
22	A2	1.52	43.2	A2	1.5	33.83
23	A2	2.37	45.57	A2	2.29	36.12
24	A2	1.94	47.51	A2	2.12	38.24
25	A3	1.85	49.36	A3	1.78	40.02
26	A3	2.61	51.97	A3	2.12	42.14
27	A3	2.76	54.74	A3	3.28	45.42
28	A3	2.34	57.08	A3	2.46	47.88
29	A3	3.49	60.57	A3	3.62	51.5
30	A3	4.1	64.67	A3	4.48	55.98
31	P1	5.07	69.74	P1	4.03	60.01
32	P1	6.23	75.97	P1	8.41	68.42
33	P1	6.23	82.2	P1	7.38	75.8
34	P2	4.34	86.54	P2	5.78	81.58
35	P2	4.92	91.46	P2	6.08	87.66
36	P2	8.54	100	P2	12.34	100

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

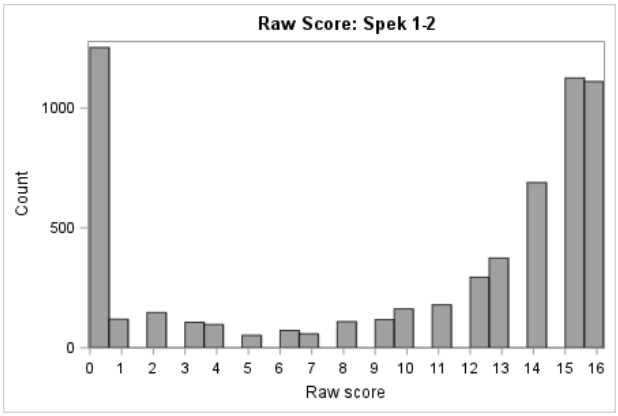


Table 6.1.3A

Raw Score Descriptive Statistics: Spek 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3257	0	16	9.42	6.43
2	2896	0	16	10.49	6.13
Total	6153	0	16	9.93	6.31

Figure 6.1.3B

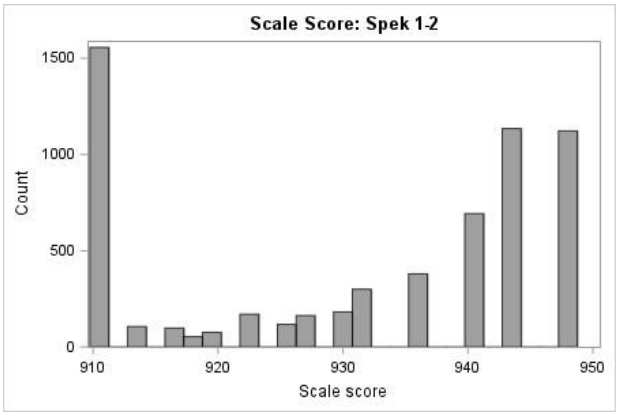


Table 6.1.3B

Scale Score Descriptive Statistics: Spek 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3257	910	948	930.07	15.03
2	2896	910	948	932.64	14.59
Total	6153	910	948	931.28	14.88

Figure 6.1.3C

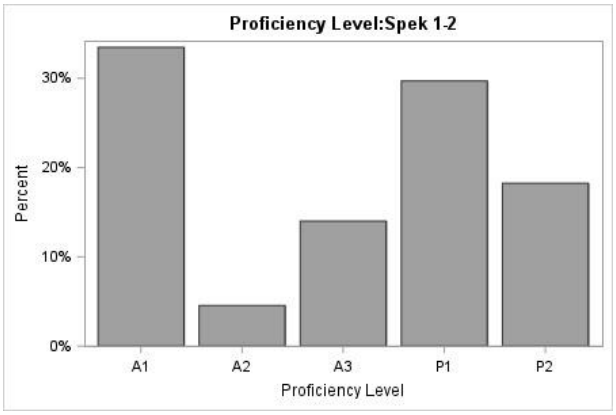


Table 6.1.3C

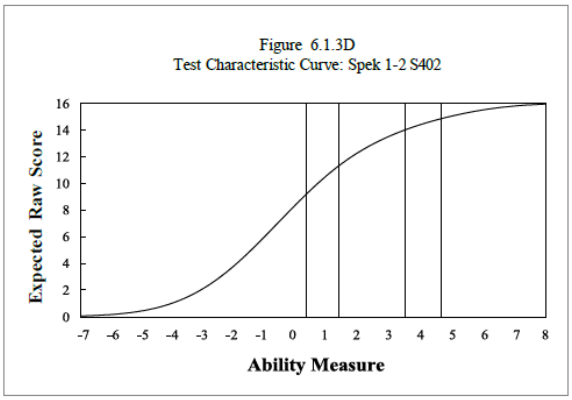
Proficiency Level Distribution: Spek 1-2

	Grade 1		Grade 2		Total	
Level	Count	Percent	Count	Percent	Count	Percent
A1	1207	37.06%	853	29.45%	2060	33.48%
A2	150	4.61%	130	4.49%	280	4.55%
A3	438	13.45%	424	14.64%	862	14.01%
P1	957	29.38%	871	30.08%	1828	29.71%
P2	505	15.51%	618	21.34%	1123	18.25%
Total	3257	100.0%	2896	100.0%	6153	100.0%

Table 6.1.3D

Equating Summary: Spek 1-2

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).



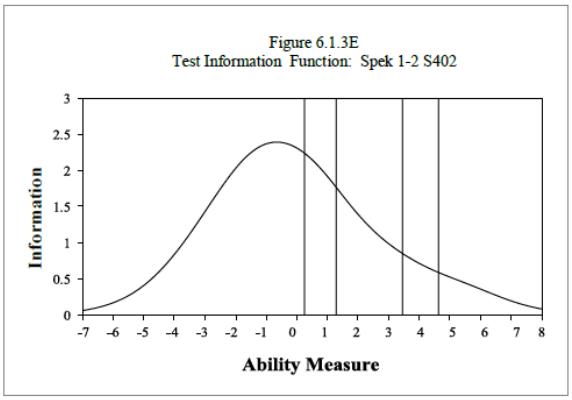


Table 6.1.3E

Reliability: Spek 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
6153	8	0.9632	1.21

Table 6.1.3F

Item Analysis Summary: Spek 1-2

Note: The contents of this table have been redacted in this version of the document.

Table 6.1.3G

Complete Item Analysis: Spek 1-2

Note: The contents of this table have been redacted in this version of the document.

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	23.52	23.52	A1	17.99	17.99
1	A1	1.96	25.48	A1	1.97	19.96
2	A1	2.7	28.19	A1	2.11	22.06
3	A1	1.84	30.03	A1	1.59	23.65
4	A1	1.63	31.65	A1	1.55	25.21
5	A1	0.92	32.58	A1	0.79	26
6	A1	1.54	34.11	A1	0.9	26.9
7	A1	1.11	35.22	A1	0.83	27.73
8	A1	1.84	37.06	A1	1.73	29.45
9	A2	1.96	39.02	A2	1.83	31.28
10	A2	2.64	41.66	A2	2.66	33.94
11	A3	2.61	44.27	A3	3.35	37.29
12	A3	4.82	49.09	A3	4.94	42.23
13	A3	6.02	55.11	A3	6.35	48.58
14	P1	11.27	66.38	P1	11.26	59.84
15	P1	18.11	84.49	P1	18.82	78.66
16	P2	15.51	100	P2	21.34	100

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

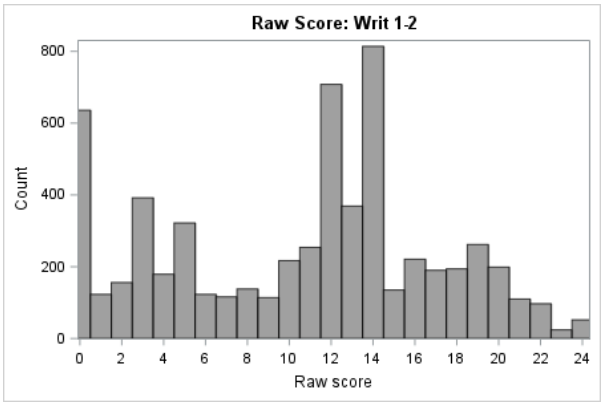


Table 6.1.4A

Raw Score Descriptive Statistics: Writ 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3250	0	24	9.92	6.29
2	2895	0	24	11.23	6.52
Total	6145	0	24	10.54	6.43

Figure 6.1.4B

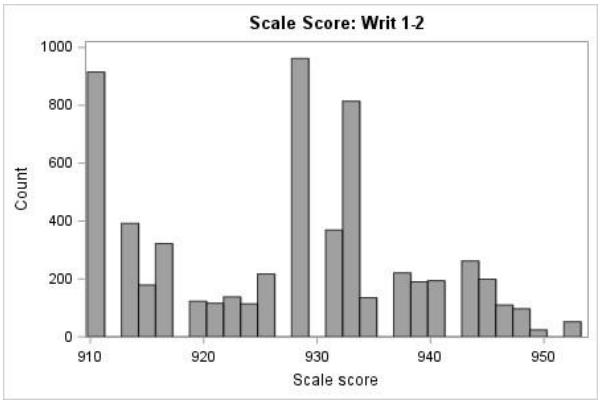


Table 6.1.4B

Scale Score Descriptive Statistics: Writ 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3250	910	953	926.01	11.19
2	2895	910	953	928.38	11.71
Total	6145	910	953	927.13	11.5

Figure 6.1.4C

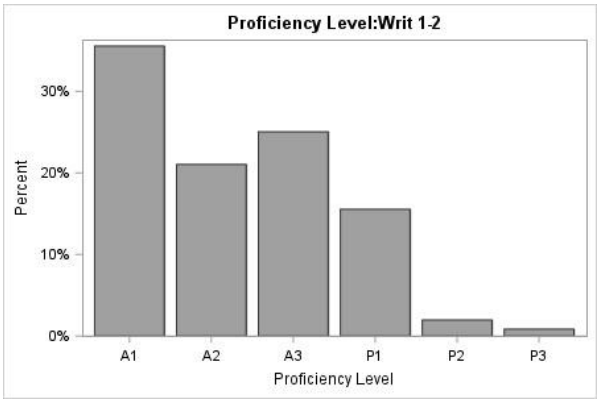


Table 6.1.4C

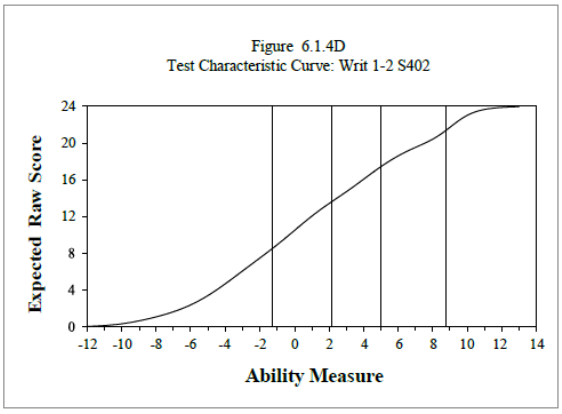
Proficiency Level Distribution: Writ 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	1264	38.89%	921	31.81%	2185	35.56%
A2	718	22.09%	575	19.86%	1293	21.04%
A3	787	24.22%	752	25.98%	1539	25.04%
P1	417	12.83%	538	18.58%	955	15.54%
P2	44	1.35%	77	2.66%	121	1.97%
P3	20	0.62%	32	1.11%	52	0.85%
Total	3250	100.0%	2895	100.0%	6145	100.0%

Table 6.1.4D

Equating Summary: Writ 1-2

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).



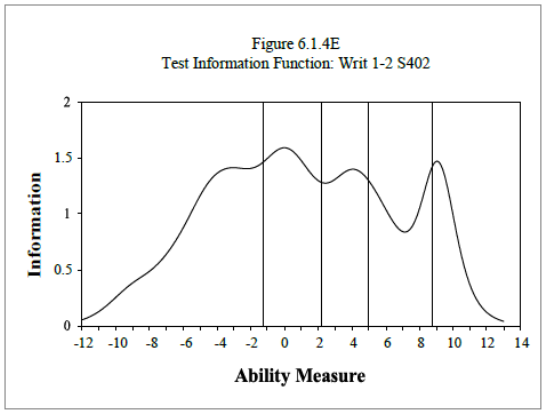


Table 6.1.4E

Reliability: Writ 1-2

No. of Students	No. of Items	Cronbach's Alpha	SEM
6145	10	0.932	1.6773

Table 6.1.4F

Item Analysis Summary: Writ 1-2

Note: The contents of this table have been redacted in this version of the document.

Table 6.1.4G

Complete Item Analysis: Writ 1-2

Note: The contents of this table have been redacted in this version of the document.

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	11.63	11.63	A1	8.91	8.91
1	A1	1.88	13.51	A1	2.14	11.05
2	A1	2.95	16.46	A1	2.07	13.13
3	A1	6.49	22.95	A1	6.25	19.38
4	A1	3.54	26.49	A1	2.21	21.59
5	A1	5.38	31.88	A1	5.08	26.67
6	A1	2.09	33.97	A1	1.9	28.57
7	A1	2.28	36.25	A1	1.45	30.02
8	A1	2.65	38.89	A1	1.8	31.81
9	A2	2	40.89	A2	1.69	33.51
10	A2	3.91	44.8	A2	3.11	36.61
11	A2	4.62	49.42	A2	3.59	40.21
12	A2	11.57	60.98	A2	11.47	51.68
13	A3	6.09	67.08	A3	5.91	57.58
14	A3	12.71	79.78	A3	13.85	71.43
15	A3	2.15	81.94	A3	2.25	73.68
16	A3	3.26	85.2	A3	3.97	77.65
17	P1	2.46	87.66	P1	3.8	81.45
18	P1	2.95	90.62	P1	3.39	84.84
19	P1	3.54	94.15	P1	5.08	89.91
20	P1	2.74	96.89	P1	3.8	93.71
21	P1	1.14	98.03	P1	2.52	96.23
22	P2	1.14	99.17	P2	2.07	98.31
23	P2	0.22	99.38	P2	0.59	98.89
24	P3	0.62	100	P3	1.11	100

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

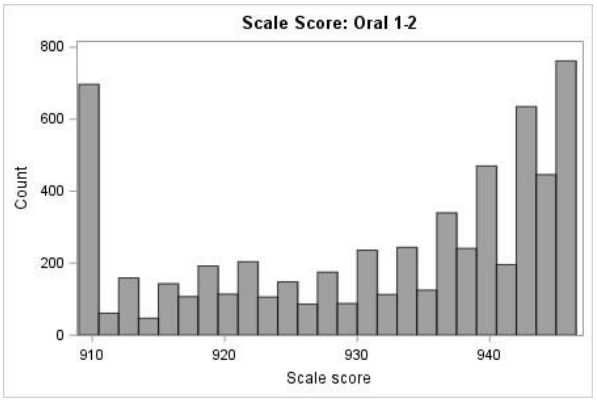


Table 6.1.5A

Scale Score Descriptive Statistics: Oral 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3247	910	946	930.29	12.52
2	2888	910	946	932.83	12.04
Total	6135	910	946	931.49	12.36

Figure 6.1.5B

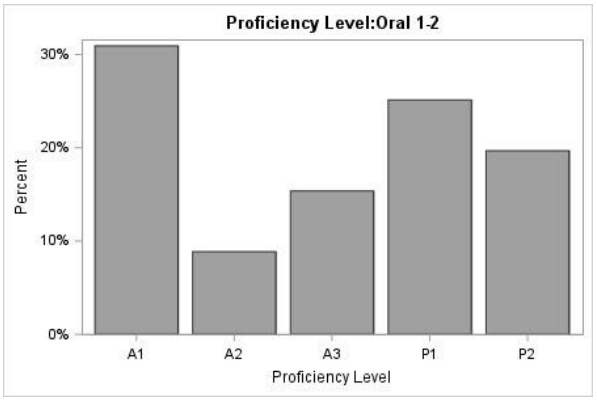


Table 6.1.5C

Proficiency Level Distribution: Oral 1-2

	Grade 1		Grade 2		Total	
Level	Count	Percent	Count	Percent	Count	Percent
A1	1114	34.31%	784	27.15%	1898	30.94%
A2	310	9.55%	234	8.1%	544	8.87%
A3	503	15.49%	440	15.24%	943	15.37%
P1	782	24.08%	760	26.32%	1542	25.13%
P2	538	16.57%	670	23.2%	1208	19.69%
Total	3247	100.0%	2888	100.0%	6135	100.0%

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	130.09	0.941
Speaking	0.5	221.3746	0.9632
Oral		152.7402	0.9741

*Variances from students who had results in all four domains

Table 6.1.5F
n/a

Table 6.1.5G
n/a

Table 6.1.5H
n/a

Table 6.1.5I
n/a

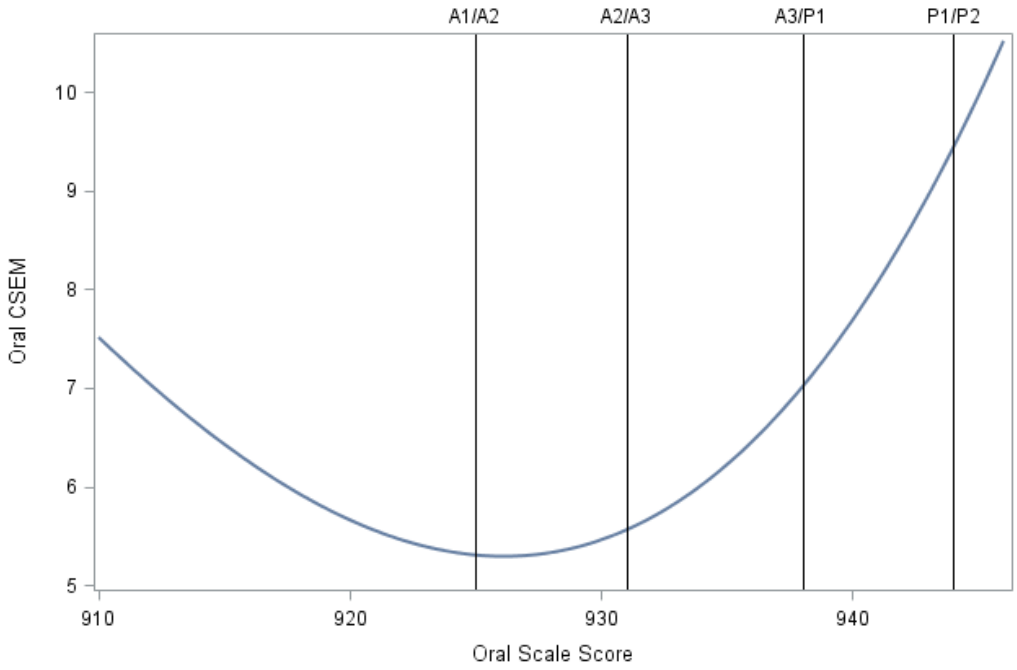
Table 6.1.5J

Accuracy and Consistency of Classification Indices: Oral 1-2

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.740	0.656	0.552
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.947	0.921
	A2	0.639	0.517
	A3	0.739	0.629
	P1	0.632	0.507
	P2	0.679	0.620

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.974	0.014	0.013	0.962
	A2/A3	0.967	0.017	0.016	0.953
	A3/P1	0.958	0.017	0.025	0.941
	P1/P2	0.841	0.067	0.092	0.794

Figure 6.1.5F CSEM for Oral Composite 1-2



6.1.6 Literacy Composite 1-2

Figure 6.1.6A

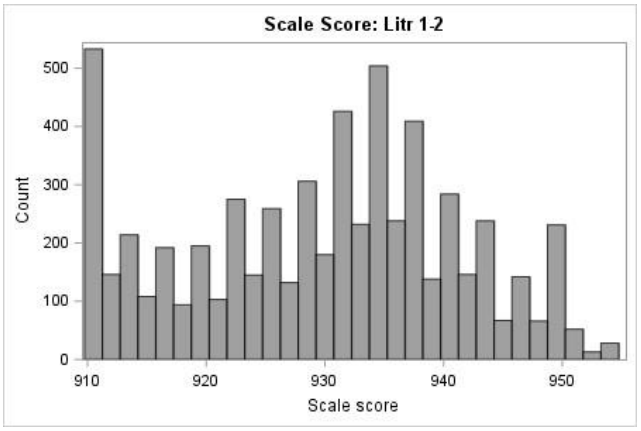


Table 6.1.6A

Scale Score Descriptive Statistics: Litr1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3229	910	954	928.49	11.34
2	2867	910	954	931.13	11.56
Total	6096	910	954	929.73	11.52

Figure 6.1.6B

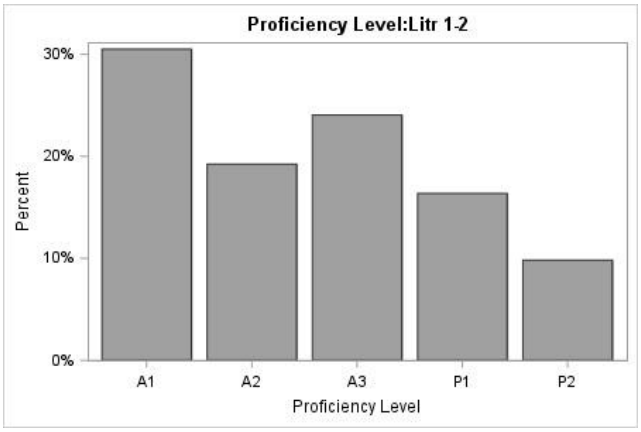


Table 6.1.6C

Proficiency Level Distribution: Litr 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	1095	33.91%	765	26.68%	1860	30.51%
A2	669	20.72%	504	17.58%	1173	19.24%
A3	754	23.35%	712	24.83%	1466	24.05%
P1	471	14.59%	527	18.38%	998	16.37%
P2	240	7.43%	359	12.52%	599	9.83%
Total	3229	100.0%	2867	100.0%	6096	100.0%

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	176.7667	0.9477
Writing	0.5	132.1917	0.932
Literacy		132.6442	0.9656

*Variances from students who had results in all four domains

Table 6.1.6F
n/a

Table 6.1.6G
n/a

Table 6.1.6H
n/a

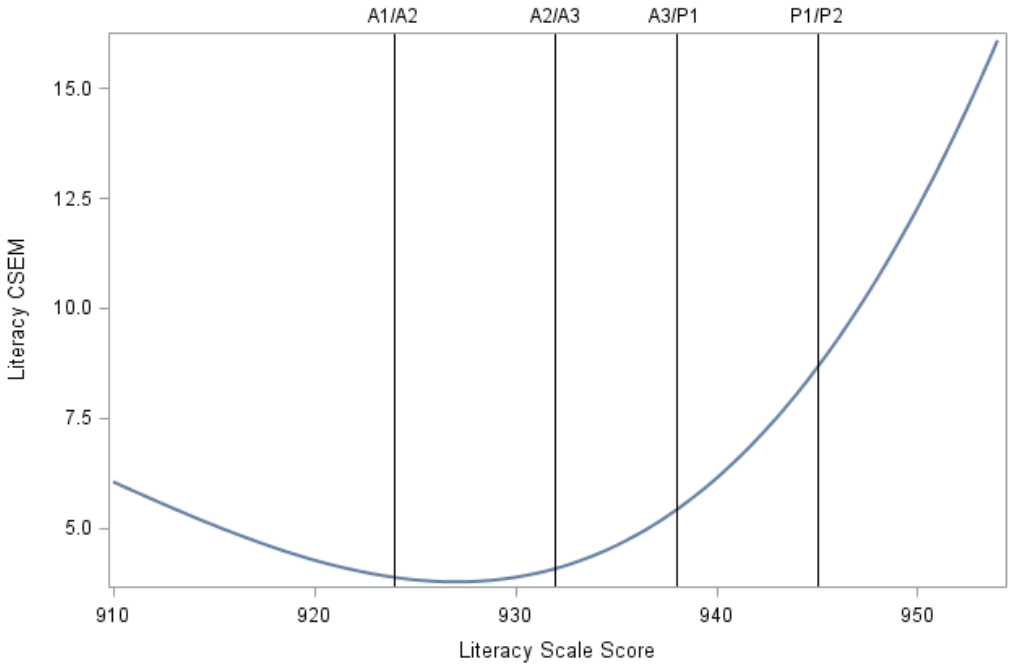
Table 6.1.6I
n/a

Table 6.1.6J

Accuracy and Consistency of Classification Indices: Litr 1-2

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

Figure 6.1.6F CSEM for Literacy Composite 1-2



6.1.7 Comprehension Composite 1-2

Figure 6.1.7A

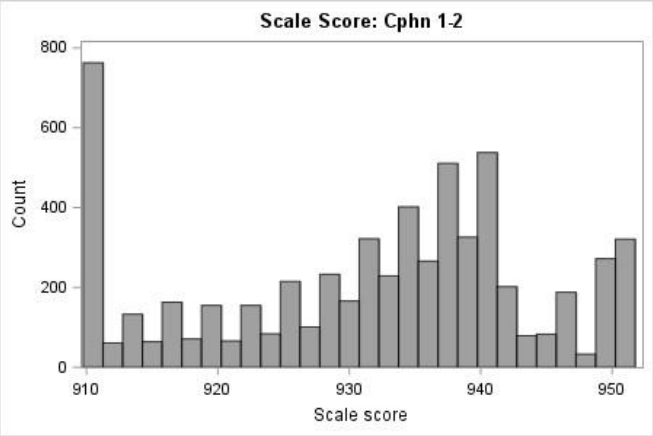


Table 6.1.7A

Scale Score Descriptive Statistics: Cphn 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3284	910	951	930.25	12.36
2	2918	910	951	932.97	12.15
Total	6202	910	951	931.53	12.34

