World-Class Instructional Design and Assessment

Annual Technical Report for ACCESS for ELLs Online English Language Proficiency Test Series 403, 2018–2019 Administration

Annual Technical Report No. 15A

Prepared by:

Center for Applied Linguistics

Language Assessment Division
Psychometrics and Quantitative Research Team

April 2020

© 2020 Board of Regents of the University of Wisconsin System on behalf of the WIDA Consortium.

The WIDA ACCESS for ELLs Technical Advisory Committee

This report has been reviewed by the WIDA ACCESS for ELLs Technical Advisory Committee (TAC), which comprises the following members:

- Jamal Abedi, Ph.D., Professor, Graduate School of Education, University of California at Davis and a research partner at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST)
- Lyle Bachman, Ph.D., Professor Emeritus, Applied Linguistics, University of California, Los Angeles
- Gregory J. Cizek, Ph.D., Guy B. Phillips Distinguished Professor, Educational Measurement and Evaluation, University of North Carolina at Chapel Hill
- Claudia Flowers, Ph.D., Professor, Educational Leadership, University of North Carolina at Charlotte
- Akihito Kamata, Ph.D., Professor, Department of Education Policy and Leadership, Department of Psychology, Southern Methodist University
- Timothy Kurtz, Hanover High School, Hanover, New Hampshire
- Carol Myford, Ph.D., Professor Emerita, Educational Psychology, University of Illinois at Chicago

Executive Summary

This is the 15th annual technical report on the ACCESS for ELLs English Language Proficiency Test and the fourth report on the ACCESS for ELLs assessment as given in Online format.

This technical report is produced as a service to members and potential members of the WIDA Consortium. The technical information herein is intended for use by those who have technical knowledge of test construction and measurement procedures, as stated in *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014).

ACCESS for ELLs is intended to assess reliably and validly the English language development of English language learners (ELLs) in Grades K–12 according to the WIDA 2012 Amplification of the English Language Development Standards Kindergarten–Grade 12 (WIDA Consortium, 2012). Results on ACCESS for ELLs are used by WIDA Consortium states for monitoring the progress of students, for making decisions about exiting students from language support services, and for accountability. WIDA additionally provides screening instruments for initial identification purposes; however, decision processes on how these are incorporated into identification decisions are at individual states' discretion.

ACCESS for ELLs assesses students in the four domains of Listening, Reading, Writing, and Speaking, as required by federal law (Elementary and Secondary Education Act of 1965, amended 2015; §1111(b)(1)(F); §1111(b)(2)(G)) and provides composite scores as required by the same statute (§3121).

ACCESS for ELLs Online Series 403 was administered in school year 2018–2019 in 34 states, the Bureau of Indian Education, the District of Columbia, the Commonwealth of the Northern Marianas, and the U.S. Virgin Islands for a total of 38 state entities (henceforth "states").

The Series 403 Online data set included the results of 1,541,961 students. The largest grade was Grade 3 with 195,584 students, while the smallest was Grade 12 with 56,193 students. Of the participating WIDA states, the largest was Illinois with 167,716 students, while the smallest was the Bureau of Indian Education with 266 students.

ACCESS for ELLs Series 403 was offered in two administrative formats, an online format (Grades 1–12) and a paper format (Kindergarten–Grade 12). The current report (WIDA ACCESS Technical Report 15A) provides technical information pertaining to ACCESS for ELLs Series 403 Online. A second report (WIDA ACCESS Technical Report 15B) provides technical information for the ACCESS for ELLs Series 403 Paper assessment, including the Kindergarten assessment.