Figure 6.1.7B

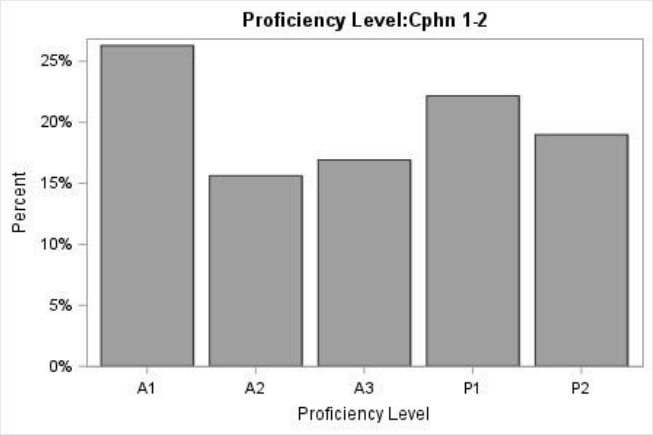


Table 6.1.7C

Proficiency Level Distribution: Cphn 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	979	29.81%	652	22.34%	1631	26.3%
A2	549	16.72%	420	14.39%	969	15.62%
A3	549	16.72%	500	17.14%	1049	16.91%
P1	678	20.65%	697	23.89%	1375	22.17%
P2	529	16.11%	649	22.24%	1178	18.99%
Total	3284	100.0%	2918	100.0%	6202	100.0%

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	130.09	0.941
Reading	0.7	176.7667	0.9477
Comprehension		152.1833	0.9657

*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

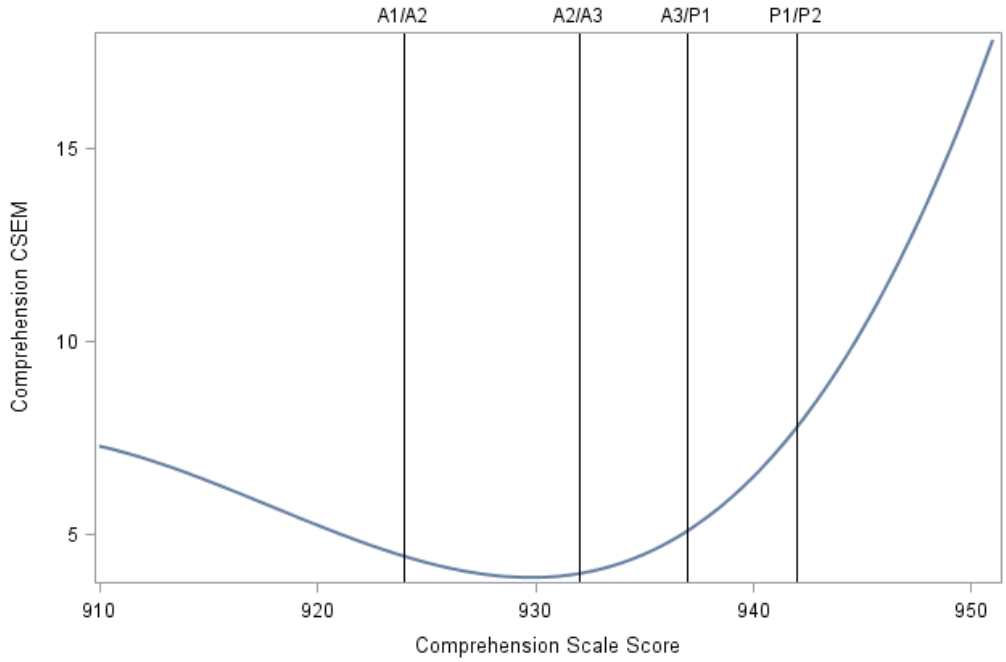
Table 6.1.7J

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

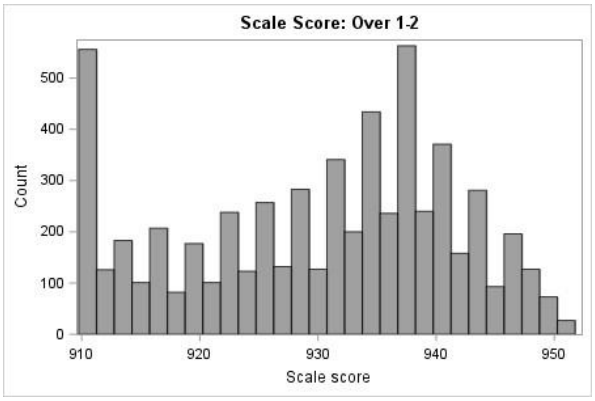


Table 6.1.8A

Scale Score Descriptive Statistics: Over 1-2

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	3198	910	951	928.86	11.25
2	2835	910	951	931.46	11.23
Total	6033	910	951	930.08	11.32

Figure 6.1.8B

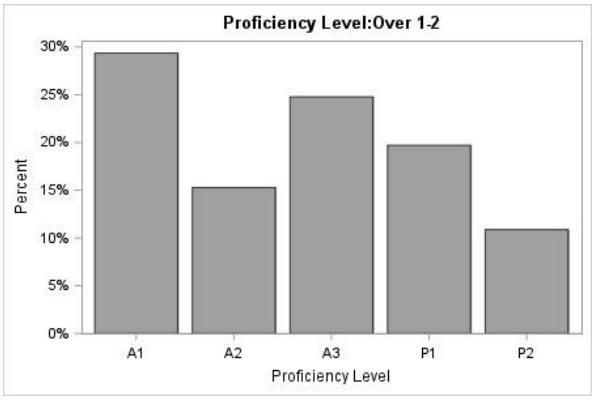


Table 6.1.8C

Proficiency Level Distribution: Over 1-2

Level	Grade 1		Grade 2		Total	
	Count	Percent	Count	Percent	Count	Percent
A1	1048	32.77%	723	25.5%	1771	29.36%
A2	529	16.54%	393	13.86%	922	15.28%
A3	784	24.52%	711	25.08%	1495	24.78%
P1	578	18.07%	611	21.55%	1189	19.71%
P2	259	8.1%	397	14.0%	656	10.87%
Total	3198	100.0%	2835	100.0%	6033	100.0%

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	130.09	0.941
Reading	0.35	176.7667	0.9477
Speaking	0.15	221.3746	0.9632
Writing	0.35	132.1917	0.932
Overall Composite		128.0387	0.9798

*Variances from students who had results in all four domains

Table 6.1.8F
n/a

Table 6.1.8G
n/a

Table 6.1.8H
n/a

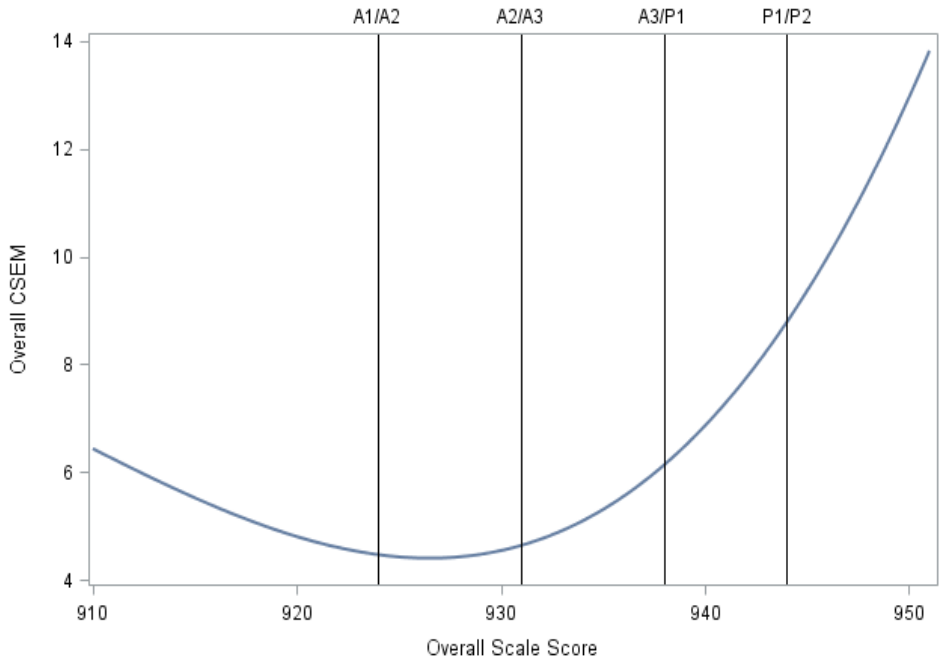
Table 6.1.8I
n/a

Table 6.1.8J

Accuracy and Consistency of Classification Indices: Over 1-2

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

Figure 6.1.8F CSEM for Overall Composite 1-2



6.2 Grades: 3-5

6.2.1 Listening 3-5

Figure 6.2.1A

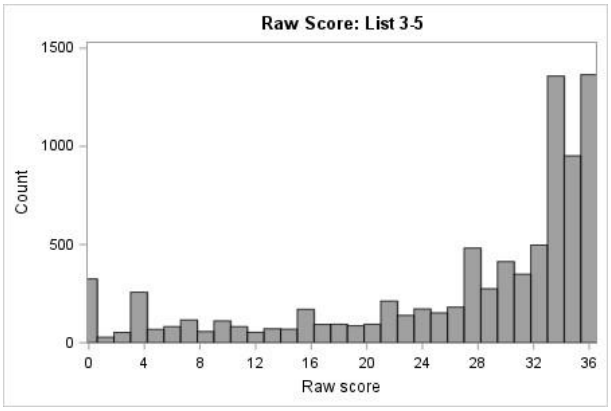


Table 6.2.1A

Raw Score Descriptive Statistics: List 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	3067	0	36	25.54	11.03
4	2842	0	36	26.82	10.58
5	2546	0	36	27.96	10.26
Total	8455	0	36	26.7	10.7

Figure 6.2.1B

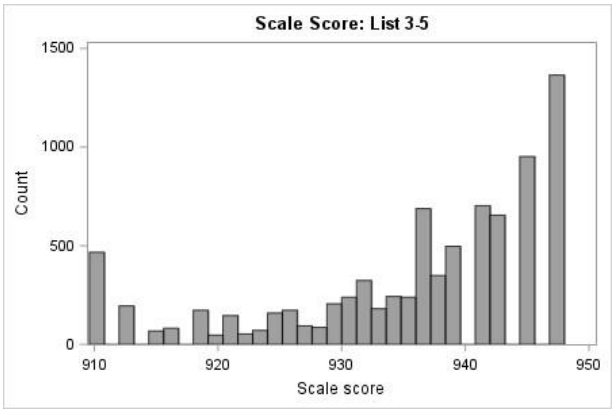


Table 6.2.1B

Scale Score Descriptive Statistics: List 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	3067	910	947	934.33	10.96
4	2842	910	947	935.71	10.62
5	2546	910	947	936.97	10.44
Total	8455	910	947	935.59	10.75

Figure 6.2.1C

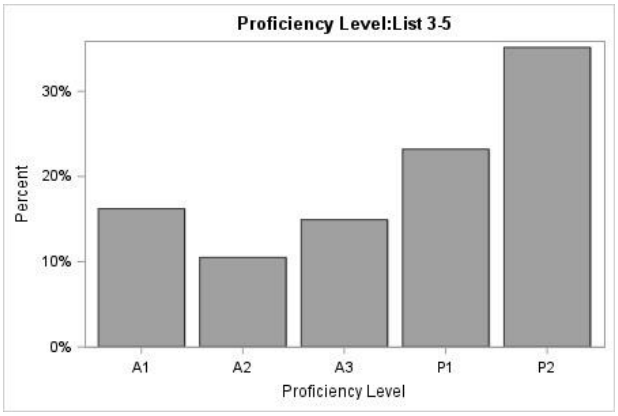


Table 6.2.1C

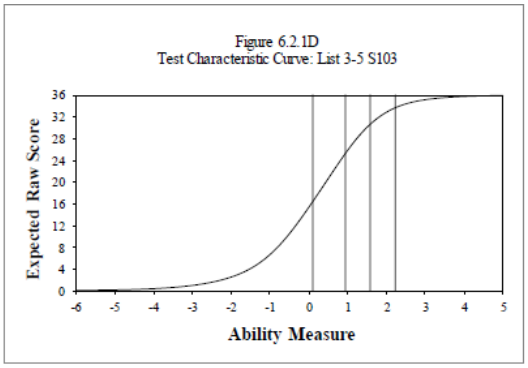
Proficiency Level Distribution: List 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	582	18.98%	444	15.62%	345	13.55%	1371	16.22%
A2	368	12.0%	301	10.59%	219	8.6%	888	10.5%
A3	493	16.07%	437	15.38%	332	13.04%	1262	14.93%
P1	713	23.25%	666	23.43%	583	22.9%	1962	23.21%
P2	911	29.7%	994	34.98%	1067	41.91%	2972	35.15%
Total	3067	100.0%	2842	100.0%	2546	100.0%	8455	100.0%

Table 6.2.1D

Equating Summary: List 3-5

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 102. Thus, the results from the S102 of the Alternate ACCESS were used to determine raw-to-scale score conversions.



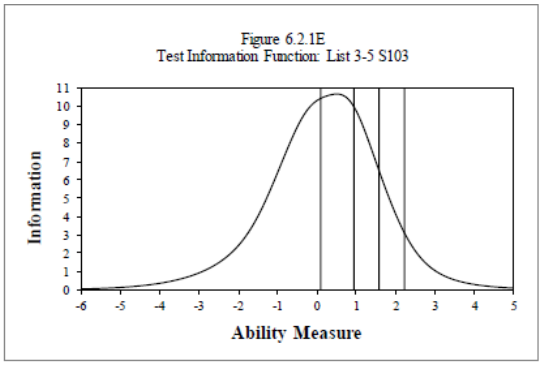


Table 6.2.1E

Reliability: List 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
8455	9	0.9369	2.6867

Table 6.2.1F

Item Analysis Summary: List 3-5

Note: The contents of this table have been redacted in this version of the document.

Table 6.2.1G

Complete Item Analysis: List 3-5

Note: The contents of this table have been redacted in this version of the document.

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	4.01	4.01	A1	3.94	3.94	A1	3.53	3.53
1	A1	0.42	4.43	A1	0.28	4.22	A1	0.27	3.81
2	A1	0.82	5.25	A1	0.42	4.64	A1	0.59	4.4
3	A1	0.91	6.16	A1	0.67	5.31	A1	0.59	4.99
4	A1	3.26	9.42	A1	2.11	7.42	A1	1.37	6.36
5	A1	0.78	10.21	A1	0.91	8.34	A1	0.71	7.07
6	A1	1.14	11.35	A1	0.77	9.11	A1	0.98	8.05
7	A1	1.37	12.72	A1	1.3	10.42	A1	1.45	9.51
8	A1	0.85	13.56	A1	0.81	11.22	A1	0.31	9.82
9	A1	0.65	14.22	A1	0.67	11.89	A1	0.31	10.13
10	A1	0.88	15.1	A1	0.63	12.53	A1	0.75	10.88
11	A1	1.04	16.14	A1	0.7	13.23	A1	1.18	12.06
12	A1	0.82	16.95	A1	0.56	13.79	A1	0.47	12.53
13	A1	0.98	17.93	A1	1.02	14.81	A1	0.47	13
14	A1	1.04	18.98	A1	0.81	15.62	A1	0.55	13.55
15	A2	1.14	20.12	A2	1.13	16.75	A2	0.9	14.45
16	A2	1.17	21.29	A2	1.09	17.84	A2	0.51	14.96
17	A2	1.17	22.46	A2	1.06	18.9	A2	1.06	16.03
18	A2	1.3	23.77	A2	1.09	19.99	A2	0.9	16.93
19	A2	1.34	25.11	A2	0.91	20.9	A2	0.75	17.67
20	A2	1.11	26.21	A2	1.3	22.2	A2	0.9	18.58
21	A2	1.57	27.78	A2	1.16	23.36	A2	1.22	19.8
22	A2	1.43	29.21	A2	1.2	24.56	A2	0.86	20.66
23	A2	1.76	30.97	A2	1.65	26.21	A2	1.49	22.15
24	A3	2.28	33.26	A3	2.08	28.29	A3	1.69	23.84
25	A3	1.83	35.08	A3	1.97	30.26	A3	1.57	25.41
26	A3	2.15	37.24	A3	2.22	32.48	A3	2.04	27.45
27	A3	3.29	40.53	A3	2.78	35.26	A3	2.47	29.93
28	A3	3.1	43.63	A3	3.03	38.28	A3	2.28	32.21
29	A3	3.42	47.05	A3	3.31	41.59	A3	2.99	35.19
30	P1	4.79	51.84	P1	4.89	46.48	P1	4.99	40.18
31	P1	4.14	55.98	P1	3.91	50.39	P1	4.36	44.54
32	P1	6.16	62.15	P1	5.74	56.12	P1	5.7	50.24
33	P1	8.15	70.3	P1	8.9	65.02	P1	7.86	58.09
34	P2	7.11	77.4	P2	7.78	72.8	P2	8.48	66.58
35	P2	10.5	87.9	P2	10.73	83.53	P2	12.77	79.34
36	P2	12.1	100	P2	16.47	100	P2	20.66	100

Table 6.2.1J

Accuracy and Consistency of Classification Indices: List 3-5

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.483	0.506	0.338
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.899	0.158
	A2	0.625	0.222
	A3	0.633	0.100
	P1	0.369	0.361
	P2	-	0.598

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.976	0.014	0.010	0.966
	A2/A3	0.961	0.019	0.020	0.947
	A3/P1	0.937	0.011	0.052	0.910
	P1/P2	0.605	0.395	0.000	0.652

6.2.2 Reading 3-5

Figure 6.2.2A

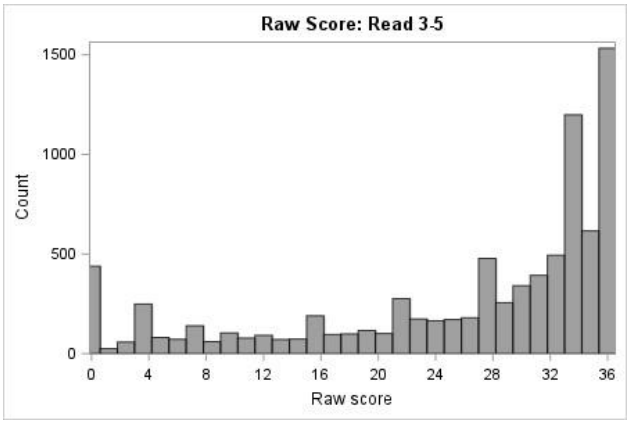


Table 6.2.2A

Raw Score Descriptive Statistics: Read 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	3050	0	36	24.54	11.54
4	2834	0	36	26.03	11.02
5	2536	0	36	27.09	10.69
Total	8420	0	36	25.81	11.16

Figure 6.2.2B

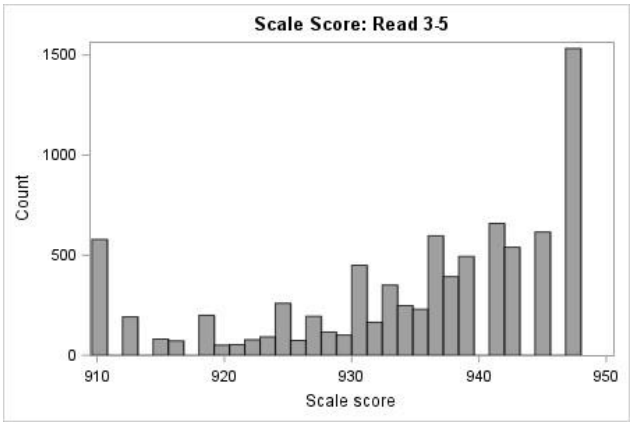


Table 6.2.2B

Scale Score Descriptive Statistics: Read 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	3050	910	947	933.46	11.35
4	2834	910	947	935.03	10.99
5	2536	910	947	936.17	10.78
Total	8420	910	947	934.81	11.12

Figure 6.2.2C

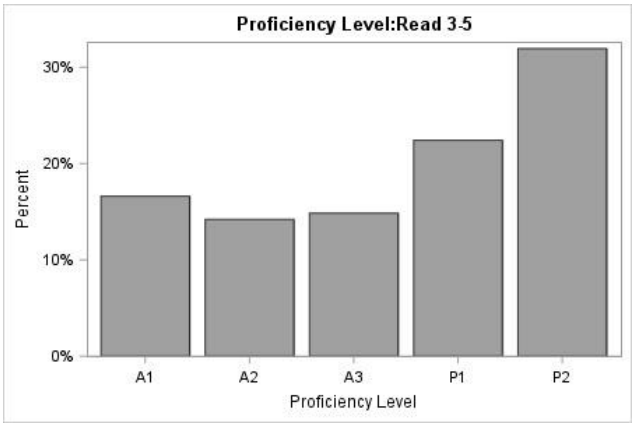


Table 6.2.2C

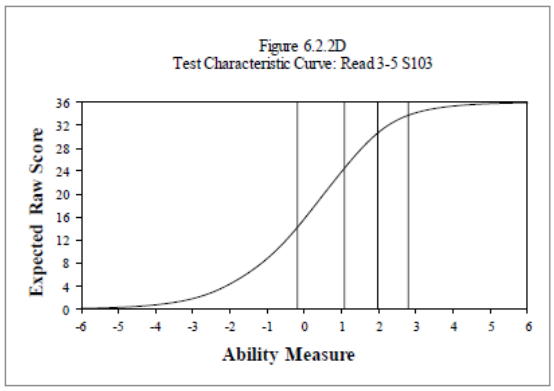
Proficiency Level Distribution: Read 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	602	19.74%	455	16.06%	341	13.45%	1398	16.6%
A2	466	15.28%	397	14.01%	333	13.13%	1196	14.2%
A3	475	15.57%	437	15.42%	338	13.33%	1250	14.85%
P1	690	22.62%	615	21.7%	582	22.95%	1887	22.41%
P2	817	26.79%	930	32.82%	942	37.15%	2689	31.94%
Total	3050	100.0%	2834	100.0%	2536	100.0%	8420	100.0%

Table 6.2.2D

Equating Summary: Read 3-5

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 102. Thus, the results from the S102 of the Alternate ACCESS were used to determine raw-to-scale score conversions.



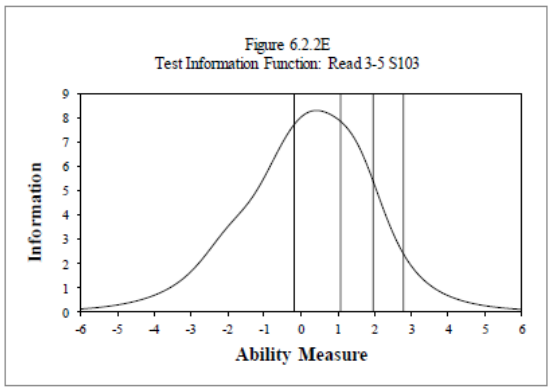


Table 6.2.2E

Reliability: Read 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
8420	9	0.9478	2.5511

Table 6.2.2F

Item Analysis Summary: Read 3-5

Note: The contents of this table have been redacted in this version of the document.

Table 6.2.2G

Complete Item Analysis: Read 3-5

Note: The contents of this table have been redacted in this version of the document.

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	5.93	5.93	A1	4.8	4.8	A1	4.81	4.81
1	A1	0.36	6.3	A1	0.28	5.08	A1	0.24	5.05
2	A1	1.02	7.31	A1	0.49	5.58	A1	0.51	5.56
3	A1	0.75	8.07	A1	0.71	6.28	A1	0.55	6.11
4	A1	3.02	11.08	A1	2.12	8.4	A1	1.58	7.69
5	A1	0.92	12	A1	1.06	9.46	A1	0.91	8.6
6	A1	0.92	12.92	A1	1.09	10.55	A1	0.51	9.11
7	A1	1.87	14.79	A1	1.76	12.31	A1	1.3	10.41
8	A1	0.82	15.61	A1	0.88	13.2	A1	0.39	10.8
9	A1	0.85	16.46	A1	0.46	13.66	A1	0.47	11.28
10	A1	0.75	17.21	A1	0.74	14.4	A1	0.35	11.63
11	A1	0.98	18.2	A1	0.95	15.35	A1	0.83	12.46
12	A1	1.54	19.74	A1	0.71	16.06	A1	0.99	13.45
13	A2	0.89	20.62	A2	0.81	16.87	A2	0.83	14.27
14	A2	0.98	21.61	A2	0.95	17.82	A2	0.63	14.91
15	A2	1.8	23.41	A2	0.92	18.74	A2	1.38	16.29
16	A2	0.98	24.39	A2	0.81	19.55	A2	0.83	17.11
17	A2	1.51	25.9	A2	1.13	20.68	A2	0.71	17.82
18	A2	1.21	27.11	A2	1.09	21.77	A2	1.22	19.05
19	A2	1.54	28.66	A2	1.34	23.11	A2	1.22	20.27
20	A2	1.28	29.93	A2	1.16	24.28	A2	1.14	21.41
21	A2	1.64	31.57	A2	1.8	26.08	A2	1.81	23.23
22	A2	1.41	32.98	A2	1.76	27.84	A2	1.42	24.65
23	A2	2.03	35.02	A2	2.22	30.06	A2	1.93	26.58
24	A3	2	37.02	A3	2.05	32.11	A3	1.81	28.39
25	A3	1.84	38.85	A3	2.26	34.37	A3	2.01	30.4
26	A3	2.36	41.21	A3	2.15	36.52	A3	1.85	32.26
27	A3	3.54	44.75	A3	3.03	39.56	A3	2.13	34.38
28	A3	2.89	47.64	A3	2.79	42.34	A3	2.48	36.87
29	A3	2.95	50.59	A3	3.14	45.48	A3	3.04	39.91
30	P1	4.23	54.82	P1	3.56	49.05	P1	4.38	44.28
31	P1	5.25	60.07	P1	4.45	53.49	P1	4.22	48.5
32	P1	6.03	66.1	P1	5.68	59.17	P1	5.88	54.38
33	P1	7.11	73.21	P1	8.01	67.18	P1	8.48	62.85
34	P2	5.54	78.75	P2	6.92	74.1	P2	6.9	69.76
35	P2	7.28	86.03	P2	7.55	81.65	P2	7.1	76.85
36	P2	13.97	100	P2	18.35	100	P2	23.15	100

Table 6.2.2J

Accuracy and Consistency of Classification Indices: Read 3-5

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.527	0.528	0.389
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.896	0.125
	A2	0.706	0.253
	A3	0.653	0.110
	P1	0.385	0.373
	P2	-	0.591

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.973	0.016	0.011	0.961
	A2/A3	0.954	0.022	0.024	0.937
	A3/P1	0.939	0.012	0.049	0.912
	P1/P2	0.659	0.341	0.000	0.693

6.2.3 Speaking 3-5

Figure 6.2.3A

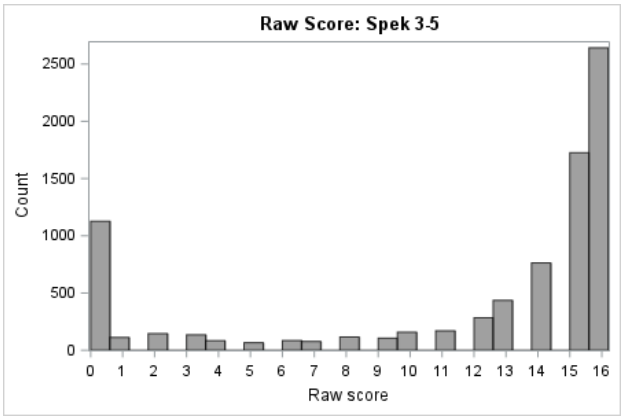


Table 6.2.3A

Raw Score Descriptive Statistics: Spek 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	3033	0	16	11.17	6.05
4	2815	0	16	11.61	5.77
5	2521	0	16	12.09	5.63
Total	8369	0	16	11.59	5.84

Figure 6.2.3B

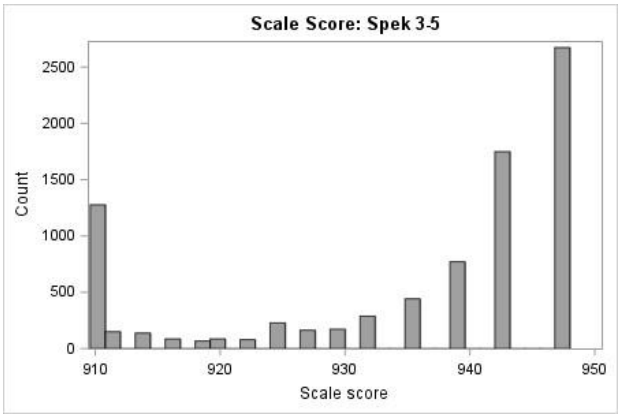


Table 6.2.3B

Scale Score Descriptive Statistics: Spek 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	3033	910	947	933.98	14.09
4	2815	910	947	934.97	13.63
5	2521	910	947	936.27	13.37
Total	8369	910	947	935	13.75

Figure 6.2.3C

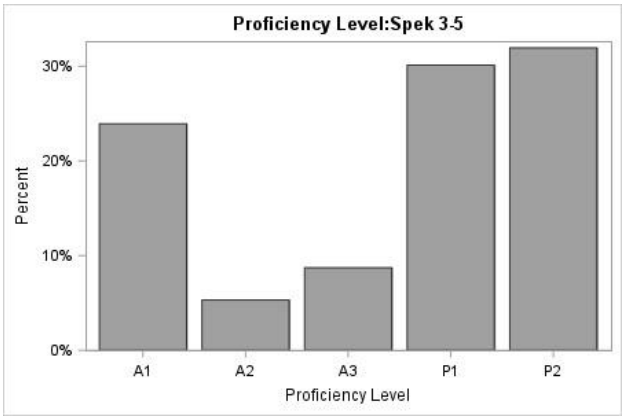


Table 6.2.3C

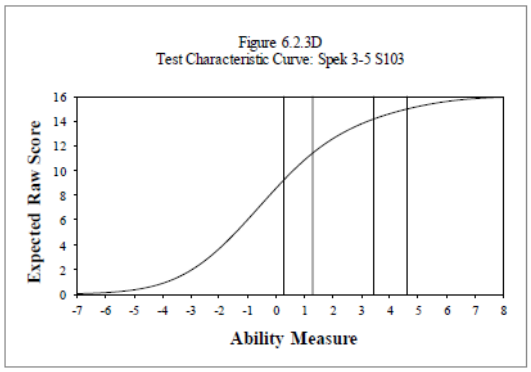
Proficiency Level Distribution: Spek 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	812	26.77%	668	23.73%	522	20.71%	2002	23.92%
A2	148	4.88%	164	5.83%	132	5.24%	444	5.31%
A3	273	9.0%	254	9.02%	203	8.05%	730	8.72%
P1	943	31.09%	846	30.05%	731	29.0%	2520	30.11%
P2	857	28.26%	883	31.37%	933	37.01%	2673	31.94%
Total	3033	100.0%	2815	100.0%	2521	100.0%	8369	100.0%

Table 6.2.3D

Equating Summary: Spek 3-5

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).



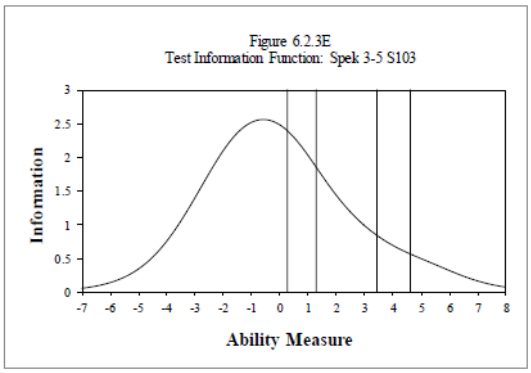


Table 6.2.3E

Reliability: Spek 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
8369	8	0.9664	1.0716

Table 6.2.3F

Item Analysis Summary: Spek 3-5

Note: The contents of this table have been redacted in this version of the document.

Table 6.2.3G

Complete Item Analysis: Spek 3-5

Note: The contents of this table have been redacted in this version of the document.

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	16.19	16.19	A1	13.07	13.07	A1	12.06	12.06
1	A1	1.38	17.57	A1	1.28	14.35	A1	1.43	13.49
2	A1	1.65	19.22	A1	2.1	16.45	A1	1.63	15.11
3	A1	1.68	20.9	A1	1.95	18.4	A1	1.27	16.38
4	A1	1.32	22.22	A1	0.89	19.29	A1	0.83	17.22
5	A1	0.92	23.15	A1	0.82	20.11	A1	0.63	17.85
6	A1	1.29	24.43	A1	1.03	21.14	A1	0.71	18.56
7	A1	0.82	25.26	A1	1.03	22.17	A1	0.99	19.56
8	A1	1.52	26.77	A1	1.56	23.73	A1	1.15	20.71
9	A2	1.06	27.83	A2	1.53	25.26	A2	1.35	22.05
10	A2	1.98	29.81	A2	1.81	27.07	A2	2.06	24.12
11	A2	1.85	31.65	A2	2.49	29.56	A2	1.82	25.94
12	A3	3.63	35.28	A3	3.52	33.07	A3	3.13	29.08
13	A3	5.37	40.65	A3	5.51	38.58	A3	4.92	33.99
14	P1	9.86	50.51	P1	9.38	47.96	P1	8.25	42.25
15	P1	21.23	71.74	P1	20.67	68.63	P1	20.75	62.99
16	P2	28.26	100	P2	31.37	100	P2	37.01	100

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

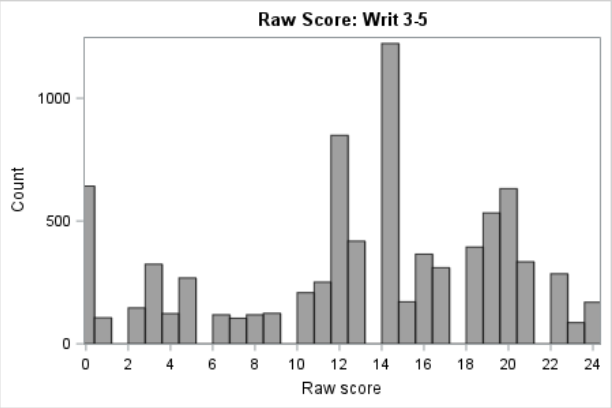


Table 6.2.4A

Raw Score Descriptive Statistics: Writ 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	3011	0	24	12.22	6.57
4	2794	0	24	12.98	6.6
5	2496	0	24	13.65	6.59
Total	8301	0	24	12.91	6.61

Figure 6.2.4B

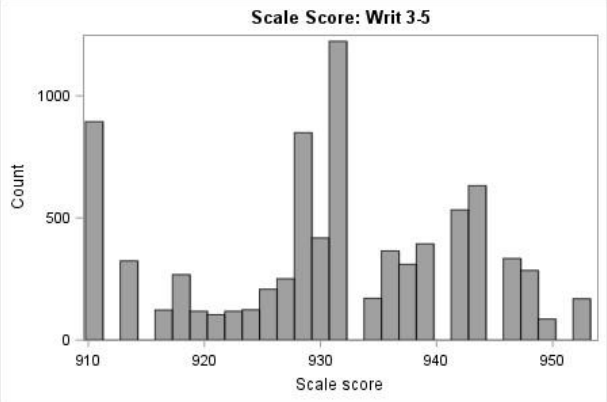


Table 6.2.4B

Scale Score Descriptive Statistics: Writ 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	3011	910	953	929.62	11.39
4	2794	910	953	931.04	11.58
5	2496	910	953	932.24	11.66
Total	8301	910	953	930.88	11.59

Figure 6.2.4C

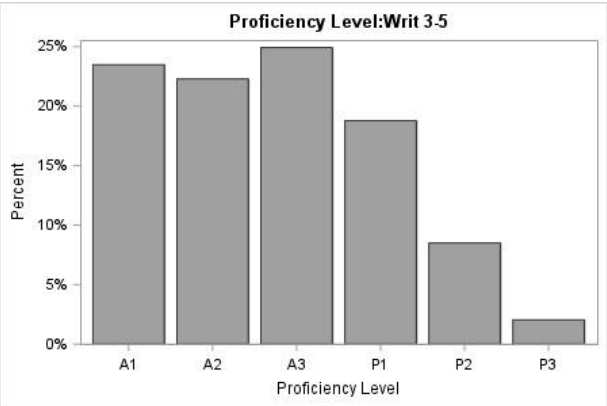


Table 6.2.4C

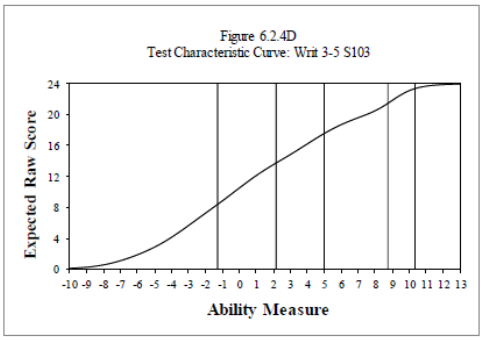
Proficiency Level Distribution: Writ 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	794	26.37%	646	23.12%	509	20.39%	1949	23.48%
A2	736	24.44%	605	21.65%	509	20.39%	1850	22.29%
A3	726	24.11%	704	25.2%	639	25.6%	2069	24.92%
P1	523	17.37%	538	19.26%	498	19.95%	1559	18.78%
P2	188	6.24%	241	8.63%	276	11.06%	705	8.49%
P3	44	1.46%	60	2.15%	65	2.6%	169	2.04%
Total	3011	100.0%	2794	100.0%	2496	100.0%	8301	100.0%

Table 6.2.4D

Equating Summary: Writ 3-5

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).



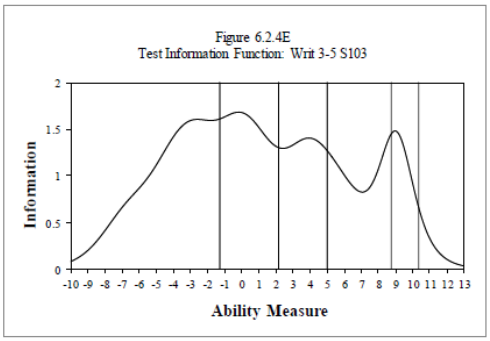


Table 6.2.4E

Reliability: Writ 3-5

No. of Students	No. of Items	Cronbach's Alpha	SEM
8301	10	0.9379	1.648

Table 6.2.4F

Item Analysis Summary: Writ 3-5

Note: The contents of this table have been redacted in this version of the document.

Table 6.2.4G

Complete Item Analysis: Writ 3-5

Note: The contents of this table have been redacted in this version of the document.

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	8.7	8.7	A1	7.41	7.41	A1	6.93	6.93
1	A1	1.53	10.23	A1	1.18	8.59	A1	1.08	8.01
2	A1	2.19	12.42	A1	1.47	10.06	A1	1.56	9.58
3	A1	4.15	16.57	A1	4.08	14.14	A1	3.41	12.98
4	A1	1.36	17.93	A1	1.9	16.03	A1	1.16	14.14
5	A1	3.35	21.29	A1	3.65	19.69	A1	2.6	16.75
6	A1	1.89	23.18	A1	1.18	20.87	A1	1.12	17.87
7	A1	1.46	24.64	A1	1.11	21.98	A1	1.16	19.03
8	A1	1.73	26.37	A1	1.15	23.12	A1	1.36	20.39
9	A2	1.69	28.06	A2	1.75	24.87	A2	0.96	21.35
10	A2	2.79	30.85	A2	2.72	27.59	A2	1.92	23.28
11	A2	3.39	34.24	A2	2.93	30.53	A2	2.68	25.96
12	A2	10.96	45.2	A2	9.7	40.23	A2	9.94	35.9
13	A2	5.61	50.81	A2	4.55	44.77	A2	4.89	40.79
14	A3	14.02	64.83	A3	14.92	59.7	A3	15.38	56.17
15	A3	1.86	66.69	A3	2.29	61.99	A3	2.04	58.21
16	A3	3.99	70.67	A3	4.4	66.39	A3	4.89	63.1
17	A3	4.25	74.93	A3	3.58	69.97	A3	3.29	66.39
18	P1	4.98	79.91	P1	4.94	74.91	P1	4.25	70.63
19	P1	6.61	86.52	P1	6.73	81.64	P1	5.85	76.48
20	P1	5.78	92.29	P1	7.59	89.23	P1	9.86	86.34
21	P2	3.29	95.58	P2	4.22	93.45	P2	4.69	91.03
22	P2	2.19	97.77	P2	3.36	96.81	P2	5.01	96.03
23	P2	0.76	98.54	P2	1.04	97.85	P2	1.36	97.4
24	P3	1.46	100	P3	2.15	100	P3	2.6	100

Table 6.2.4J

Accuracy and Consistency of Classification Indices: Writ 3-5

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.624	0.561	0.445
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.870	0.158
	A2	0.639	0.245
	A3	0.668	0.196
	P1	0.474	0.454
	P2	-	0.500

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.949	0.029	0.022	0.927
	A2/A3	0.921	0.039	0.040	0.892
	A3/P1	0.914	0.021	0.065	0.879
	P1/P2	0.835	0.165	0.000	0.835

6.2.5 Oral Language Composite 3-5

Figure 6.2.5A

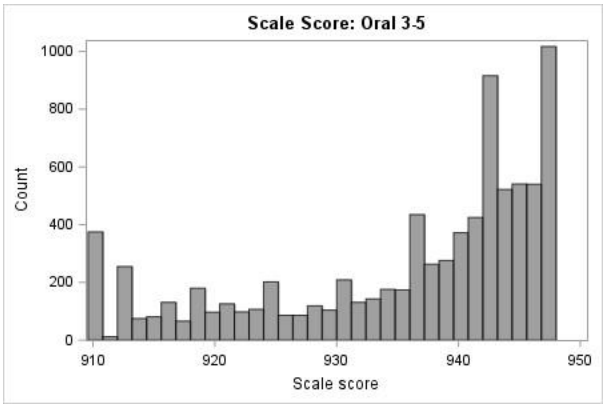


Table 6.2.5A

Scale Score Descriptive Statistics: Oral 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	3025	910	947	934.38	11.63
4	2806	910	947	935.54	11.23
5	2509	910	947	936.82	11.01
Total	8340	910	947	935.5	11.35

Figure 6.2.5B

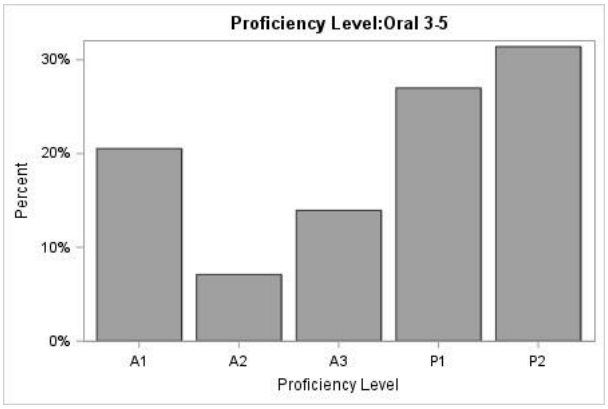


Table 6.2.5C

Proficiency Level Distribution: Oral 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	726	24.0%	561	19.99%	426	16.98%	1713	20.54%
A2	202	6.68%	213	7.59%	177	7.05%	592	7.1%
A3	435	14.38%	400	14.26%	328	13.07%	1163	13.94%
P1	856	28.3%	770	27.44%	626	24.95%	2252	27.0%
P2	806	26.64%	862	30.72%	952	37.94%	2620	31.41%
Total	3025	100.0%	2806	100.0%	2509	100.0%	8340	100.0%

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	115.4736	0.9369
Speaking	0.5	189.0765	0.9664
Oral		128.9117	0.9735

*Variances from students who had results in all four domains

Table 6.2.5F
n/a

Table 6.2.5G
n/a

Table 6.2.5H
n/a

Table 6.2.5I
n/a

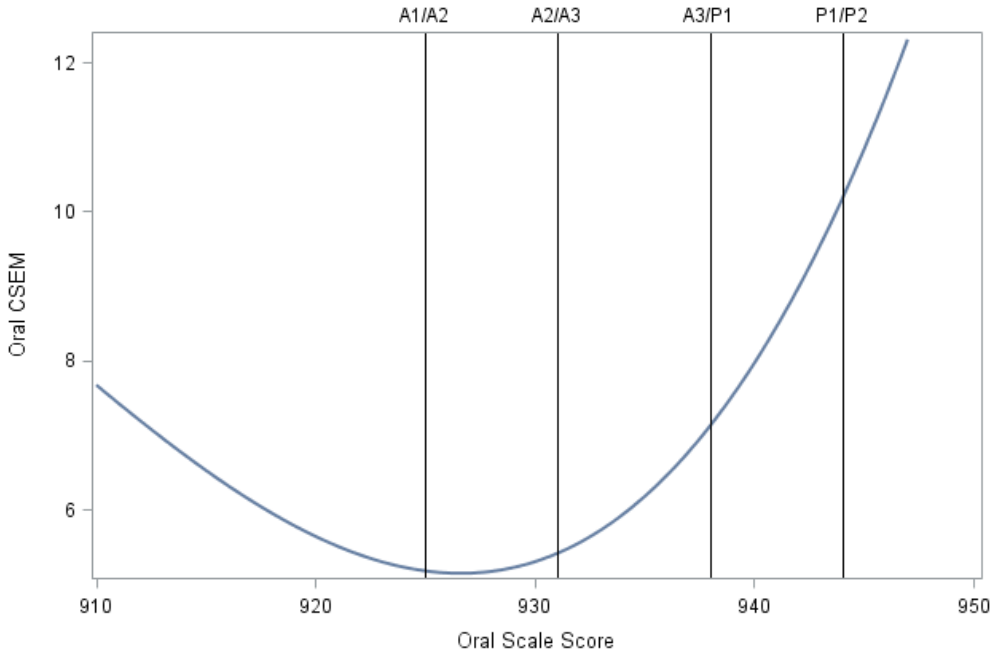
Table 6.2.5J

Accuracy and Consistency of Classification Indices: Oral 3-5

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.607	0.608	0.467
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.949	0.923
	A2	0.632	0.509
	A3	0.799	0.706
	P1	0.487	0.500
	P2	-	0.569

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.983	0.009	0.009	0.975
	A2/A3	0.976	0.013	0.011	0.966
	A3/P1	0.969	0.010	0.021	0.957
	P1/P2	0.680	0.320	0.000	0.708

Figure 6.2.5F CSEM for Oral Composite 3-5



6.2.6 Literacy Composite 3-5

Figure 6.2.6A

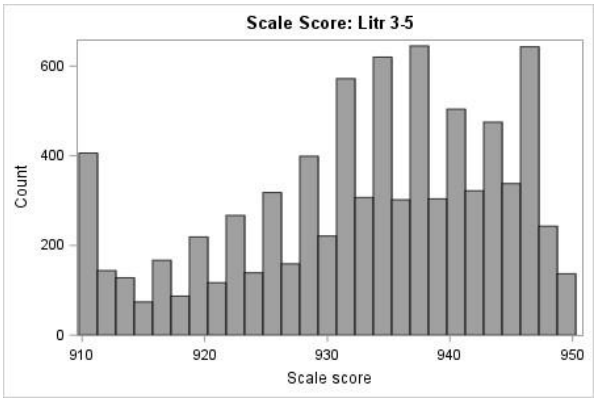


Table 6.2.6A

Scale Score Descriptive Statistics: Litr 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2985	910	950	931.84	10.65
4	2787	910	950	933.35	10.64
5	2485	910	950	934.6	10.51
Total	8257	910	950	933.18	10.66

Figure 6.2.6B

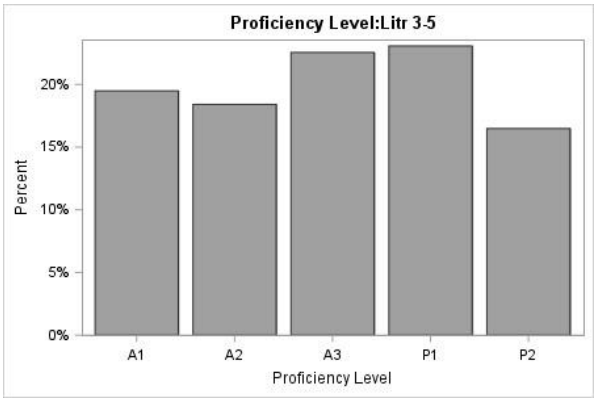


Table 6.2.6C

Proficiency Level Distribution: Litr 3-5

Level	Grade 3		Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	667	22.35%	539	19.34%	403	16.22%	1609	19.49%
A2	616	20.64%	522	18.73%	382	15.37%	1520	18.41%
A3	675	22.61%	598	21.46%	589	23.7%	1862	22.55%
P1	662	22.18%	655	23.5%	588	23.66%	1905	23.07%
P2	365	12.23%	473	16.97%	523	21.05%	1361	16.48%
Total	2985	100.0%	2787	100.0%	2485	100.0%	8257	100.0%

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	123.5688	0.9478
Writing	0.5	134.2701	0.9379
Literacy		113.7256	0.9675