Part 1:
Purpose, Design, Implementation

Contents

1. Purpose and Design of ACCESS

1.1. Purpose 1-1
1.2. The WIDA Standards 1-1
1.3. The WIDA Proficiency Levels 1-2
1.4. Language Domains 1-4
1.5. Grade-Level Clusters 1-4
1.6. Tiers 1-4

2. Test Development

2.1. Test Design 2-1
  2.1.1. Listening 2-1
  2.1.2. Reading 2-2
  2.1.3. Writing 2-3
  2.1.4. Speaking 2-4
2.2. Test Development and Field Testing 2-5
  2.2.1. Listening 2-6
  2.2.2. Reading 2-7
  2.2.3. Writing 2-8
  2.2.4. Speaking 2-9
2.3. Item and Task Design 2-10
  2.3.1. Listening Items 2-10
  2.3.2. Reading Items 2-11
  2.3.3. Writing Tasks 2-12
  2.3.4. Speaking Tasks 2-12

3. Assessment Performance: The Implementation of ACCESS

3.1. Test Delivery 3-1
  3.1.1. Listening and Reading 3-1
  3.1.2. Writing 3-1
  3.1.3. Speaking 3-1
3.2. Scoring Procedures 3-2
  3.2.1. Multiple-Choice Scoring: Listening and Reading 3-2
  3.2.2. Scoring Performance-Based Tasks: Writing and Speaking 3-2
  3.2.3. Writing Scoring Scale 3-4
  3.2.4. Speaking Scoring Scale 3-5
3.3. Operational Administration 3-6
  3.3.1. Administering the Test Practice 3-6
  3.3.2. Listening Test Administration 3-7

3.3.3. Reading Test Administration 3-7
3.3.4. Writing Test Administration 3-8
3.3.5. Speaking Test Administration 3-8
3.3.6. Test Security 3-9
3.4. Procedures for Including Students with Disabilities 3-10
3.4.1. Support Provided to All ELLs 3-10
3.4.2. Support Provided to ELLs with IEP or 504 Plans 3-10

4. Summary of Score Reports

4.1. Individual Student Report 4-1
4.2. Other Reports 4-3

1. Purpose and Design of ACCESS

1.1. Purpose

The overarching purpose of ACCESS for ELLs is to assess the developing English language proficiency of English language learners (ELLs) in Grades K–12 in the United States as defined by the multistate WIDA Consortium, first in the English Language Proficiency Standards (Gottlieb, 2004; WIDA Consortium, 2007) and then in the amplified 2012 English Language Development (ELD) Standards (WIDA Consortium, 2012). The WIDA ELD Standards, which correspond to the academic language identified in state academic content standards, describe six levels of developing English language proficiency and form the core of the WIDA Consortium's approach to instructing and testing ELLs. ACCESS may thus be described as a standards-based English language proficiency test designed to measure the social and academic language proficiency of ELLs in English. It assesses social and instructional English as well as the academic language associated with language arts, mathematics, science, and social studies, within the school context, across the four language domains (Listening, Reading, Writing, and Speaking).

Other major purposes of ACCESS include

- Identifying the English language proficiency level of students with respect to the WIDA ELD Standards used in all member states of the WIDA Consortium
- Identifying students who have attained English language proficiency
- Assessing annual English language proficiency gains using a standards-based assessment instrument
- Providing districts with information that will help them to evaluate the effectiveness of their language instructional educational programs and determine staffing requirements
- Providing data for meeting federal and state statutory requirements with respect to student assessment
- Providing information that enhances instruction and learning in programs for English language learners

ACCESS for ELLs is offered in two formats: ACCESS Online, described in this report, and ACCESS Paper, described in a companion report.

1.2. The WIDA Standards

Five foundational WIDA ELD Standards inform the design, structure, and content of ACCESS for ELLs:

- *Standard 1*: ELLs communicate in English for **Social and Instructional** purposes within the school setting.

- *Standard 2*: ELLs communicate information, ideas, and concepts necessary for academic success in the content area of **Language Arts.**
- *Standard 3*: ELLs communicate information, ideas, and concepts necessary for academic success in the content area of **Mathematics.**
- *Standard 4*: ELLs communicate information, ideas, and concepts necessary for academic success in the content area of **Science.**
- *Standard 5*: ELLs communicate information, ideas, and concepts necessary for academic success in the content area of **Social Studies.**

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

- Social and Instructional Language: SIL
- Language of Language Arts: LoLA
- Language of Math: LoMA
- Language of Science: LoSC
- Language of Social Studies: LoSS

Every selected response item and every performance-based task on ACCESS for ELLs targets at least one of these five Standards. In the cases of some test items and tasks, the Standards are combined as follows:

- Integrated Social and Instructional Language (SIL), Language of Language Arts (LoLA), and Language of Social Studies (LoSS): IT
- Language of Math (LoMA) and Language of Science (LoSC): MS
- Language of Language Arts (LoLA) and Language of Social Studies (LoSS): LS

1.3. The WIDA Proficiency Levels

The WIDA ELD Standards describe the continuum of language development via five language proficiency levels (PLs) that are fully delineated in the WIDA ELD Standards document (WIDA Consortium, 2012), with scores indicating progression through each level. These levels are *Entering*, *Emerging*, *Developing*, *Expanding*, and *Bridging*. There is also a final stage known as *Reaching,* which is used to describe students who have progressed across the entire WIDA English language proficiency continuum; as this is the end of the continuum, scores do not indicate progression through this level. The proficiency levels are shown graphically in Figure 1.

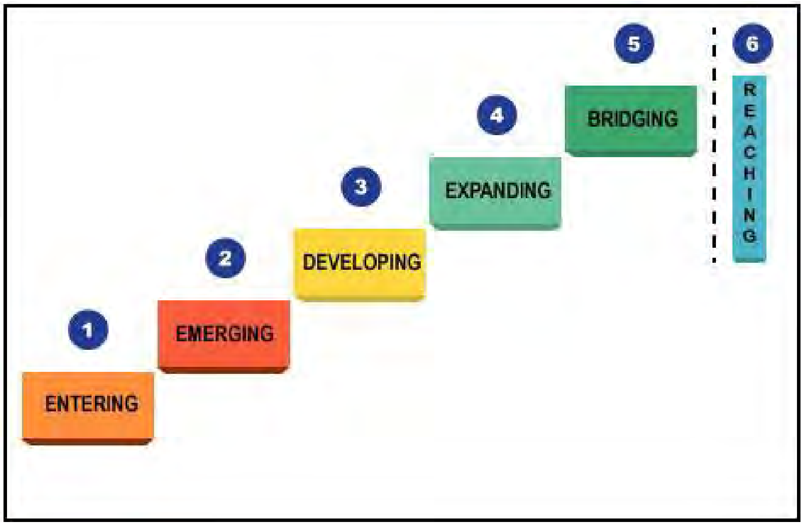


Figure 1. The language proficiency levels of the WIDA ELD Standards.

These language proficiency levels are 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. The performance definitions complement the **model performance indicators** (MPIs; see below) for each language proficiency level. The performance definitions are based on three criteria: (a) vocabulary usage at the word/phrase dimension; (b) language forms and conventions at the sentence dimension; and (c) linguistic complexity at the discourse dimension. Vocabulary usage refers to students' increasing comprehension and production of the technical language required for success in the academic content areas. Language forms and conventions refers to the increasing development of phonological, syntactic, and semantic understanding in receptive skills or control of usage in productive language skills. Linguistic complexity refers to students' demonstration of oral interaction or writing of increasing quantity and variety.

Second, the language proficiency levels of the WIDA ELD Standards are represented through connections to the accompanying MPIs. The MPIs describe the expectations for ELL students in each of the five Standards**,** by grade-level cluster, across the four language domains. That is, an MPI at each of the five language proficiency levels can be found within each combination of Standard, grade-level cluster, and language domain. *Reaching* (PL 6) represents the end of the continuum rather than another level of language proficiency. The sequence of these five MPIs together describes a logical progression and accumulation of skills on the path from the lowest level of English language proficiency to full English language proficiency for academic success. The grouping of five MPIs in logical progression is called a "strand."

ACCESS for ELLs is based on individual MPIs organized into strands within the WIDA ELD Standards. Each selected-response item or performance-based task on ACCESS for ELLs is carefully developed, reviewed, piloted, and field tested to ensure that it allows students to demonstrate accomplishment of the targeted MPI.

1.4. Language Domains

The WIDA ELD Standards describe developing English language proficiency for each of the four language domains: Listening, Reading, Writing, and Speaking. Thus, ACCESS for ELLs contains four sections, each assessing an individual language domain.