*Variances from students who had results in all four domains

Table 6.2.6F
n/a

Table 6.2.6G
n/a

Table 6.2.6H
n/a

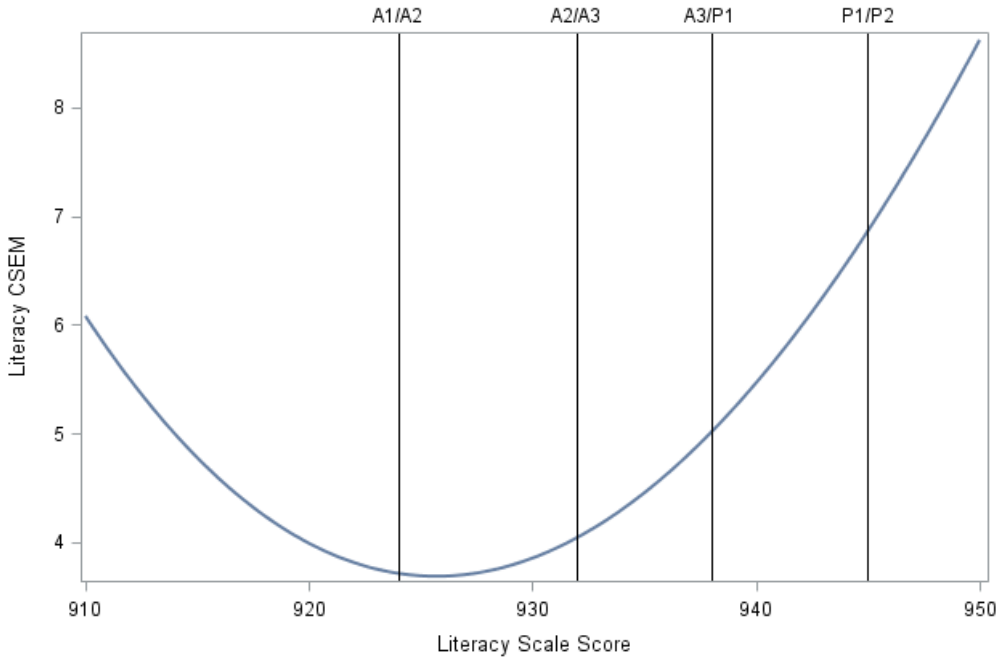
Table 6.2.6I
n/a

Table 6.2.6J

Accuracy and Consistency of Classification Indices: Litr 3-5

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

Figure 6.2.6F CSEM for Literacy Composite 3-5



6.2.7 Comprehension Composite 3-5

Figure 6.2.7A

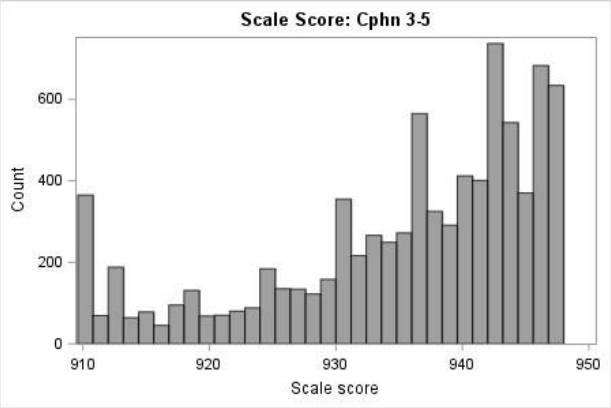


Table 6.2.7A

Scale Score Descriptive Statistics: Cphn 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	3043	910	947	933.79	10.86
4	2821	910	947	935.32	10.52
5	2529	910	947	936.47	10.36
Total	8393	910	947	935.11	10.65

Figure 6.2.7B

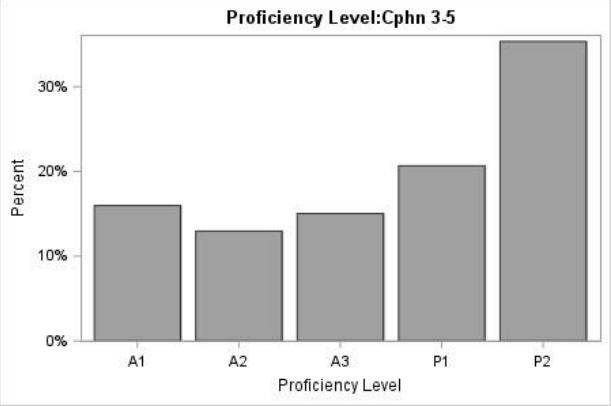


Table 6.2.7C

Proficiency Level Distribution: Cphn 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	582	19.13%	433	15.35%	326	12.89%	1341	15.98%
A2	431	14.16%	370	13.12%	287	11.35%	1088	12.96%
A3	474	15.58%	435	15.42%	353	13.96%	1262	15.04%
P1	661	21.72%	552	19.57%	522	20.64%	1735	20.67%
P2	895	29.41%	1031	36.55%	1041	41.16%	2967	35.35%
Total	3043	100.0%	2821	100.0%	2529	100.0%	8393	100.0%

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	115.4736	0.9369
Reading	0.7	123.5688	0.9478
Comprehension		113.5225	0.9664

*Variances from students who had results in all four domains

Table 6.2.7F
n/a

Table 6.2.7G
n/a

Table 6.2.7H
n/a

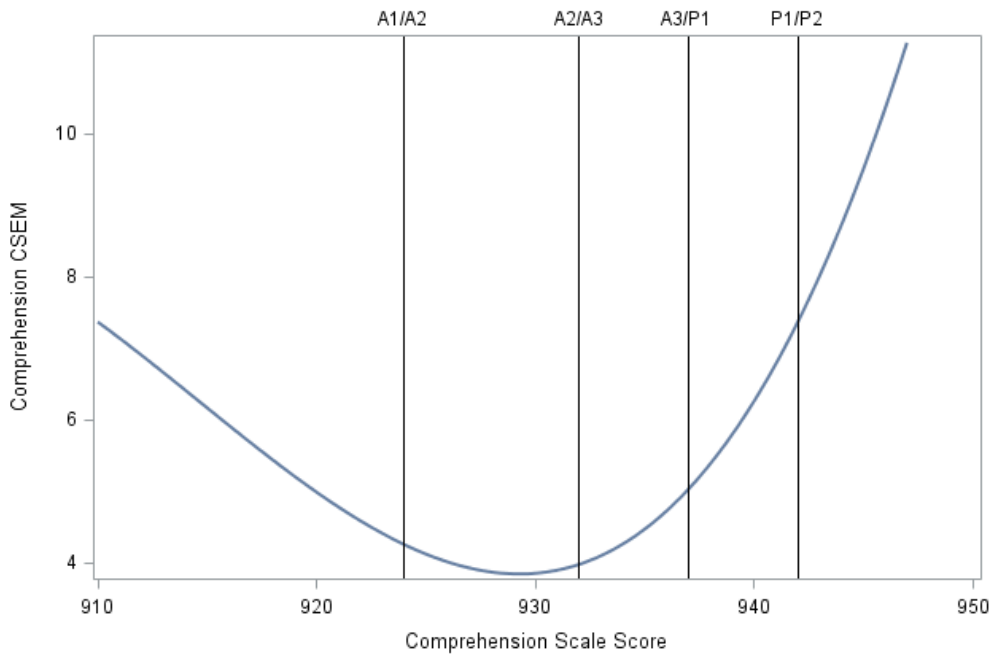
Table 6.2.7I
n/a

Table 6.2.7J

Accuracy and Consistency of Classification Indices: Cphn 3-5

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

Figure 6.2.7F CSEM for Comprehension Composite 3-5



6.2.8 Overall Composite 3-5

Figure 6.2.8A

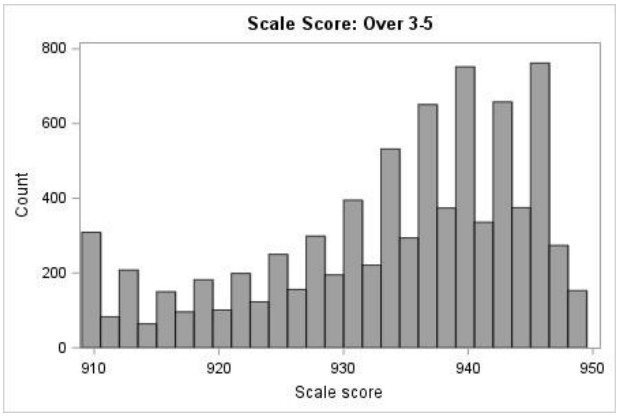


Table 6.2.8A

Scale Score Descriptive Statistics: Over 3-5

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
3	2961	910	949	932.42	10.57
4	2765	910	949	933.85	10.43
5	2466	910	949	935.12	10.29
Total	8192	910	949	933.72	10.5

Figure 6.2.8B

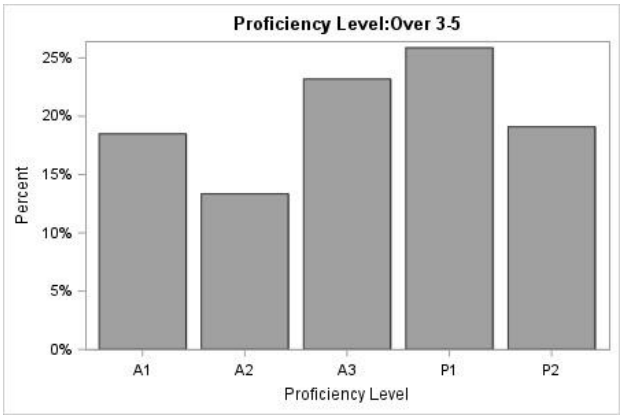


Table 6.2.8C

Proficiency Level Distribution: Over 3-5

	Grade 3		Grade 4		Grade 5		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	648	21.88%	489	17.69%	378	15.33%	1515	18.49%
A2	420	14.18%	387	14.0%	286	11.6%	1093	13.34%
A3	705	23.81%	642	23.22%	553	22.42%	1900	23.19%
P1	764	25.8%	701	25.35%	655	26.56%	2120	25.88%
P2	424	14.32%	546	19.75%	594	24.09%	1564	19.09%
Total	2961	100.0%	2765	100.0%	2466	100.0%	8192	100.0%

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	115.4736	0.9369
Reading	0.35	123.5688	0.9478
Speaking	0.15	189.0765	0.9664
Writing	0.35	134.2701	0.9379
Overall Composite		110.1649	0.9808

*Variances from students who had results in all four domains

Table 6.2.8F
n/a

Table 6.2.8G
n/a

Table 6.2.8H
n/a

Table 6.2.8I
n/a

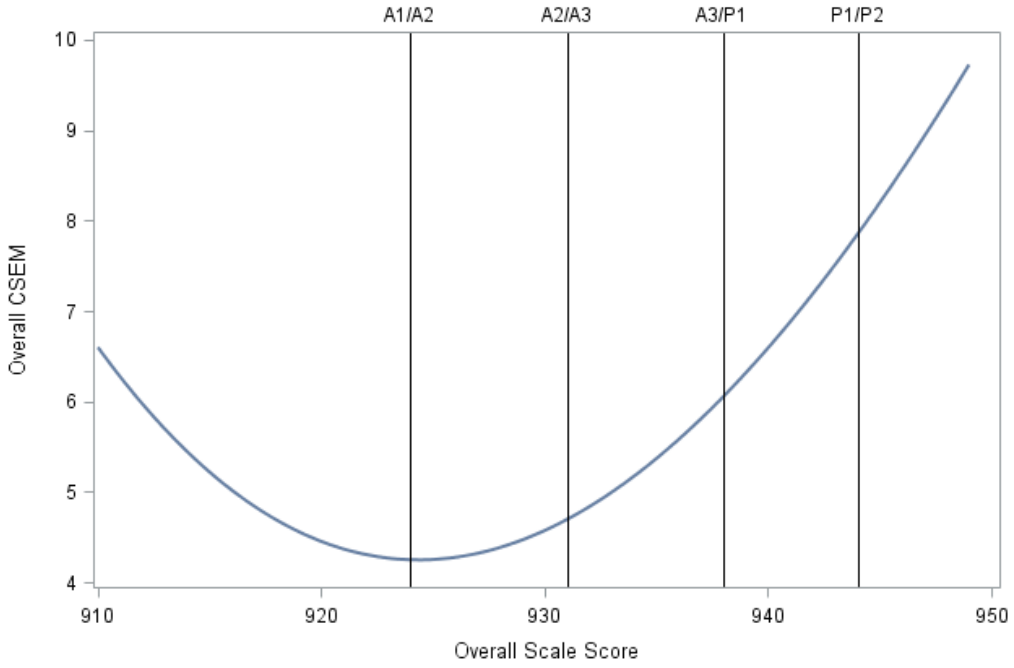
Table 6.2.8J

Accuracy and Consistency of Classification Indices: Over 3-5

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.708	0.671	0.567
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.941	0.913
	A2	0.763	0.666
	A3	0.876	0.814
	P1	0.561	0.550
	P2	-	0.499

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.982	0.010	0.008	0.974
	A2/A3	0.969	0.017	0.014	0.955
	A3/P1	0.958	0.010	0.031	0.943
	P1/P2	0.799	0.201	0.000	0.799

Figure 6.2.8F CSEM for Overall Composite 3-5



6.3 Grades: 6-8

6.3.1 Listening 6-8

Figure 6.3.1A

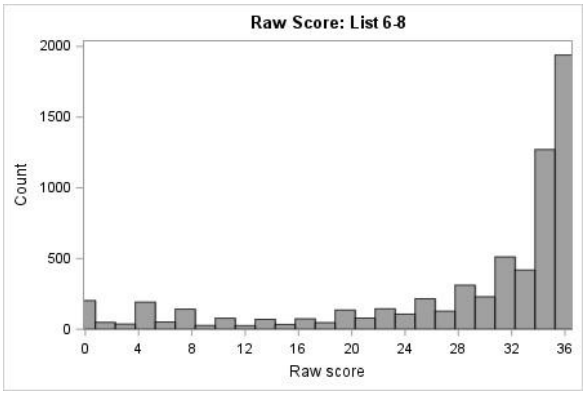


Table 6.3.1A

Raw Score Descriptive Statistics: List 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2286	0	36	27.82	10.63
7	2160	0	36	28.69	10.31
8	2078	0	36	29.09	10.17
Total	6524	0	36	28.51	10.39

Figure 6.3.1B

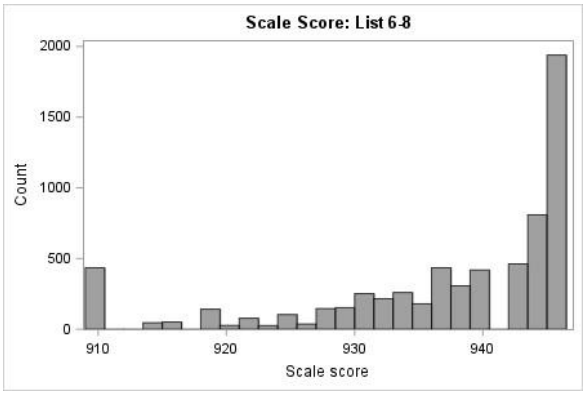


Table 6.3.1B

Scale Score Descriptive Statistics: List 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2286	910	946	936.25	10.64
7	2160	910	946	937.1	10.5
8	2078	910	946	937.66	10.36
Total	6524	910	946	936.98	10.52

Figure 6.3.1C

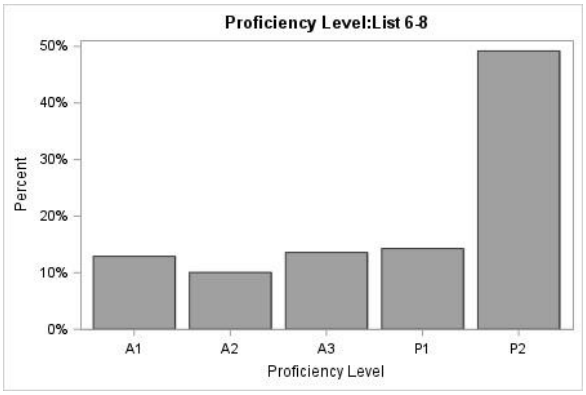


Table 6.3.1C

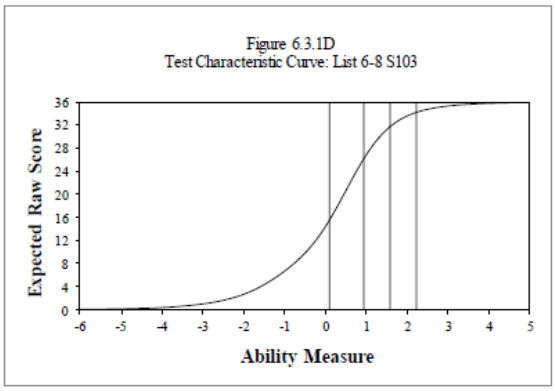
Proficiency Level Distribution: List 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	334	14.61%	267	12.36	240	11.55%	841	12.89%
A2	252	11.02	214	9.91%	190	9.14%	656	10.06%
A3	326	14.26%	288	13.33	272	13.09	886	13.58%
P1	325	14.22%	327	15.14	279	13.43	931	14.27%
P2	1049	45.89%	1064	49.26	1097	52.79	3210	49.2%
Total	2286	100.0%	2160	100.0	2078	100.0	6524	100.0%

Table 6.3.1D

Equating Summary: List 6-8

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 102. Thus, the results from the S102 of the Alternate ACCESS were used to determine raw-to-scale score conversion.



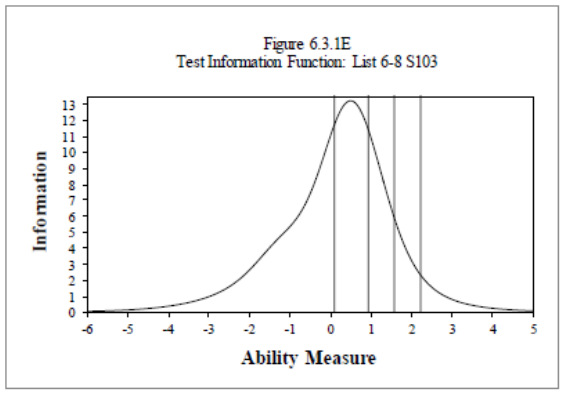


Table 6.3.1E

Reliability: List 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
6524	9	0.9438	2.4624

Table 6.3.1F

Item Analysis Summary: List 6-8

Note: The contents of this table have been redacted in this version of the document.

Table 6.3.1G

Complete Item Analysis: List 6-8

Note: The contents of this table have been redacted in this version of the document.

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	3.02	3.02	A1	3.06	3.06	A1	3.27	3.27
1	A1	0.22	3.24	A1	0.32	3.38	A1	0.38	3.66
2	A1	0.48	3.72	A1	0.37	3.75	A1	0.48	4.14
3	A1	0.61	4.33	A1	0.83	4.58	A1	0.24	4.38
4	A1	2.32	6.65	A1	2.31	6.9	A1	2.07	6.45
5	A1	0.48	7.13	A1	1.02	7.92	A1	0.63	7.07
6	A1	1.22	8.36	A1	0.6	8.52	A1	0.48	7.56
7	A1	2.19	10.54	A1	1.57	10.09	A1	1.49	9.05
8	A1	0.66	11.2	A1	0.32	10.42	A1	0.24	9.29
9	A1	0.79	11.99	A1	0.23	10.65	A1	0.19	9.48
10	A1	0.7	12.69	A1	0.28	10.93	A1	0.43	9.91
11	A1	0.83	13.52	A1	0.56	11.48	A1	0.77	10.68
12	A1	0.61	14.13	A1	0.42	11.9	A1	0.14	10.83
13	A1	0.48	14.61	A1	0.46	12.36	A1	0.72	11.55
14	A2	0.57	15.18	A2	0.42	12.78	A2	0.58	12.13
15	A2	0.74	15.92	A2	0.42	13.19	A2	0.38	12.51
16	A2	0.61	16.54	A2	0.42	13.61	A2	0.67	13.19
17	A2	0.57	17.1	A2	0.42	14.03	A2	0.72	13.91
18	A2	1.01	18.11	A2	0.65	14.68	A2	0.48	14.39
19	A2	0.83	18.94	A2	0.97	15.65	A2	1.06	15.45
20	A2	1.18	20.12	A2	1.16	16.81	A2	1.06	16.51
21	A2	1.4	21.52	A2	1.25	18.06	A2	0.96	17.47
22	A2	0.87	22.4	A2	1.39	19.44	A2	0.77	18.24
23	A2	1.27	23.67	A2	1.16	20.6	A2	1.2	19.44
24	A2	1.97	25.63	A2	1.67	22.27	A2	1.25	20.69
25	A3	1.4	27.03	A3	1.34	23.61	A3	1.73	22.43
26	A3	2.06	29.09	A3	1.57	25.19	A3	1.83	24.25
27	A3	2.23	31.32	A3	1.99	27.18	A3	1.64	25.89
28	A3	2.06	33.38	A3	1.9	29.07	A3	2.12	28.01
29	A3	2.67	36.05	A3	2.64	31.71	A3	2.98	30.99
30	A3	3.85	39.9	A3	3.89	35.6	A3	2.79	33.78
31	P1	3.28	43.18	P1	3.1	38.7	P1	3.03	36.81
32	P1	4.55	47.73	P1	5.6	44.31	P1	3.95	40.76
33	P1	6.39	54.11	P1	6.44	50.74	P1	6.45	47.21
34	P2	7.39	61.5	P2	7.04	57.78	P2	6.79	53.99
35	P2	13.04	74.54	P2	12.27	70.05	P2	11.84	65.83
36	P2	25.46	100	P2	29.95	100	P2	34.17	100

Table 6.3.1J

Accuracy and Consistency of Classification Indices: List 6-8

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.748	0.639	0.464
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.890	0.129
	A2	0.606	0.229
	A3	0.651	0.161
	P1	0.448	0.202
	P2	0.807	0.781

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.978	0.014	0.008	0.967
	A2/A3	0.959	0.023	0.018	0.945
	A3/P1	0.950	0.013	0.038	0.931
	P1/P2	0.855	0.029	0.116	0.770

6.3.2 Reading 6-8

Figure 6.3.2A

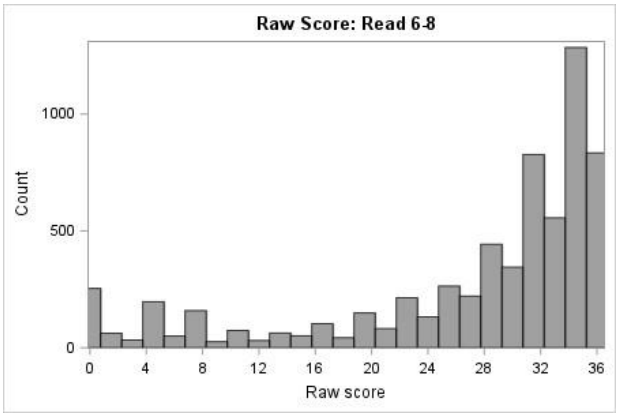


Table 6.3.2A

Raw Score Descriptive Statistics: Read 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2273	0	36	26.26	10.39
7	2150	0	36	27.1	10.38
8	2075	0	36	27.55	10.28
Total	6498	0	36	26.95	10.36

Figure 6.3.2B

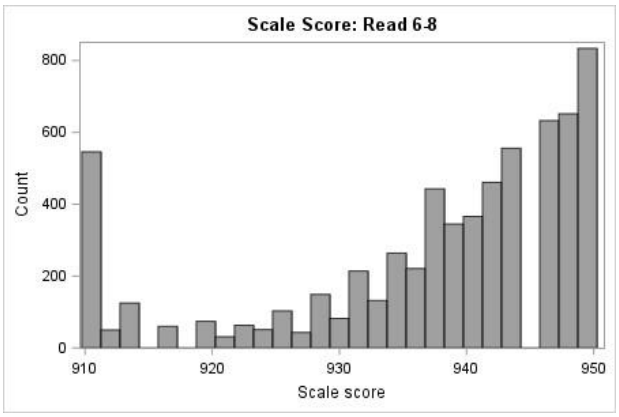


Table 6.3.2B

Scale Score Descriptive Statistics: Read 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2273	910	950	936.48	11.92
7	2150	910	950	937.58	12.01
8	2075	910	950	938.16	11.89
Total	6498	910	950	937.38	11.96

Figure 6.3.2C

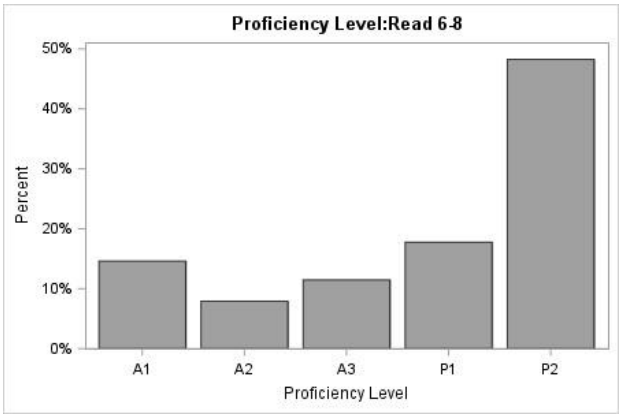


Table 6.3.2C

Proficiency Level Distribution: Read 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	361	15.88%	304	14.14%	284	13.69%	949	14.6%
A2	203	8.93%	172	8.0%	139	6.7%	514	7.91%
A3	289	12.71%	234	10.88%	222	10.7%	745	11.47%
P1	429	18.87%	379	17.63%	346	16.67%	1154	17.76%
P2	991	43.6%	1061	49.35%	1084	52.24%	3136	48.26%
Total	2273	100.0%	2150	100.0%	2075	100.0%	6498	100.0%

Table 6.3.2D

Equating Summary: Read 6-8

No equating summary is presented because the Alternate ACCESS Series 601 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

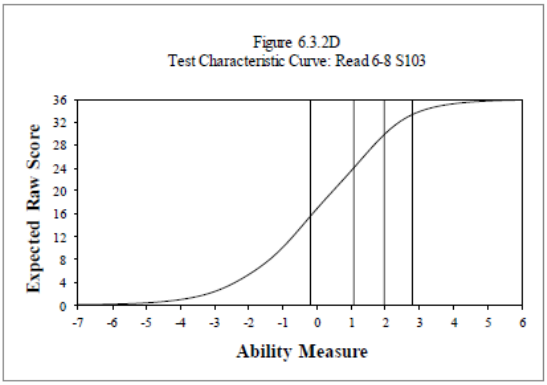


Figure 6.3.2E
Test Information Function: Read 6-8 S103

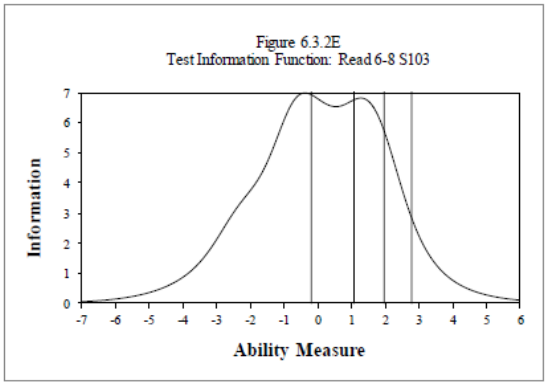


Table 6.3.2E

Reliability: Read 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
6498	9	0.942	2.4957

Table 6.3.2F

Item Analysis Summary: Read 6-8

Note: The contents of this table have been redacted in this version of the document.

Table 6.3.2G

Complete Item Analysis: Read 6-8

Note: The contents of this table have been redacted in this version of the document.

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	3.87	3.87	A1	3.81	3.81	A1	4.05	4.05
1	A1	0.31	4.18	A1	0.28	4.09	A1	0.19	4.24
2	A1	0.66	4.84	A1	0.79	4.88	A1	0.63	4.87
3	A1	0.44	5.28	A1	0.51	5.4	A1	0.58	5.45
4	A1	2.38	7.66	A1	2.28	7.67	A1	1.93	7.37
5	A1	1.19	8.84	A1	0.74	8.42	A1	0.53	7.9
6	A1	0.84	9.68	A1	0.88	9.3	A1	0.58	8.48
7	A1	1.54	11.22	A1	2.05	11.35	A1	2.22	10.7
8	A1	0.48	11.7	A1	0.51	11.86	A1	0.58	11.28
9	A1	0.53	12.23	A1	0.42	12.28	A1	0.24	11.52
10	A1	0.4	12.63	A1	0.33	12.6	A1	0.39	11.9
11	A1	0.92	13.55	A1	0.51	13.12	A1	0.87	12.77
12	A1	0.75	14.3	A1	0.19	13.3	A1	0.48	13.25
13	A1	0.66	14.96	A1	0.28	13.58	A1	0.19	13.45
14	A1	0.92	15.88	A1	0.56	14.14	A1	0.24	13.69
15	A2	0.75	16.63	A2	0.88	15.02	A2	0.72	14.41
16	A2	1.01	17.64	A2	0.65	15.67	A2	0.29	14.7
17	A2	1.14	18.79	A2	0.98	16.65	A2	0.63	15.33
18	A2	0.62	19.4	A2	0.6	17.26	A2	0.77	16.1
19	A2	1.41	20.81	A2	1.07	18.33	A2	0.67	16.77
20	A2	1.28	22.09	A2	1.16	19.49	A2	1.25	18.02
21	A2	1.28	23.36	A2	1.26	20.74	A2	1.25	19.28
22	A2	1.45	24.81	A2	1.4	22.14	A2	1.11	20.39
23	A3	1.98	26.79	A3	2	24.14	A3	1.93	22.31
24	A3	2.55	29.34	A3	1.63	25.77	A3	1.88	24.19
25	A3	2.42	31.76	A3	1.91	27.67	A3	1.59	25.78
26	A3	2.33	34.1	A3	1.81	29.49	A3	2.07	27.86
27	A3	3.43	37.53	A3	3.53	33.02	A3	3.23	31.08
28	P1	3.96	41.49	P1	3.26	36.28	P1	2.22	33.3
29	P1	3.78	45.27	P1	3.67	39.95	P1	3.47	36.77
30	P1	5.54	50.81	P1	5.07	45.02	P1	5.3	42.07
31	P1	5.59	56.4	P1	5.63	50.65	P1	5.69	47.76
32	P2	7.22	63.62	P2	6.84	57.49	P2	7.23	54.99
33	P2	7.79	71.4	P2	8.93	66.42	P2	9.01	64
34	P2	9.46	80.86	P2	9.49	75.91	P2	10.31	74.31
35	P2	8.97	89.84	P2	10.74	86.65	P2	10.46	84.77
36	P2	10.16	100	P2	13.35	100	P2	15.23	100

Table 6.3.2J

Accuracy and Consistency of Classification Indices: Read 6-8

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.786	0.720	0.561
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.927	0.152
	A2	0.564	0.184
	A3	0.538	0.197
	P1	0.561	0.157
	P2	0.847	0.825

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.980	0.011	0.009	0.971
	A2/A3	0.963	0.022	0.015	0.948
	A3/P1	0.944	0.024	0.032	0.925
	P1/P2	0.888	0.019	0.093	0.845

6.3.3 Speaking 6-8

Figure 6.3.3A

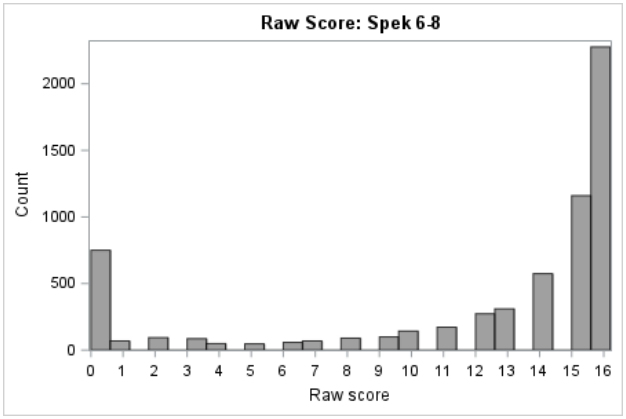


Table 6.3.3A

Raw Score Descriptive Statistics: Spek 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2244	0	16	11.8	5.6
7	2129	0	16	12.02	5.61
8	2055	0	16	12.07	5.54
Total	6428	0	16	11.96	5.59

Figure 6.3.3B

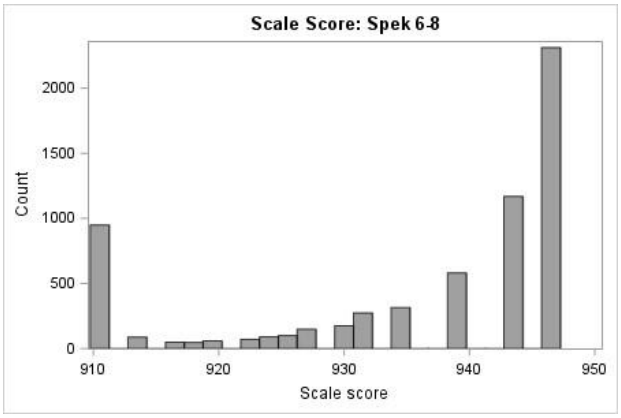


Table 6.3.3B

Scale Score Descriptive Statistics: Spek 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2244	910	947	935.37	13.26
7	2129	910	947	936.07	13.35
8	2055	910	947	936.18	13.22
Total	6428	910	947	935.86	13.28

Figure 6.3.3C

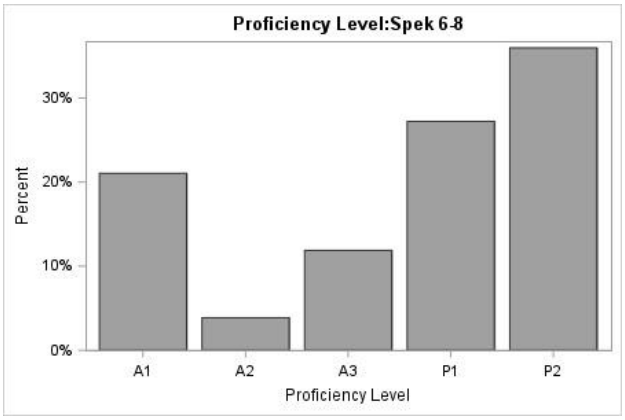


Table 6.3.3C

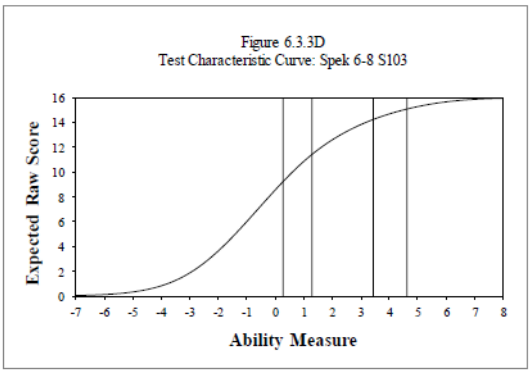
Proficiency Level Distribution: Spek 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	491	21.88%	434	20.39%	428	20.83%	1353	21.05%
A2	82	3.65%	87	4.09%	80	3.89%	249	3.87%
A3	297	13.24%	242	11.37%	225	10.95%	764	11.89%
P1	637	28.39%	576	27.05%	537	26.13%	1750	27.22%
P2	737	32.84%	790	37.11%	785	38.2%	2312	35.97%
Total	2244	100.0%	2129	100.0%	2055	100.0%	6428	100.0%

Table 6.3.3D

Equating Summary: Spek 6-8

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).



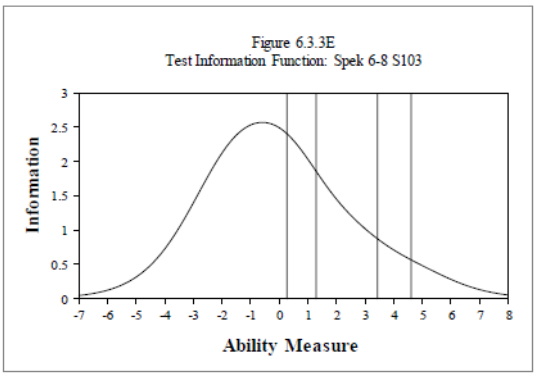


Table 6.3.3E

Reliability: Spek 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
6428	8	0.9645	1.0521

Table 6.3.3F

Item Analysis Summary: Spek 6-8

Note: The contents of this table have been redacted in this version of the document.

Table 6.3.3G

Complete Item Analysis: Spek 6-8

Note: The contents of this table have been redacted in this version of the document.

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	12.3	12.3	A1	12.17	12.17	A1	11.92	11.92
1	A1	1.25	13.55	A1	1.17	13.34	A1	0.97	12.9
2	A1	1.43	14.97	A1	1.78	15.12	A1	1.27	14.16
3	A1	1.47	16.44	A1	1.36	16.49	A1	1.27	15.43
4	A1	1.02	17.47	A1	0.61	17.1	A1	0.63	16.06
5	A1	0.76	18.23	A1	0.7	17.8	A1	0.73	16.79
6	A1	0.89	19.12	A1	0.61	18.41	A1	1.27	18.05
7	A1	1.34	20.45	A1	0.94	19.35	A1	1.02	19.08
8	A1	1.43	21.88	A1	1.03	20.39	A1	1.75	20.83
9	A2	1.25	23.13	A2	1.74	22.12	A2	1.7	22.53
10	A2	2.41	25.53	A2	2.35	24.47	A2	2.19	24.72
11	A3	3.16	28.7	A3	2.63	27.1	A3	2.34	27.06
12	A3	4.63	33.33	A3	3.85	30.95	A3	4.28	31.34
13	A3	5.44	38.77	A3	4.88	35.84	A3	4.33	35.67
14	P1	10.07	48.84	P1	8.13	43.96	P1	8.86	44.53
15	P1	18.32	67.16	P1	18.93	62.89	P1	17.27	61.8
16	P2	32.84	100	P2	37.11	100	P2	38.2	100

Table 6.3.3J

Accuracy and Consistency of Classification Indices: Spek 6-8

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.560	0.579	0.428
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.945	0.239
	A2	0.513	0.149
	A3	0.708	0.080
	P1	0.419	0.406
	P2	-	0.568

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.980	0.012	0.009	0.971
	A2/A3	0.975	0.012	0.014	0.965
	A3/P1	0.957	0.011	0.032	0.935
	P1/P2	0.646	0.354	0.000	0.684

6.3.4 Writing 6-8

Figure 6.3.4A

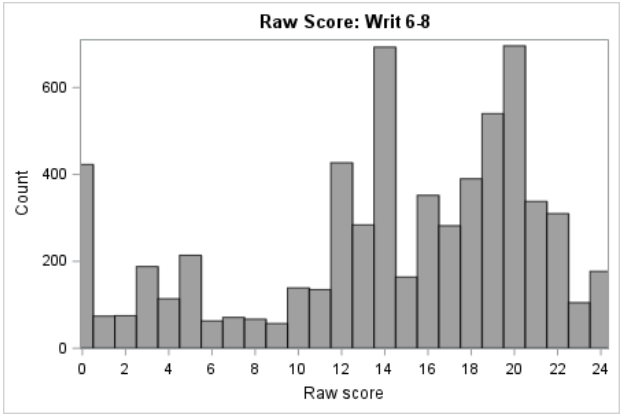


Table 6.3.4A

Raw Score Descriptive Statistics: Writ 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2224	0	24	13.62	6.61
7	2114	0	24	14.33	6.72
8	2040	0	24	14.56	6.72
Total	6378	0	24	14.16	6.69

Figure 6.3.4B

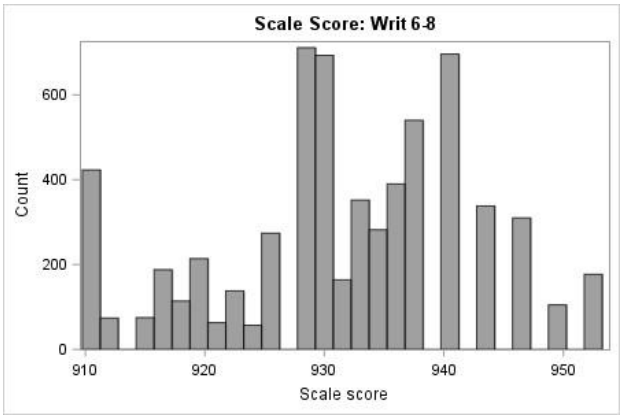


Table 6.3.4B

Scale Score Descriptive Statistics: Writ 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2224	910	953	930.77	10.07
7	2114	910	953	932.01	10.55
8	2040	910	953	932.43	10.58
Total	6378	910	953	931.71	10.42

Figure 6.3.4C

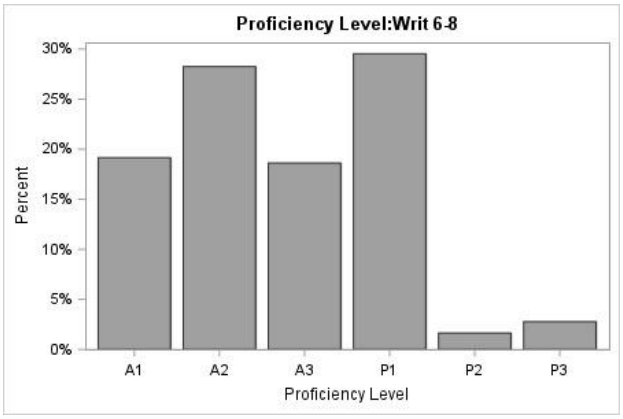


Table 6.3.4C

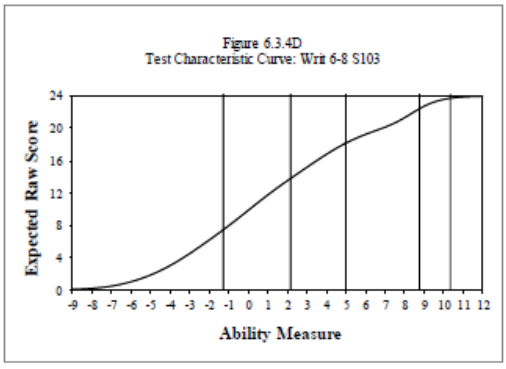
Proficiency Level Distribution: Writ 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	462	20.77%	387	18.31%	373	18.28%	1222	19.16%
A2	667	29.99%	597	28.24%	538	26.37%	1802	28.25%
A3	455	20.46%	371	17.55%	362	17.75%	1188	18.63%
P1	573	25.76%	652	30.84%	659	32.3%	1884	29.54%
P2	30	1.35%	39	1.84%	36	1.76%	105	1.65%
P3	37	1.66%	68	3.22%	72	3.53%	177	2.78%
Total	2224	100.0%	2114	100.0%	2040	100.0%	6378	100.0%

Table 6.3.4D

Equating Summary: Writ 6-8

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).



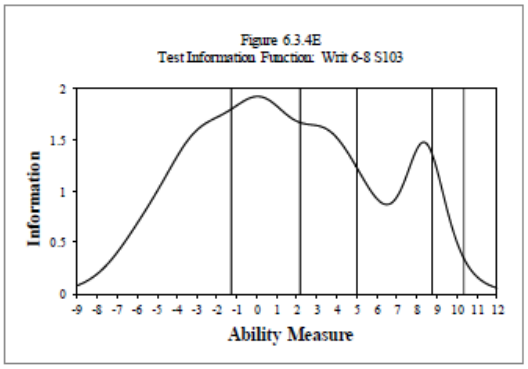


Table 6.3.4E

Reliability: Writ 6-8

No. of Students	No. of Items	Cronbach's Alpha	SEM
6378	10	0.9398	1.6426

Table 6.3.4F

Item Analysis Summary: Writ 6-8

Note: The contents of this table have been redacted in this version of the document.

Table 6.3.4G

Complete Item Analysis: Writ 6-8

Note: The contents of this table have been redacted in this version of the document.

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	6.79	6.79	A1	6.62	6.62	A1	6.47	6.47
1	A1	1.48	8.27	A1	1.32	7.95	A1	0.64	7.11
2	A1	1.12	9.4	A1	1.09	9.04	A1	1.32	8.43
3	A1	3.24	12.63	A1	2.89	11.92	A1	2.7	11.13
4	A1	1.8	14.43	A1	1.94	13.86	A1	1.62	12.75
5	A1	3.46	17.9	A1	2.93	16.79	A1	3.68	16.42
6	A1	1.39	19.29	A1	0.71	17.5	A1	0.83	17.25
7	A1	1.48	20.77	A1	0.8	18.31	A1	1.03	18.28
8	A2	1.21	21.99	A2	0.95	19.25	A2	0.98	19.26
9	A2	1.03	23.02	A2	0.9	20.15	A2	0.74	20
10	A2	2.61	25.63	A2	1.75	21.9	A2	2.16	22.16
11	A2	2.43	28.06	A2	2.22	24.12	A2	1.67	23.82
12	A2	6.83	34.89	A2	6.86	30.98	A2	6.37	30.2
13	A2	4.32	39.21	A2	5.2	36.19	A2	3.82	34.02
14	A2	11.56	50.76	A2	10.36	46.55	A2	10.64	44.66
15	A3	2.65	53.42	A3	2.74	49.29	A3	2.3	46.96
16	A3	6.43	59.85	A3	4.92	54.21	A3	5.15	52.11
17	A3	4.59	64.43	A3	4.64	58.85	A3	4.02	56.13
18	A3	6.79	71.22	A3	5.25	64.1	A3	6.27	62.4
19	P1	7.01	78.24	P1	9.27	73.37	P1	9.22	71.62
20	P1	10.3	88.53	P1	10.97	84.34	P1	11.52	83.14
21	P1	5.08	93.62	P1	5.44	89.78	P1	5.39	88.53
22	P1	3.37	96.99	P1	5.16	94.94	P1	6.18	94.71
23	P2	1.35	98.34	P2	1.84	96.78	P2	1.76	96.47
24	P3	1.66	100	P3	3.22	100	P3	3.53	100

Table 6.3.4J

Accuracy and Consistency of Classification Indices: Writ 6-8

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.748	0.652	0.531
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.858	0.129
	A2	0.776	0.300
	A3	0.524	0.121
	P1	0.765	0.775
	P2	-	0.190

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.953	0.030	0.018	0.932
	A2/A3	0.932	0.022	0.046	0.907
	A3/P1	0.925	0.031	0.044	0.888
	P1/P2	0.936	0.064	0.000	0.908

6.3.5 Oral Language Composite 6-8

Figure 6.3.5A

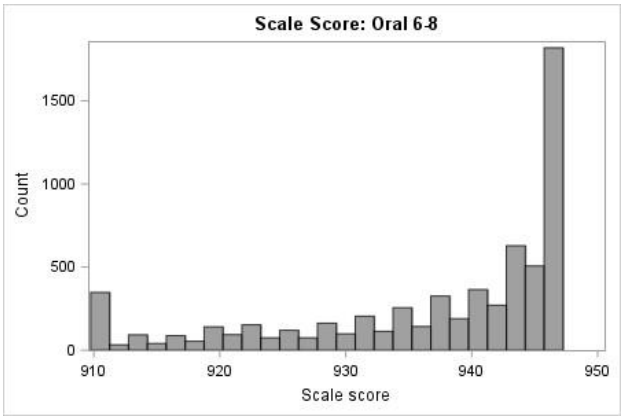


Table 6.3.5A

Scale Score Descriptive Statistics: Oral 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2235	910	947	936.26	11.27
7	2122	910	947	937.06	11.32
8	2045	910	947	937.39	11.07
Total	6402	910	947	936.89	11.23

Figure 6.3.5B

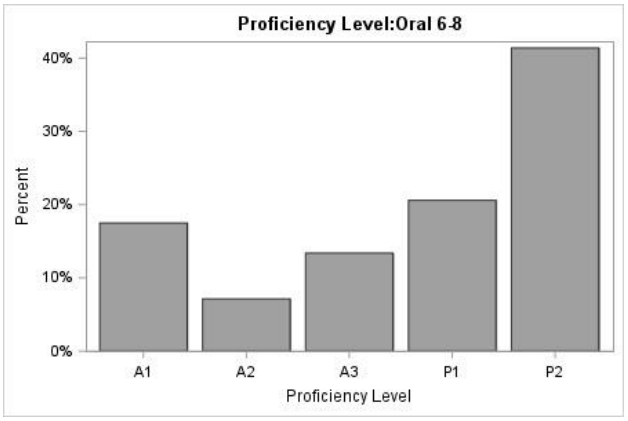


Table 6.3.5C

Proficiency Level Distribution: Oral 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	415	18.57%	373	17.58%	332	16.23%	1120	17.49%
A2	174	7.79%	131	6.17%	152	7.43%	457	7.14%
A3	321	14.36%	276	13.01%	260	12.71%	857	13.39%
P1	479	21.43%	447	21.07%	392	19.17%	1318	20.59%
P2	846	37.85%	895	42.18%	909	44.45%	2650	41.39%
Total	2235	100.0%	2122	100.0%	2045	100.0%	6402	100.0%

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	110.6927	0.9438
Speaking	0.5	176.3391	0.9645
Oral		126.1716	0.9753

*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

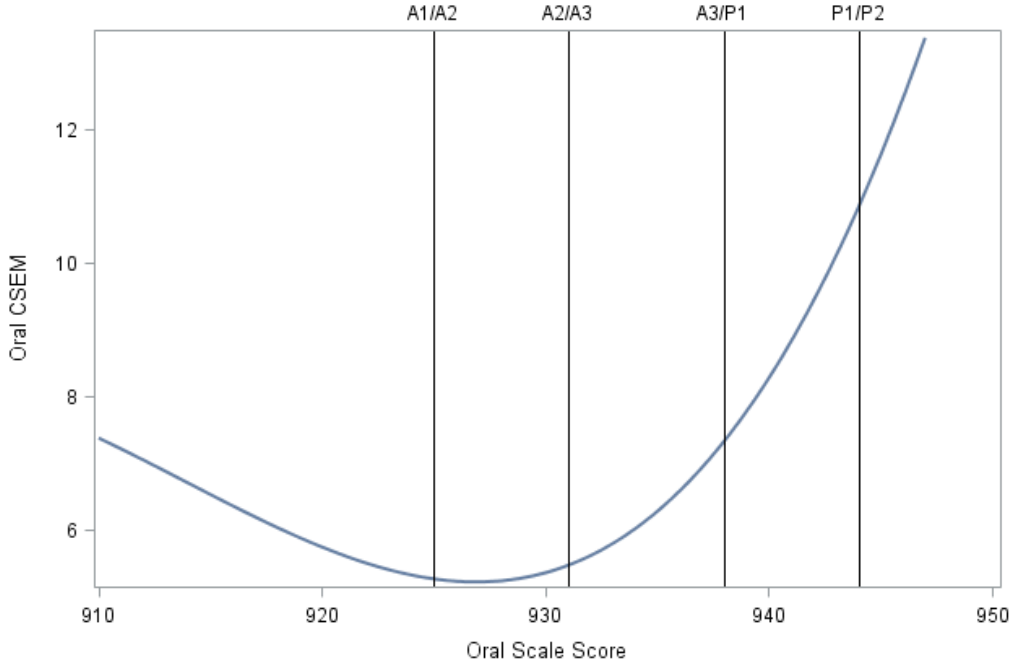
Table 6.3.5I
n/a

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

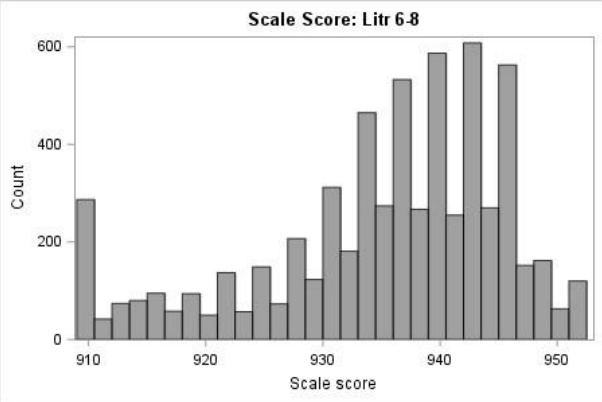


Table 6.3.6A

Scale Score Descriptive Statistics: Litr 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2210	910	952	933.95	10.23
7	2099	910	952	935.14	10.54
8	2029	910	952	935.63	10.48
Total	6338	910	952	934.88	10.44

Figure 6.3.6B

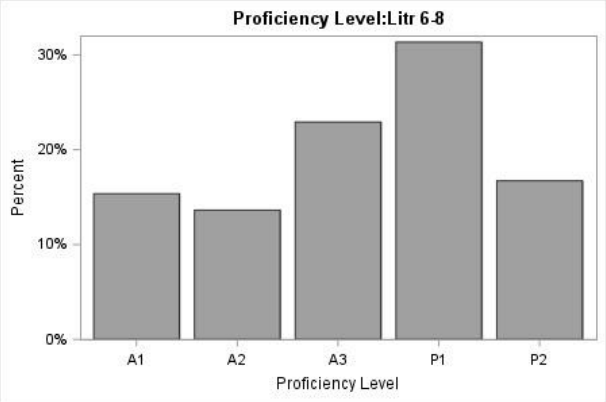


Table 6.3.6C

Proficiency Level Distribution: Litr 6-8

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	363	16.43%	319	15.2%	292	14.39%	974	15.37%
A2	355	16.06%	258	12.29%	251	12.37%	864	13.63%
A3	535	24.21%	480	22.87%	438	21.59%	1453	22.93%
P1	673	30.45%	676	32.21%	638	31.44%	1987	31.35%
P3	284	12.85%	366	17.44%	410	20.21%	1060	16.72%
Total	2210	100.0%	2099	100.0%	2029	100.0%	6338	100.0%

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	143.0367	0.942
Writing	0.5	108.5641	0.9398
Literacy		108.8919	0.9659

*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

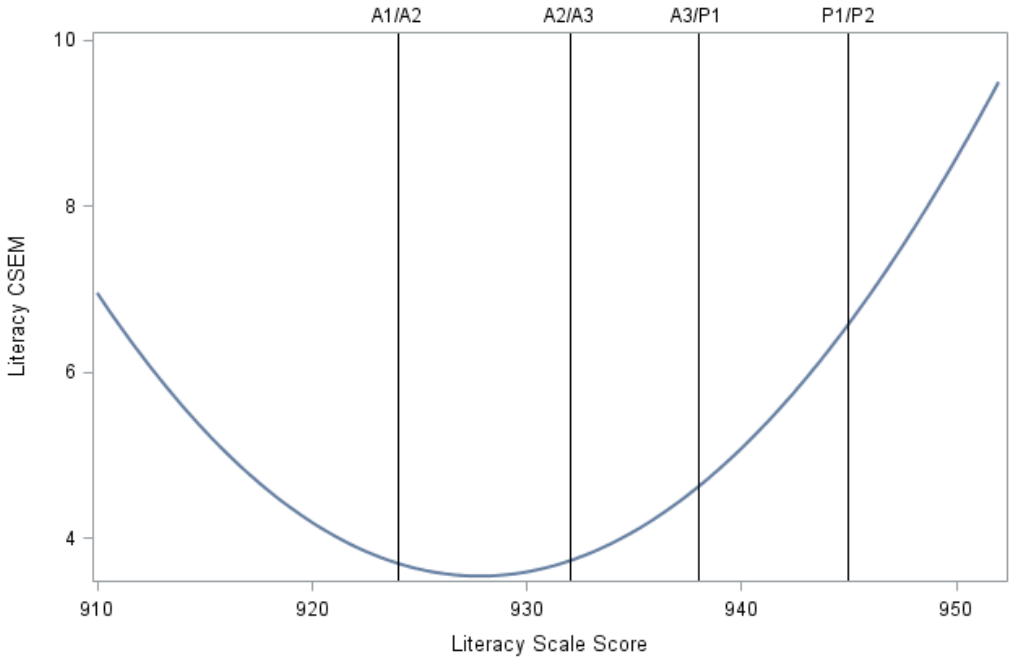
Table 6.3.6I
n/a

Table 6.3.6J

Accuracy and Consistency of Classification Indices: Litr 6-8

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**	
	0.638	0.616	0.601	
Conditional on Level	**Level**	**Accuracy**	**Consistency**	
	A1	0.922	0.883	
	A2	0.740	0.636	
	A3	0.794	0.696	
	P1	0.499	0.601	
	P2	-	0.557	
Indices at Cut Points		**Accuracy**		
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.979	0.012	0.009	0.970
	A2/A3	0.959	0.023	0.018	0.943
	A3/P1	0.948	0.013	0.039	0.929
	P1/P2	0.751	0.249	0.000	0.771

Figure 6.3.6F CSEM for Literacy Composite 6-8



6.3.7 Comprehension Composite 6-8

Figure 6.3.7A

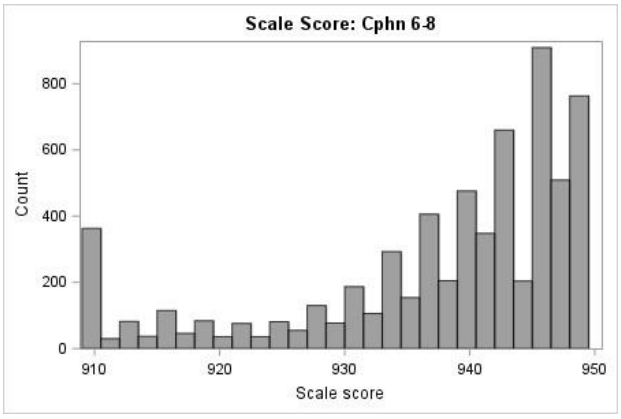


Table 6.3.7A

Scale Score Descriptive Statistics: Cphn 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2263	910	949	936.52	11.1
7	2141	910	949	937.53	11.18
8	2064	910	949	938.1	11.06
Total	6468	910	949	937.36	11.13

Figure 6.3.7B

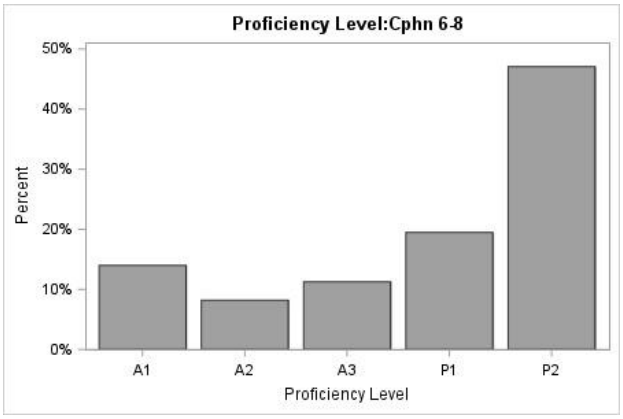


Table 6.3.7C

Proficiency Level Distribution: Cphn 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	337	14.89%	296	13.83%	272	13.18%	905	13.99%
A2	219	9.68%	169	7.89%	142	6.88%	530	8.19%
A3	286	12.64%	226	10.56%	216	10.47%	728	11.26%
P1	463	20.46%	408	19.06%	389	18.85%	1260	19.48%
P2	958	42.33%	1042	48.67%	1045	50.63%	3045	47.08%
Total	2263	100.0%	2141	100.0%	2064	100.0%	6468	100.0%

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	110.6927	0.9438
Reading	0.7	143.0367	0.942
Comprehension		123.9779	0.9627

*Variances from students who had results in all four domains

Table 6.3.7F
n/a

Table 6.3.7G
n/a

Table 6.3.7H
n/a

Table 6.3.7I
n/a

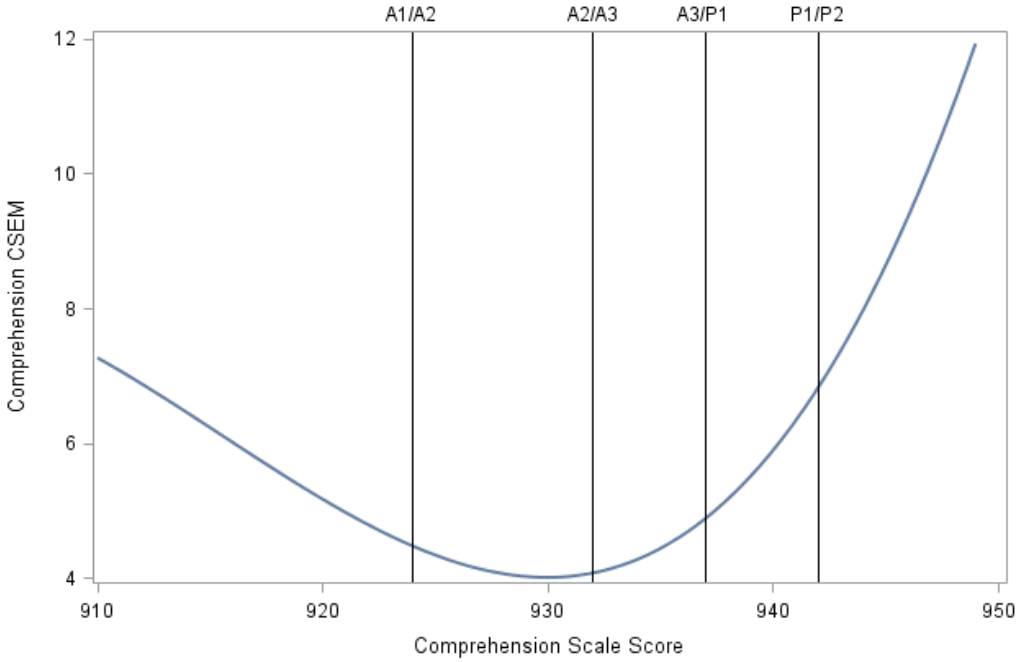
Table 6.3.7J

Accuracy and Consistency of Classification Indices: Cphn 6-8

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.823	0.763	0.633
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.948	0.920
	A2	0.645	0.518
	A3	0.603	0.478
	P1	0.707	0.542
	P2	0.870	0.849

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.985	0.007	0.008	0.979
	A2/A3	0.974	0.016	0.010	0.962
	A3/P1	0.955	0.022	0.023	0.938
	P1/P2	0.907	0.017	0.076	0.873

Figure 6.3.7F CSEM for Comprehension Composite 6-8



6.3.8 Overall Composite 6-8

Figure 6.3.8A

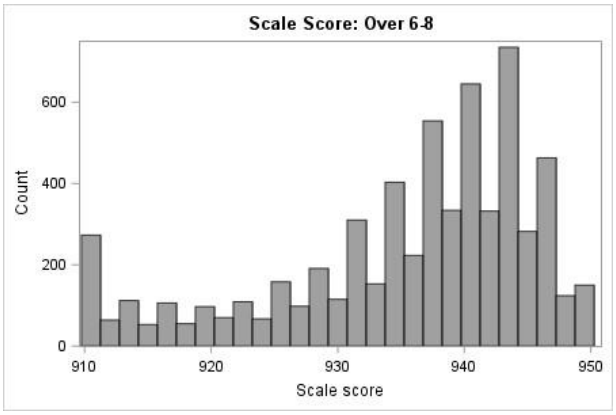


Table 6.3.8A

Scale Score Descriptive Statistics: Over 6-8

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2187	910	950	934.5	10.15
7	2081	910	950	935.54	10.39
8	2008	910	950	936.01	10.24
Total	6276	910	950	935.33	10.27

Figure 6.3.8B

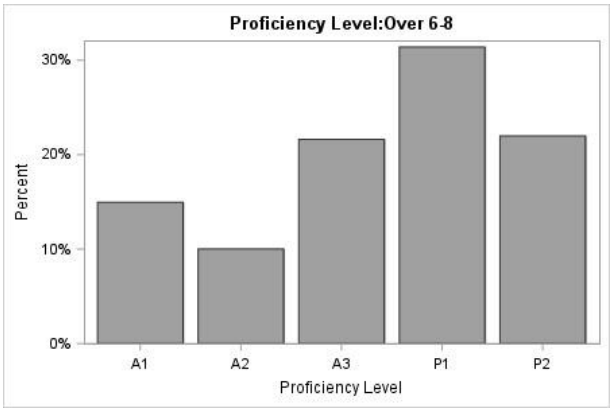


Table 6.3.8C

Proficiency Level Distribution: Over 6-8

	Grade 6		Grade 7		Grade 8		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	348	15.91%	313	15.04%	278	13.84%	939	14.96%
A2	256	11.71%	182	8.75%	191	9.51%	629	10.02%
A3	516	23.59%	435	20.9%	406	20.22%	1357	21.62%
P1	682	31.18%	675	32.44%	614	30.58%	1971	31.41%
P2	385	17.6%	476	22.87%	519	25.85%	1380	21.99%
Total	2187	100.0%	2081	100.0%	2008	100.0%	6276	100.0%

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	110.6927	0.9438
Reading	0.35	143.0367	0.942
Speaking	0.15	176.3391	0.9645
Writing	0.35	108.5641	0.9398
Overall Composite		105.5472	0.9801

*Variances from students who had results in all four domains

Table 6.3.8F
n/a

Table 6.3.8G
n/a

Table 6.3.8H
n/a

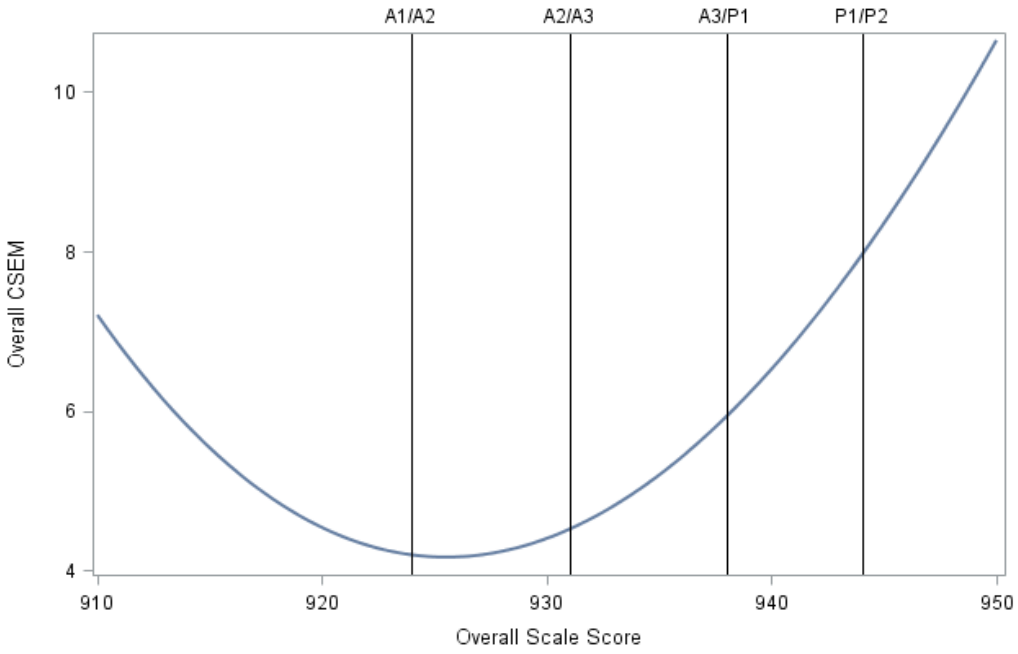
Table 6.3.8I
n/a

Table 6.3.8J

Accuracy and Consistency of Classification Indices: Over 6-8

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

Figure 6.3.8F CSEM for Overall Composite 6-8



6.4 Grades: 9-12

6.4.1 Listening 9-12

Figure 6.4.1A

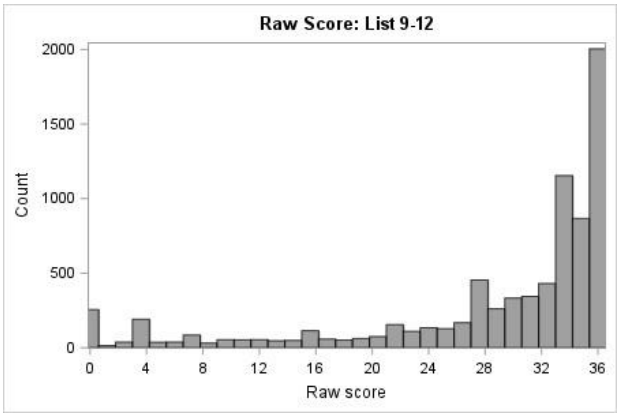


Table 6.4.1A

Raw Score Descriptive Statistics: List 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1960	0	36	27.91	10.27
10	1748	0	36	28.61	9.87
11	1534	0	36	28.74	9.78
12	2608	0	36	28.75	9.76
Total	7850	0	36	28.51	9.93

Figure 6.4.1B

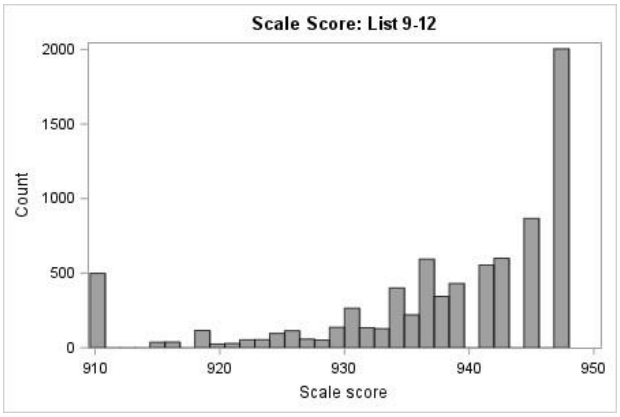


Table 6.4.1B

Scale Score Descriptive Statistics: List 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1960	910	947	936.98	10.57
10	1748	910	947	937.79	10.2
11	1534	910	947	937.89	10.21
12	2608	910	947	937.87	10.14
Total	7850	910	947	937.63	10.28

Figure 6.4.1C

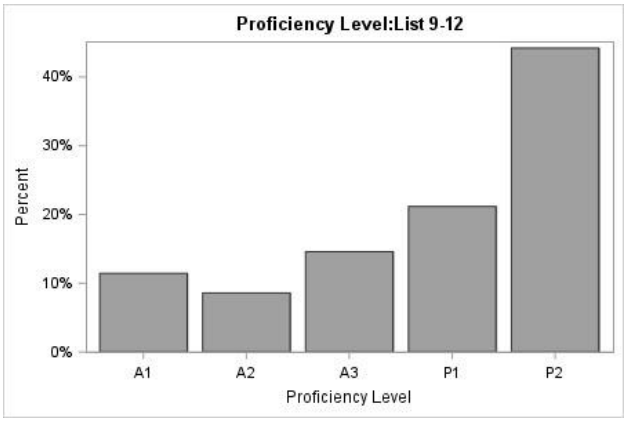


Table 6.4.1C

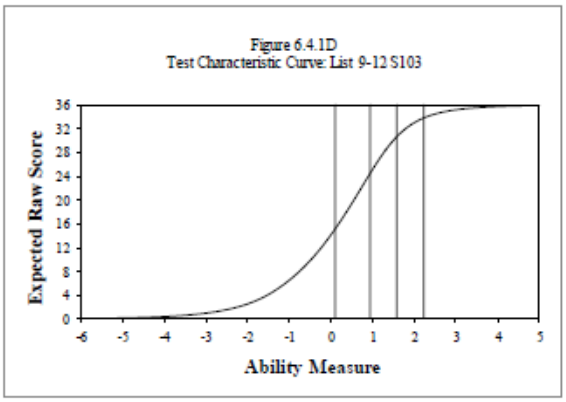
Proficiency Level Distribution: List 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	249	12.70%	201	11.50%	169	11.02%	279	10.70%	898	11.44%
A2	189	9.64%	146	8.35%	119	7.76%	220	8.44%	674	8.59%
A3	289	14.74%	236	13.50%	242	15.78%	378	14.49%	1145	14.59%
P1	418	21.33%	374	21.40%	307	20.01%	563	21.59%	1662	21.17%
P2	815	41.58%	791	45.25%	697	45.44%	1168	44.79%	3471	44.22%
Total	1960	100.00%	1748	100.00%	1534	100.00%	2608	100.00%	7850	100.00%

Table 6.4.1D

Equating Summary: List 9-12

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 102. Thus, the results from the S102 of the Alternate ACCESS were used to determine raw-to-scale score conversion.



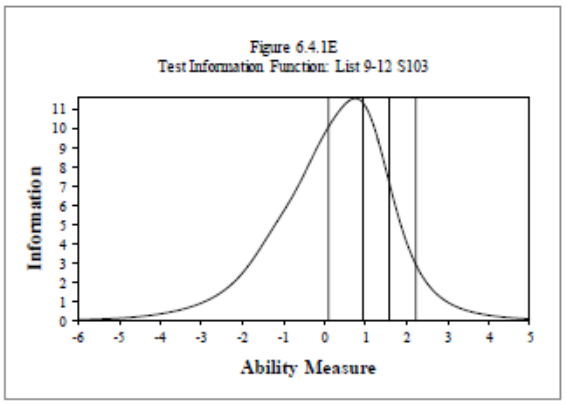


Table 6.4.1E

Reliability: List 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
7850	9	0.939	2.4518

Table 6.4.1F

Item Analysis Summary: List 9-12

Note: The contents of this table have been redacted in this version of the document.

Table 6.4.1G

Complete Item Analysis: List 9-12

Note: The contents of this table have been redacted in this version of the document.

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	3.72	3.72	A1	3.09	3.09	A1	3.06	3.06	A1	3.14	3.14
1	A1	0.26	3.98	A1	0.06	3.15	A1	0.13	3.19	A1	0.23	3.37
2	A1	0.77	4.74	A1	0.29	3.43	A1	0.59	3.78	A1	0.35	3.72
3	A1	0.36	5.1	A1	0.57	4	A1	0.59	4.37	A1	0.54	4.26
4	A1	2.14	7.24	A1	1.77	5.78	A1	1.83	6.19	A1	1.92	6.17
5	A1	0.41	7.65	A1	0.8	6.58	A1	0.59	6.78	A1	0.23	6.4
6	A1	0.41	8.06	A1	0.34	6.92	A1	0.39	7.17	A1	0.69	7.09
7	A1	1.38	9.44	A1	1.14	8.07	A1	1.17	8.34	A1	0.77	7.86
8	A1	0.51	9.95	A1	0.4	8.47	A1	0.26	8.6	A1	0.38	8.24
9	A1	0.36	10.31	A1	0.29	8.75	A1	0.26	8.87	A1	0.35	8.59
10	A1	0.56	10.87	A1	0.34	9.1	A1	0.07	8.93	A1	0.42	9.01
11	A1	0.56	11.43	A1	0.86	9.95	A1	0.72	9.65	A1	0.61	9.62
12	A1	0.66	12.09	A1	0.92	10.87	A1	0.78	10.43	A1	0.5	10.12
13	A1	0.61	12.7	A1	0.63	11.5	A1	0.59	11.02	A1	0.58	10.7
14	A2	0.71	13.42	A2	0.57	12.07	A2	0.39	11.41	A2	0.73	11.43
15	A2	0.56	13.98	A2	0.69	12.76	A2	0.52	11.93	A2	0.92	12.35
16	A2	0.82	14.8	A2	1.09	13.84	A2	0.52	12.45	A2	0.61	12.96
17	A2	0.97	15.77	A2	0.92	14.76	A2	0.59	13.04	A2	0.54	13.5
18	A2	0.41	16.17	A2	0.8	15.56	A2	0.98	14.02	A2	0.54	14.03
19	A2	0.71	16.89	A2	1.03	16.59	A2	0.46	14.47	A2	0.88	14.92
20	A2	1.28	18.16	A2	0.46	17.05	A2	0.98	15.45	A2	1.04	15.95
21	A2	1.22	19.39	A2	0.51	17.56	A2	1.17	16.62	A2	1.11	17.06
22	A2	1.33	20.71	A2	0.97	18.54	A2	0.72	17.34	A2	0.81	17.87
23	A2	1.63	22.35	A2	1.32	19.85	A2	1.43	18.77	A2	1.27	19.13
24	A3	1.63	23.98	A3	1.26	21.11	A3	1.96	20.73	A3	1.92	21.05
25	A3	1.63	25.61	A3	1.77	22.88	A3	1.83	22.56	A3	1.42	22.47
26	A3	2.6	28.21	A3	2.29	25.17	A3	1.96	24.51	A3	1.8	24.27
27	A3	3.01	31.22	A3	2.46	27.63	A3	3.26	27.77	A3	3.11	27.38
28	A3	2.76	33.98	A3	2.75	30.38	A3	3.26	31.03	A3	2.65	30.02
29	A3	3.11	37.09	A3	2.97	33.35	A3	3.52	34.55	A3	3.6	33.63
30	P1	4.18	41.28	P1	4.69	38.04	P1	4.63	39.18	P1	3.76	37.38
31	P1	4.74	46.02	P1	4.92	42.96	P1	3.72	42.89	P1	4.14	41.53
32	P1	5.51	51.53	P1	4.63	47.6	P1	5.15	48.04	P1	6.25	47.78
33	P1	6.89	58.42	P1	7.15	54.75	P1	6.52	54.56	P1	7.44	55.21
34	P2	8.16	66.58	P2	8.01	62.76	P2	5.93	60.5	P2	8.01	63.23
35	P2	10	76.58	P2	11.16	73.91	P2	12.91	73.4	P2	10.66	73.89
36	P2	23.42	100	P2	26.09	100	P2	26.6	100	P2	26.11	100

Table 6.4.1J

Accuracy and Consistency of Classification Indices: List 9-12

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.756	0.652	0.470

Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.913	0.150
	A2	0.456	0.146
	A3	0.730	0.220
	P1	0.399	0.166
	P2	0.838	0.810

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.981	0.010	0.009	0.971
	A2/A3	0.958	0.029	0.013	0.942
	A3/P1	0.936	0.014	0.050	0.916
	P1/P2	0.872	0.030	0.098	0.791

6.4.2 Reading 9-12

Figure 6.4.2A

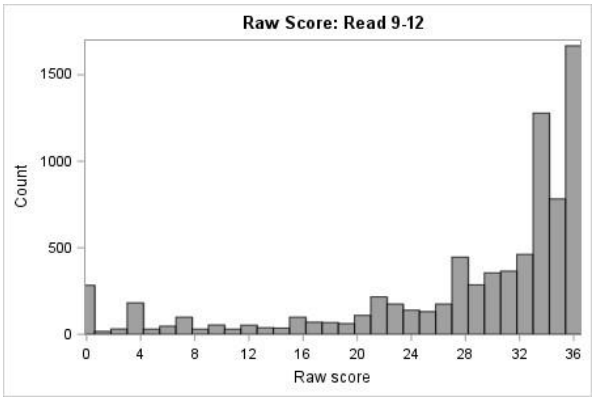


Table 6.4.2A

Raw Score Descriptive Statistics: Read 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1947	0	36	27.43	10.16
10	1745	0	36	28.14	9.91
11	1525	0	36	28.38	9.75
12	2598	0	36	28.26	9.85
Total	7815	0	36	28.05	9.93

Figure 6.4.2B

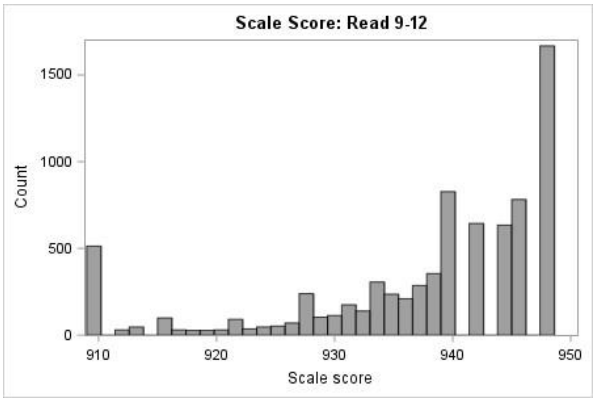


Table 6.4.2B

Scale Score Descriptive Statistics: Read 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1947	910	948	936.82	11.05
10	1745	910	948	937.61	10.86
11	1525	910	948	937.89	10.72
12	2598	910	948	937.81	10.83
Total	7815	910	948	937.54	10.88

Figure 6.4.2C

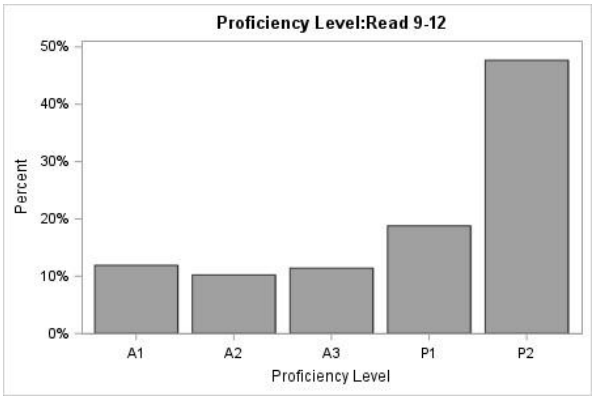


Table 6.4.2C

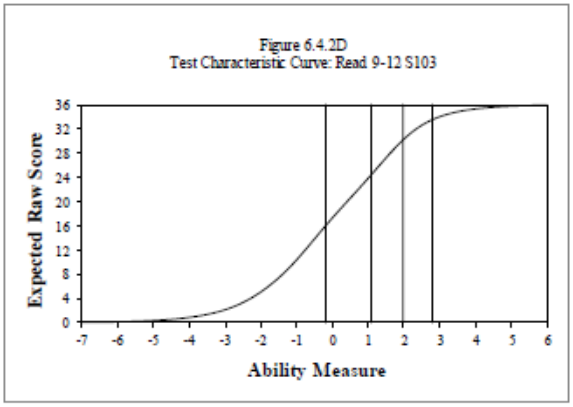
Proficiency Level Distribution: Read 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	249	12.79%	203	11.63%	173	11.34%	305	11.74%	930	11.90%
A2	227	11.66%	176	10.09%	136	8.92%	260	10.01%	799	10.22%
A3	242	12.43%	195	11.17%	180	11.80%	274	10.55%	891	11.40%
P1	358	18.39%	333	19.08%	298	19.54%	479	18.44%	1468	18.78%
P2	871	44.74%	838	48.02%	738	48.39%	1280	49.27%	3727	47.69%
Total	1947	100.00%	1745	100.00%	1525	100.00%	2598	100.00%	7815	100.00%

Table 6.4.2D

Equating Summary: Read 9-12

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).



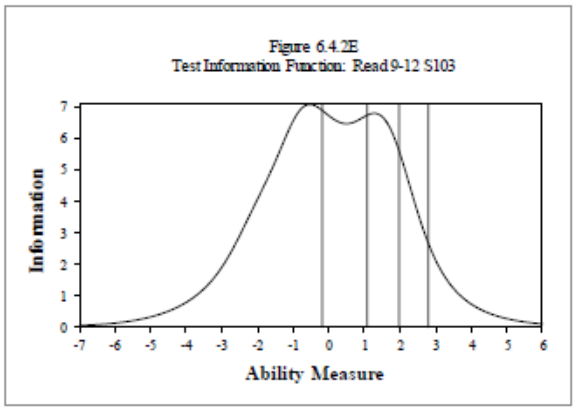


Table 6.4.2E

Reliability: Read 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
7815	9	0.9406	2.4186

Table 6.4.2F

Item Analysis Summary: Read 9-12

Note: The contents of this table have been redacted in this version of the document.

Table 6.4.2G

Complete Item Analysis: Read 9-12

Note: The contents of this table have been redacted in this version of the document.

Table 6.4.2H

Raw Score to Scale Score Conversion: Read 9-12

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

^ Truncated

* Adjusted for end of scale effect

Table 6.4.2I

Raw Score to Proficiency Level Conversion: Read 9-12

Raw Score	Grade 9			Grade 10			Grade 11			Grade 12		
	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Students	Proficiency Level Score	% of Students	Cumulative % of Student	Proficiency Level Score	% of Students	Cumulative % of Students
0	A1	4.06	4.06	A1	3.55	3.55	A1	3.41	3.41	A1	3.46	3.46
1	A1	0.21	4.26	A1	0.23	3.78	A1	0.33	3.74	A1	0.15	3.62
2	A1	0.51	4.78	A1	0.34	4.13	A1	0.39	4.13	A1	0.35	3.96
3	A1	0.72	5.5	A1	0.57	4.7	A1	0.92	5.05	A1	0.5	4.46
4	A1	1.44	6.93	A1	1.95	6.65	A1	1.44	6.49	A1	1.81	6.27
5	A1	0.41	7.34	A1	0.29	6.93	A1	0.2	6.69	A1	0.54	6.81
6	A1	0.72	8.06	A1	0.86	7.79	A1	0.59	7.28	A1	0.35	7.16
7	A1	1.54	9.6	A1	1.09	8.88	A1	1.31	8.59	A1	1.15	8.31
8	A1	0.46	10.07	A1	0.52	9.4	A1	0.33	8.92	A1	0.27	8.58
9	A1	0.31	10.37	A1	0.34	9.74	A1	0.13	9.05	A1	0.5	9.08
10	A1	0.56	10.94	A1	0.23	9.97	A1	0.39	9.44	A1	0.23	9.31
11	A1	0.36	11.3	A1	0.17	10.14	A1	0.2	9.64	A1	0.65	9.97
12	A1	0.62	11.92	A1	0.74	10.89	A1	0.52	10.16	A1	0.77	10.74
13	A1	0.51	12.43	A1	0.46	11.35	A1	0.59	10.75	A1	0.42	11.16
14	A1	0.36	12.79	A1	0.29	11.63	A1	0.59	11.34	A1	0.58	11.74
15	A2	0.67	13.46	A2	0.8	12.44	A2	0.52	11.87	A2	0.46	12.2
16	A2	0.72	14.18	A2	0.69	13.12	A2	0.39	12.26	A2	0.77	12.97
17	A2	1.13	15.31	A2	0.69	13.81	A2	0.72	12.98	A2	0.96	13.93
18	A2	1.08	16.38	A2	0.86	14.67	A2	0.59	13.57	A2	0.89	14.82
19	A2	0.72	17.1	A2	0.74	15.42	A2	0.85	14.43	A2	0.85	15.67
20	A2	2.05	19.16	A2	1.26	16.68	A2	0.52	14.95	A2	1.5	17.17
21	A2	1.49	20.65	A2	1.03	17.71	A2	1.7	16.66	A2	1.15	18.32
22	A2	1.64	22.29	A2	1.43	19.14	A2	1.44	18.1	A2	1.31	19.63
23	A2	2.16	24.45	A2	2.58	21.72	A2	2.16	20.26	A2	2.12	21.75
24	A3	2.16	26.61	A3	2.06	23.78	A3	1.57	21.84	A3	1.42	23.17
25	A3	1.64	28.25	A3	1.43	25.21	A3	2.23	24.07	A3	1.54	24.71
26	A3	2.52	30.77	A3	2.12	27.34	A3	2.56	26.62	A3	1.92	26.64
27	A3	3.49	34.26	A3	2.64	29.97	A3	2.82	29.44	A3	3.04	29.68
28	A3	2.62	36.88	A3	2.92	32.89	A3	2.62	32.07	A3	2.62	32.29
29	P1	3.8	40.68	P1	3.04	35.93	P1	3.61	35.67	P1	4	36.3
30	P1	4.42	45.1	P1	4.7	40.63	P1	5.05	40.72	P1	4.23	40.53
31	P1	4.47	49.56	P1	4.81	45.44	P1	4.66	45.38	P1	4.73	45.27
32	P1	5.7	55.26	P1	6.53	51.98	P1	6.23	51.61	P1	5.47	50.73
33	P2	7.4	62.66	P2	8.71	60.69	P2	8.52	60.13	P2	8.39	59.12
34	P2	8.99	71.65	P2	8.14	68.83	P2	7.54	67.67	P2	7.78	66.9
35	P2	10.02	81.66	P2	10.14	78.97	P2	9.18	76.85	P2	10.39	77.29
36	P2	18.34	100	P2	21.03	100	P2	23.15	100	P2	22.71	100

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

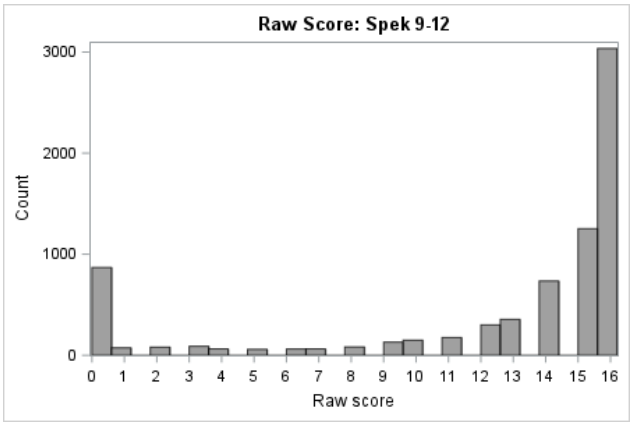


Table 6.4.3A

Raw Score Descriptive Statistics: Spek 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1929	0	16	11.94	5.67
10	1718	0	16	12.28	5.47
11	1515	0	16	12.45	5.35
12	2565	0	16	12.24	5.48
Total	7727	0	16	12.22	5.5

Figure 6.4.3B

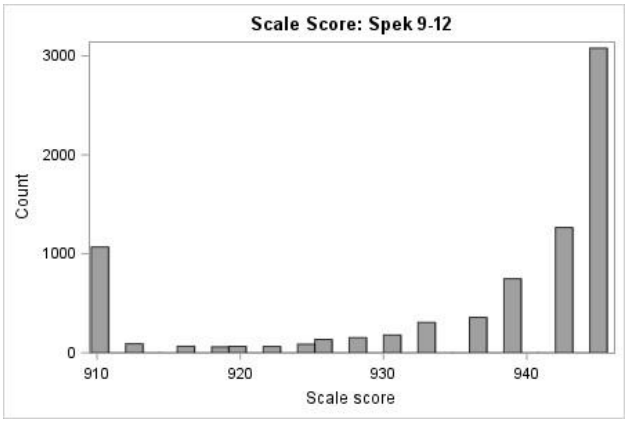


Table 6.4.3B

Scale Score Descriptive Statistics: Spek 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1929	910	945	935.08	12.77
10	1718	910	945	935.86	12.4
11	1515	910	945	936.24	12.13
12	2565	910	945	935.75	12.32
Total	7727	910	945	935.7	12.42

Figure 6.4.3C

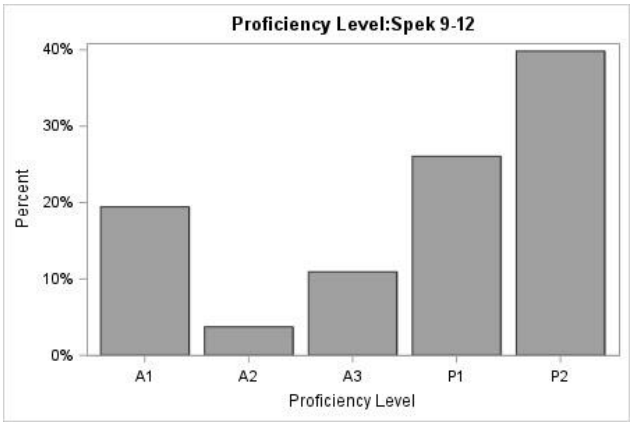


Table 6.4.3C

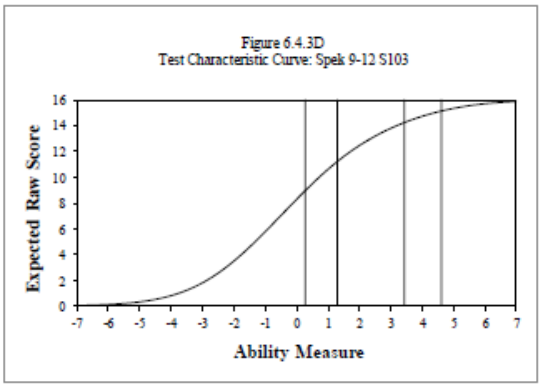
Proficiency Level Distribution: Spek 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	413	21.41%	337	19.62%	268	17.69%	484	18.87%	1502	19.44%
A2	73	3.78%	66	3.84%	55	3.63%	94	3.66%	288	3.73%
A3	197	10.21%	173	10.07%	176	11.62%	299	11.66%	845	10.94%
P1	515	26.70%	430	25.03%	374	24.69%	695	27.10%	2014	26.06%
P2	731	37.90%	712	41.44%	642	42.38%	993	38.71%	3078	39.83%
Total	1929	100.00%	1718	100.00%	1515	100.00%	2565	100.00%	7727	100.00%

Table 6.4.3D

Equating Summary: Spek 9-12

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).



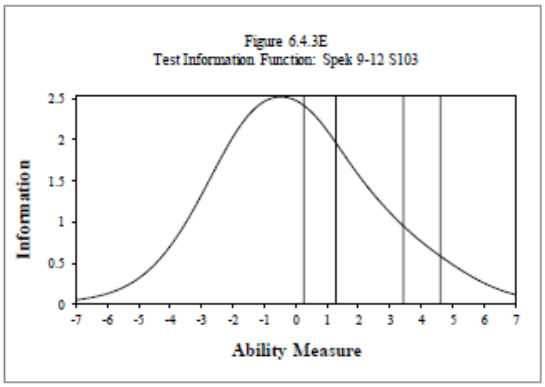


Table 6.4.3E

Reliability: Spek 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
7727	8	0.9667	1.0048

Table 6.4.3F

Item Analysis Summary: Spek 9-12

Note: The contents of this table have been redacted in this version of the document.

Table 6.4.3G

Complete Item Analysis: Spek 9-12

Note: The contents of this table have been redacted in this version of the document.

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	12.75	12.75	A1	11.23	11.23	A1	10.76	10.76	A1	11.93	11.93
1	A1	1.5	14.26	A1	0.87	12.11	A1	0.66	11.42	A1	0.9	12.83
2	A1	1.14	15.4	A1	0.87	12.98	A1	1.45	12.87	A1	0.97	13.8
3	A1	1.14	16.54	A1	1.69	14.67	A1	0.79	13.66	A1	1.09	14.89
4	A1	0.78	17.31	A1	0.93	15.6	A1	1.06	14.72	A1	0.74	15.63
5	A1	0.47	17.78	A1	0.87	16.47	A1	0.66	15.38	A1	1.01	16.65
6	A1	1.4	19.18	A1	0.93	17.4	A1	0.66	16.04	A1	0.47	17.12
7	A1	0.98	20.17	A1	0.87	18.28	A1	0.86	16.9	A1	0.66	17.78
8	A1	1.24	21.41	A1	1.34	19.62	A1	0.79	17.69	A1	1.09	18.87
9	A2	1.87	23.28	A2	1.69	21.3	A2	1.72	19.41	A2	1.72	20.58
10	A2	1.92	25.19	A2	2.15	23.46	A2	1.91	21.32	A2	1.95	22.53
11	A3	2.59	27.79	A3	2.15	25.61	A3	2.77	24.09	A3	1.99	24.52
12	A3	3.47	31.26	A3	3.67	29.28	A3	3.76	27.85	A3	4.68	29.2
13	A3	4.15	35.41	A3	4.25	33.53	A3	5.08	32.94	A3	4.99	34.19
14	P1	10.52	45.93	P1	8.38	41.91	P1	9.11	42.05	P1	10.25	44.44
15	P1	16.17	62.1	P1	16.65	58.56	P1	15.58	57.62	P1	16.84	61.29
16	P2	37.9	100	P2	41.44	100	P2	42.38	100	P2	38.71	100

Table 6.4.3J

Accuracy and Consistency of Classification Indices: Spek 9-12

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.558	0.576	0.424

Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.943	0.245
	A2	0.502	0.150
	A3	0.699	0.081
	P1	0.419	0.406
	P2	-	0.566

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.979	0.012	0.009	0.970
	A2/A3	0.974	0.012	0.014	0.964
	A3/P1	0.956	0.011	0.033	0.933
	P1/P2	0.646	0.354	0.000	0.683

6.4.4 Writing 9-12

Figure 6.4.4A

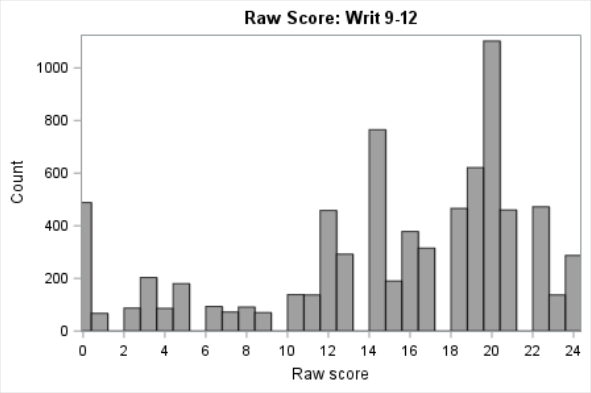


Table 6.4.4A

Raw Score Descriptive Statistics: Writ 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1908	0	24	14.49	6.76
10	1715	0	24	14.83	6.55
11	1500	0	24	15.13	6.71
12	2532	0	24	15.08	6.67
Total	7655	0	24	14.89	6.67

Figure 6.4.4B

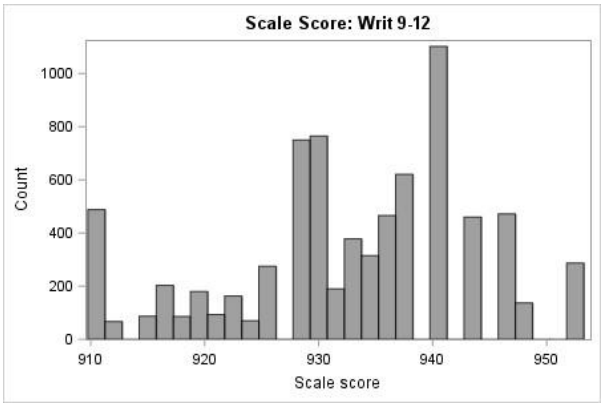


Table 6.4.4B

Scale Score Descriptive Statistics: Writ 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1908	910	953	932.27	10.73
10	1715	910	953	932.79	10.4
11	1500	910	953	933.41	10.83
12	2532	910	953	933.33	10.78
Total	7655	910	953	932.96	10.7

Figure 6.4.4C

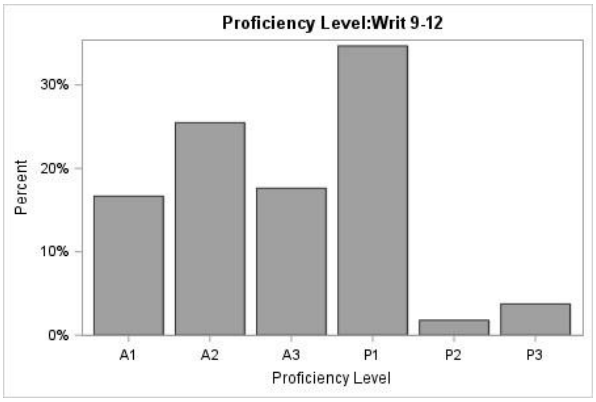


Table 6.4.4C

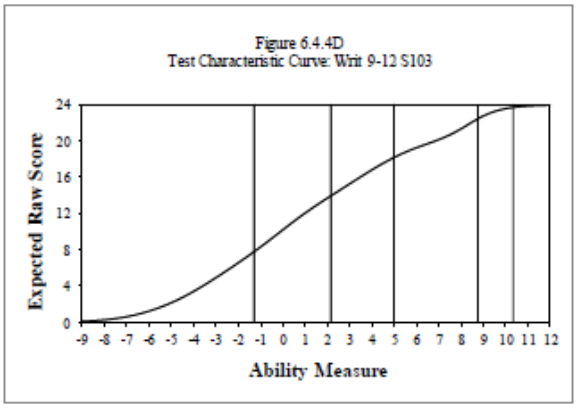
Proficiency Level Distribution: Writ 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	346	18.13%	278	16.21%	246	16.40%	406	16.03%	1276	16.67%
A2	500	26.21%	460	26.82%	351	23.40%	640	25.28%	1951	25.49%
A3	368	19.29%	294	17.14%	267	17.80%	420	16.59%	1349	17.62%
P1	599	31.39%	604	35.22%	541	36.07%	911	35.98%	2655	34.68%
P2	31	1.62%	33	1.92%	28	1.87%	45	1.78%	137	1.79%
P3	64	3.35%	46	2.68%	67	4.47%	110	4.34%	287	3.75%
Total	1908	100.00%	1715	100.00%	1500	100.00%	2532	100.00%	7655	100.00%

Table 6.4.4D

Equating Summary: Writ 9-12

No equating summary is presented because the Alternate ACCESS Series 601 was not equated. There is no change from the field test Series 100. Thus, the results from the original field test of the Alternate ACCESS were used to determine raw-to-scale score conversions. Technical details of the analysis of this process can be found in the Alternate ACCESS for ELLs™ Series 100 Development and Operational Field Test: Technical Report (2013).



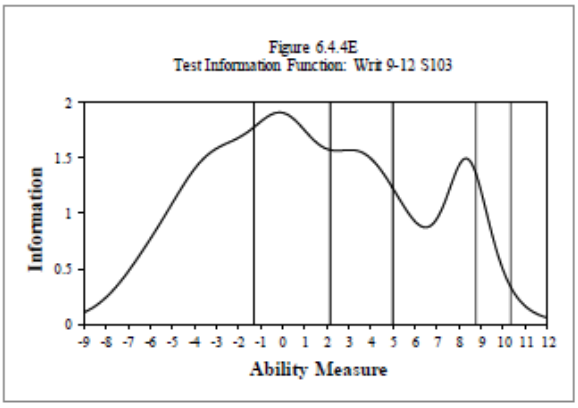


Table 6.4.4E

Reliability: Writ 9-12

No. of Students	No. of Items	Cronbach's Alpha	SEM
7655	10	0.9413	1.6169

Table 6.4.4F

Item Analysis Summary: Writ 9-12

Note: The contents of this table have been redacted in this version of the document.

Table 6.4.4G

Complete Item Analysis: Writ 9-12

Note: The contents of this table have been redacted in this version of the document.

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	6.87	6.87	A1	6.01	6.01	A1	6.27	6.27	A1	6.32	6.32
1	A1	1.21	8.07	A1	1.05	7.06	A1	0.6	6.87	A1	0.67	6.99
2	A1	1.47	9.54	A1	1.11	8.16	A1	1.33	8.2	A1	0.79	7.78
3	A1	2.36	11.9	A1	2.68	10.85	A1	2.93	11.13	A1	2.69	10.47
4	A1	1.52	13.42	A1	0.7	11.55	A1	0.73	11.87	A1	1.34	11.81
5	A1	2.83	16.25	A1	2.45	13.99	A1	2.27	14.13	A1	1.97	13.78
6	A1	0.79	17.03	A1	1.22	15.22	A1	1.4	15.53	A1	1.42	15.21
7	A1	1.1	18.13	A1	0.99	16.21	A1	0.87	16.4	A1	0.83	16.03
8	A2	1.47	19.6	A2	1.46	17.67	A2	0.8	17.2	A2	1.03	17.06
9	A2	1.05	20.65	A2	0.82	18.48	A2	1.13	18.33	A2	0.75	17.81
10	A2	1.36	22.01	A2	1.63	20.12	A2	1.8	20.13	A2	2.25	20.06
11	A2	2.31	24.32	A2	1.28	21.4	A2	1.13	21.27	A2	2.13	22.2
12	A2	5.14	29.45	A2	7.17	28.57	A2	6.07	27.33	A2	5.77	27.96
13	A2	3.93	33.39	A2	3.5	32.07	A2	3.33	30.67	A2	4.23	32.19
14	A2	10.95	44.34	A2	10.96	43.03	A2	9.13	39.8	A2	9.12	41.31
15	A3	2.78	47.12	A3	2.62	45.66	A3	2.47	42.27	A3	2.17	43.48
16	A3	4.4	51.52	A3	5.25	50.9	A3	5.27	47.53	A3	4.94	48.42
17	A3	5.19	56.71	A3	3.56	54.46	A3	4.47	52	A3	3.48	51.9
18	A3	6.92	63.63	A3	5.71	60.17	A3	5.6	57.6	A3	6	57.9
19	P1	7.7	71.33	P1	8.69	68.86	P1	8.2	65.8	P1	7.98	65.88
20	P1	12.37	83.7	P1	15.16	84.02	P1	15.07	80.87	P1	15.01	80.88
21	P1	6.45	90.15	P1	5.66	89.68	P1	5.67	86.53	P1	6.12	87.01
22	P1	4.87	95.02	P1	5.71	95.39	P1	7.13	93.67	P1	6.87	93.88
23	P2	1.62	96.65	P2	1.92	97.32	P2	1.87	95.53	P2	1.78	95.66
24	P3	3.35	100	P3	2.68	100	P3	4.47	100	P3	4.34	100

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

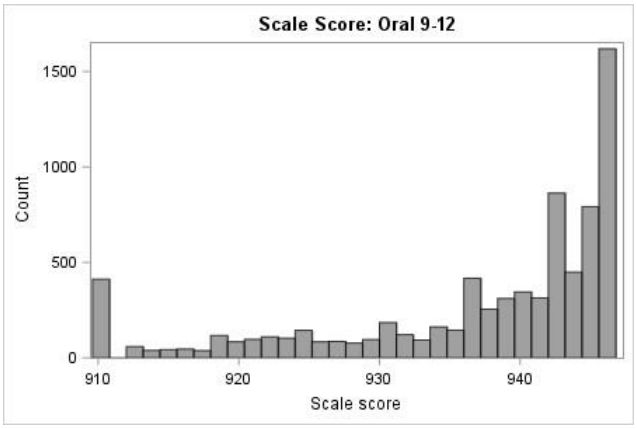


Table 6.4.5A

Scale Score Descriptive Statistics: Oral 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1921	910	946	936.3	10.8
10	1715	910	946	937.06	10.5
11	1511	910	946	937.31	10.36
12	2559	910	946	937.08	10.44
Total	7706	910	946	936.93	10.53

Figure 6.4.5B

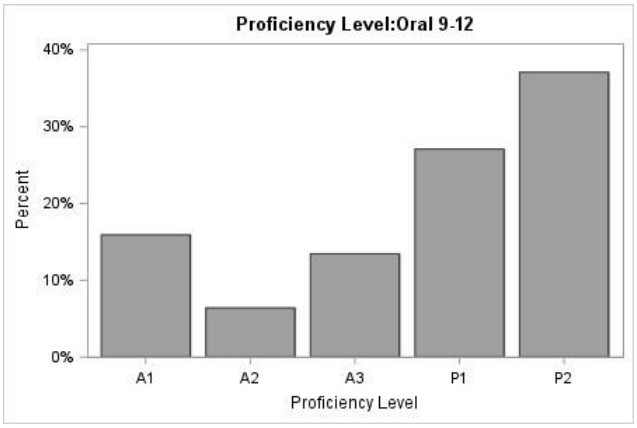


Table 6.4.5C

Proficiency Level Distribution: Oral 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	340	17.70%	262	15.28%	231	15.29%	394	15.40%	1227	15.92%
A2	135	7.03%	123	7.17%	86	5.69%	150	5.86%	494	6.41%
A3	256	13.33%	213	12.42%	213	14.10%	355	13.87%	1037	13.46%
P1	533	27.75%	452	26.36%	391	25.88%	712	27.82%	2088	27.10%
P2	657	34.20%	665	38.78%	590	39.05%	948	37.05%	2860	37.11%
Total	1921	100.00%	1715	100.00%	1511	100.00%	2559	100.00%	7706	100.00%

Table 6.4.5D

n/a

Figure 6.4.5D

n/a

Figure 6.4.5E

n/a

Table 6.4.5E

Reliability: Oral 9-12

Component	Weight	Variance	Reliability
Listening	0.5	105.7258	0.939
Speaking	0.5	154.2598	0.9667
Oral		110.8945	0.9739

*Variances from students who had results in all four domains

Table 6.4.5F
n/a

Table 6.4.5G
n/a

Table 6.4.5H
n/a

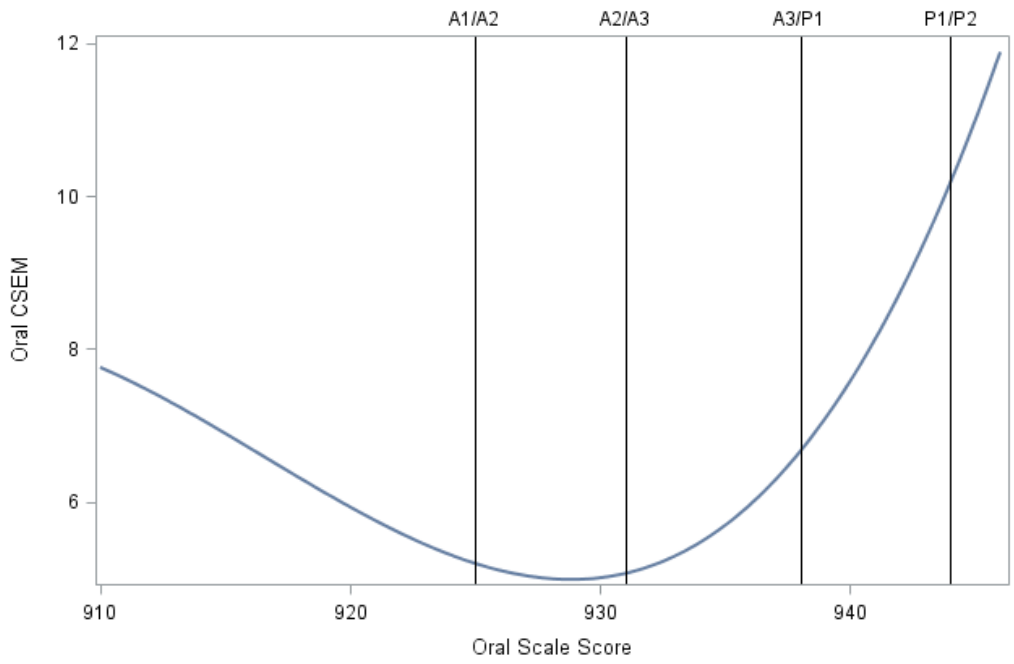
Table 6.4.5I
n/a

Table 6.4.5J

Accuracy and Consistency of Classification Indices: Oral 9-12

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

Figure 6.4.5F CSEM for Oral Composite 9-12



6.4.6 Literacy Composite 9-12

Figure 6.4.6A

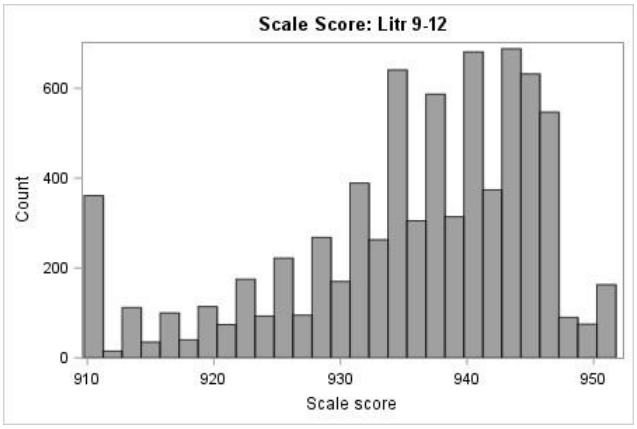


Table 6.4.6A

Scale Score Descriptive Statistics: Litr 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1898	910	951	934.82	10.19
10	1709	910	951	935.5	9.99
11	1493	910	951	935.92	10.08
12	2523	910	951	935.89	10.08
Total	7623	910	951	935.54	10.09

Figure 6.4.6B

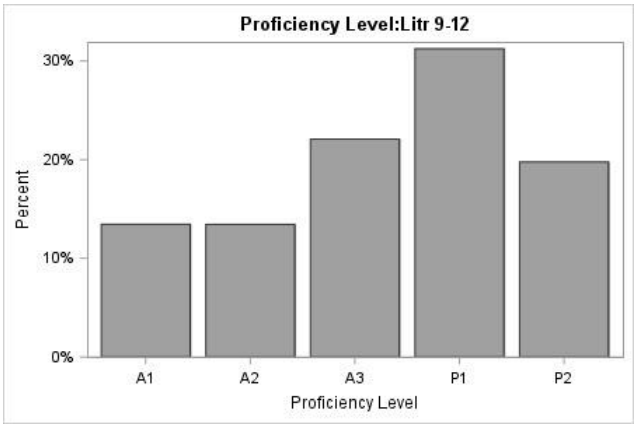


Table 6.4.6C

Proficiency Level Distribution: Litr 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	286	15.07%	227	13.28%	189	12.66%	324	12.84%	1026	13.46%
A2	270	14.23%	220	12.87%	193	12.93%	342	13.56%	1025	13.45%
A3	429	22.60%	395	23.11%	334	22.37%	525	20.81%	1683	22.08%
P1	605	31.88%	535	31.30%	459	30.74%	783	31.03%	2382	31.25%
P2	308	16.23%	332	19.43%	318	21.30%	549	21.76%	1507	19.77%
Total	1898	100.00%	1709	100.00%	1493	100.00%	2523	100.00%	7623	100.00%

Table 6.4.6D
n/a

Figure 6.4.6D
n/a

Figure 6.4.6E
n/a

Table 6.4.6E

Reliability: Litr 9-12

Component	Weight	Variance	Reliability
Reading	0.5	118.3076	0.9406
Writing	0.5	114.5028	0.9413
Literacy		101.8867	0.9663

*Variances from students who had results in all four domains

Table 6.4.6F
n/a

Table 6.4.6G
n/a

Table 6.4.6H
n/a

Table 6.4.6I
n/a

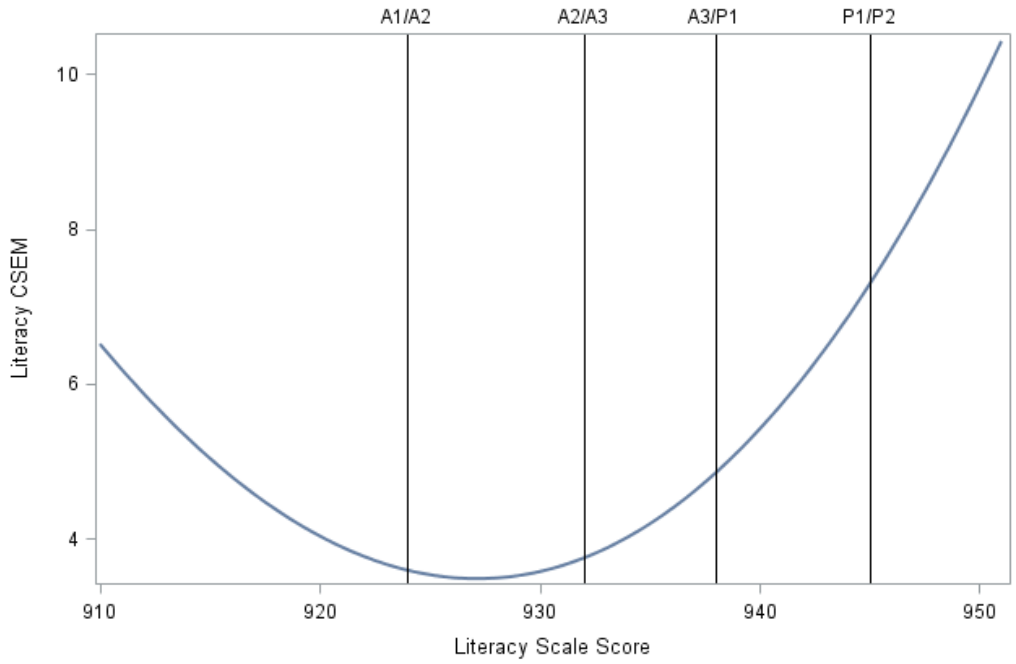
Table 6.4.6J

Accuracy and Consistency of Classification Indices: Litr 9-12

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.633	0.609	0.496
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.895	0.842
	A2	0.747	0.644
	A3	0.790	0.694
	P1	0.492	0.495
	P2	-	0.589

Indices at Cut Points		**Accuracy**			
	Cut Point	**Accuracy**	**False Positives**	**False Negatives**	**Consistency**
	A1/A2	0.976	0.014	0.010	0.965
	A2/A3	0.952	0.026	0.022	0.933
	A3/P1	0.943	0.015	0.042	0.921
	P1/P2	0.762	0.238	0.000	0.786

Figure 6.4.6F CSEM for Literacy Composite 9-12



6.4.7 Comprehension Composite 9-12

Figure 6.4.7A

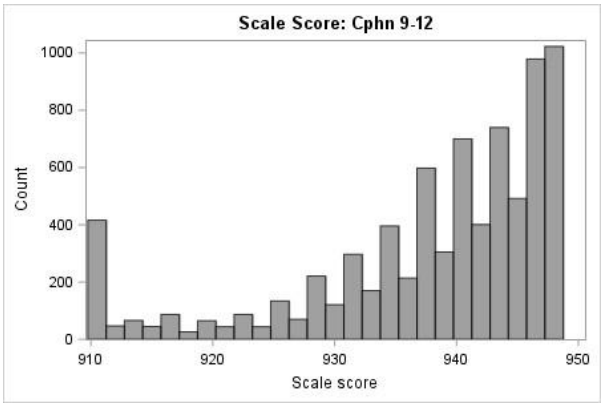


Table 6.4.7A

Scale Score Descriptive Statistics: Cphn 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1943	910	948	936.99	10.61
10	1738	910	948	937.79	10.35
11	1519	910	948	938.04	10.28
12	2588	910	948	937.95	10.39
Total	7788	910	948	937.69	10.42

Figure 6.4.7B

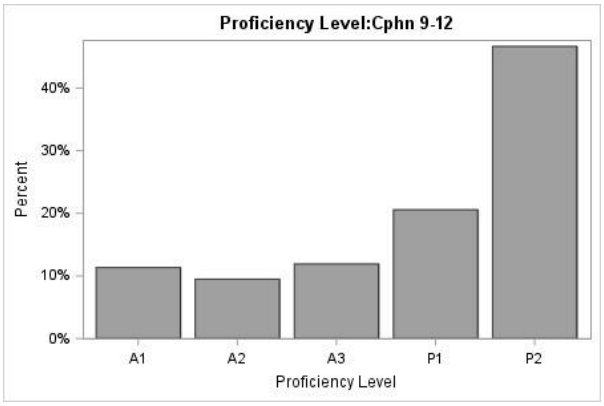


Table 6.4.7C

Proficiency Level Distribution: Cphn 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	239	12.30%	198	11.39%	158	10.40%	288	11.13%	883	11.34%
A2	215	11.07%	156	8.98%	134	8.82%	234	9.04%	739	9.49%
A3	241	12.40%	206	11.85%	185	12.18%	296	11.44%	928	11.92%
P1	397	20.43%	361	20.77%	311	20.47%	534	20.63%	1603	20.58%
P2	851	43.80%	817	47.01%	731	48.12%	1236	47.76%	3635	46.67%
Total	1943	100.00%	1738	100.00%	1519	100.00%	2588	100.00%	7788	100.00%

Table 6.4.7D
n/a

Figure 6.4.7D
n/a

Figure 6.4.7E
n/a

Table 6.4.7E

Reliability: Cphn 9-12

Component	Weight	Variance	Reliability
Listening	0.3	105.7258	0.939
Reading	0.7	118.3076	0.9406
Comprehension		108.6447	0.963

*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

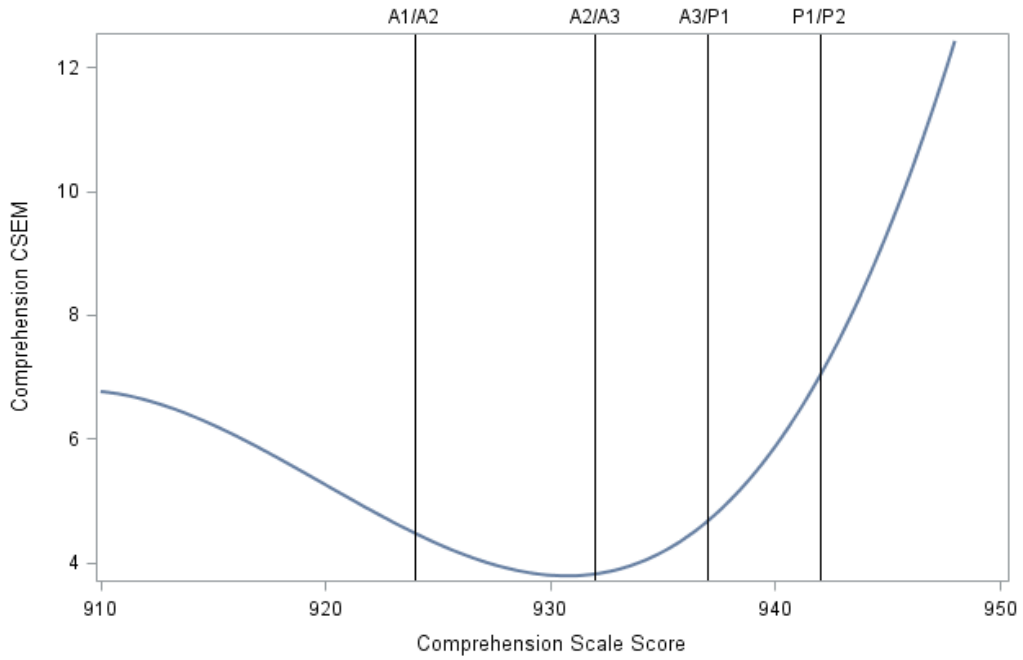
Table 6.4.7I
n/a

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

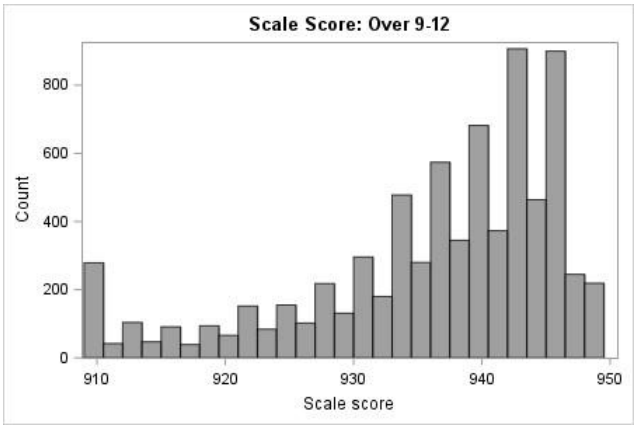


Table 6.4.8A

Scale Score Descriptive Statistics: Over 9-12

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	1878	910	949	935.19	10.04
10	1690	910	949	935.89	9.82
11	1480	910	949	936.22	9.83
12	2499	910	949	936.15	9.86
Total	7547	910	949	935.87	9.9

Figure 6.4.8B

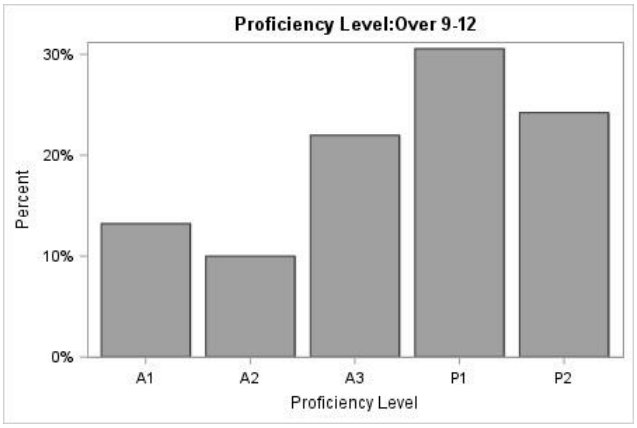


Table 6.4.8C

Proficiency Level Distribution: Over 9-12

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
A1	274	14.59%	221	13.08%	187	12.64%	316	12.65%	998	13.22%
A2	211	11.24%	169	10.00%	132	8.92%	243	9.72%	755	10.00%
A3	408	21.73%	372	22.01%	326	22.03%	553	22.13%	1659	21.98%
P1	589	31.36%	526	31.12%	449	30.34%	743	29.73%	2307	30.57%
P2	396	21.09%	402	23.79%	386	26.08%	644	25.77%	1828	24.22%
Total	1878	100.00%	1690	100.00%	1480	100.00%	2499	100.00%	7547	100.00%

Table 6.4.8D
n/a

Figure 6.4.8D
n/a

Figure 6.4.8E

n/a

Table 6.4.8E

Reliability: Over 9-12

Component	Weight	Variance	Reliability
Listening	0.15	105.7258	0.939
Reading	0.35	118.3076	0.9406
Speaking	0.15	154.2598	0.9667
Writing	0.35	114.5028	0.9413
Overall Composite		97.9669	0.9802

*Variances from students who had results in all four domains

Table 6.4.8F

n/a

Table 6.4.8G

n/a

Table 6.4.8H

n/a

Table 6.4.8I

n/a

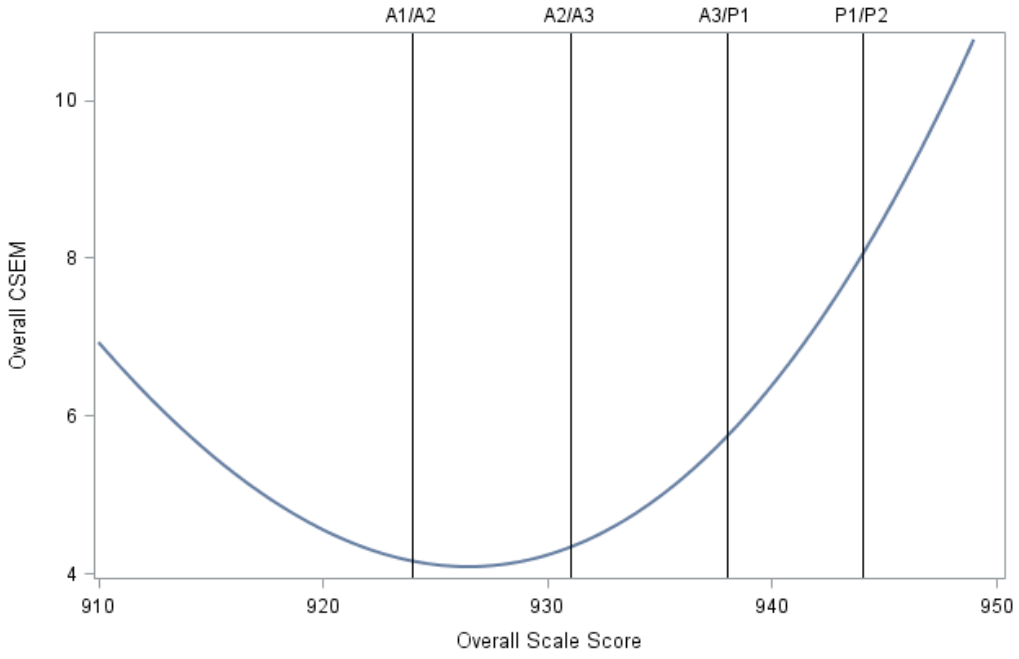
Table 6.4.8J

Accuracy and Consistency of Classification Indices: Over 9-12

Overall Indices	**Accuracy**	**Consistency**	**Kappa (k)**
	0.594	0.653	0.550
Conditional on Level	**Level**	**Accuracy**	**Consistency**
	A1	0.938	0.907
	A2	0.725	0.619
	A3	0.894	0.840
	P1	0.424	0.487
	P2	1.780	0.645

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

Figure 6.4.8F CSEM for Overall Composite 9-12



References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

WIDA Consortium. (2007). *English Language Proficiency Standards and Resource Guide, 2007 Edition, Pre-Kindergarten through Grade 12*. Madison, Wisconsin: Board of Regents of the University of Wisconsin System.

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.