1.5. Grade-Level Clusters

The WIDA ELD Standards describe developing English language proficiency within six grade-level clusters. These are K, 1, 2–3, 4–5, 6–8, and 9–12. Test forms follow this grade-level clustering (note that the Kindergarten [K] form is not administered online and thus is not covered in this report).

1.6. Tiers

Tests must be at the appropriate difficulty level for individual test takers in order to be valid and reliable. While the grade-level cluster structure is a design feature intended to ensure that the language expectations are developmentally appropriate for children at different age ranges, it is also important to attend to students' differing levels of language proficiency within each grade-level cluster. As one might expect, test items and tasks that allow Entering (PL 1) or Emerging (PL 2) students to demonstrate accomplishment of the MPIs at their proficiency level will not allow Expanding (PL 4) or Bridging (PL 5) students to demonstrate the full extent of their language proficiency. Likewise, items and tasks that allow Expanding (PL 4) and Bridging (PL 5) students to demonstrate accomplishment of the MPIs at their level would be far too challenging for Entering (PL 1) or Emerging (PL 2) students. Items that are far too easy for test takers may be boring and lead to inattentiveness on the part of students; items that are far too difficult for test takers may be frustrating and discourage them from performing their best. But more importantly, items that are too easy or too hard for a student add very little to the accuracy or quality of the measurement of that student’s language proficiency.

ACCESS is designed so that test paths or forms are appropriate to the proficiency level of individual students across the wide range of proficiencies described in the WIDA ELD Standards. In the multi-stage adaptive test, in the domains of Listening and Reading, students are routed to folders which vary in difficulty, designated as A, B, or C level folders. Tier A folders are intended for students at beginning levels of English language proficiency (PLs 1-3), Tier B folders for students at intermediate levels (PLs 2-4), and Tier C folders for students at more advanced proficiency levels (PLs 3-5). In the domain of Writing, the test forms are designated as either Tier A, which includes tasks written to elicit language up to PL 3, or Tier B/C which includes tasks written to elicit language up to PL 4 or PL 5. In the domain of Speaking, test forms are designed so that students at very beginning levels of proficiency take a pre-A form, which is designed to elicit language at PL 1; students at early levels of proficiency take the Tier A form, with tasks designed to elicit language at PL1 and PL3; and more proficient students take the Tier B/C form, with tasks designed to elicit language at PL3 and PL5.

2. Test Development

2.1. Test Design

This section describes how ACCESS Online is assembled to ensure that the evidence collected is (a) sufficient to make the intended decisions, and (b) appropriate for the student's level of proficiency. In order to tailor the test closely to student ability levels while still including items and tasks that assess all of the Standards, adaptivity has been built into the test. The Listening and Reading tests both use a multistage adaptive test design. The Writing and Speaking tests are tiered, and placement into the tiers depends on performance on the Listening and Reading tests.

2.1.1. Listening

The Listening test uses a multistage adaptive design, as illustrated in **Error! Reference source not found.**. All students begin the Listening test with two entry folders (with three items each) at Stage 1 and Stage 2, both targeting Social and Instructional Language (see Section 1.2 for the WIDA ELD Standards). At that point, the student's ability is estimated based on performance on those six items, and that ability estimate is used to determine which of the three leveled Language of Language Arts folders in Stage 3 is administered next. Students whose ability estimate predicts a PL score of 5.0 or higher are routed into the folder at the highest level (C in **Error! Reference source not found.**); students whose ability estimate predicts a PL score of 2.5 or lower are routed into the folder at the lowest level (A in **Error! Reference source not found.**); all others are routed into the B folder. Throughout the test, a student's underlying measure of ability is re-estimated with the completion of each folder, and the level of the next folder to be administered is chosen accordingly, following the decision rules above. Thus, each student will trace a tailor-made path through the test according to ability level, but the order of the stages is invariant across students. In total, there are eight possible stages, but students whose ability estimate falls below PL 2.5 after the sixth stage end the test at this point. The intent of this design is to ensure coverage of the Standards while delivering a test that closely matches the student's PL, thus minimizing measurement error.

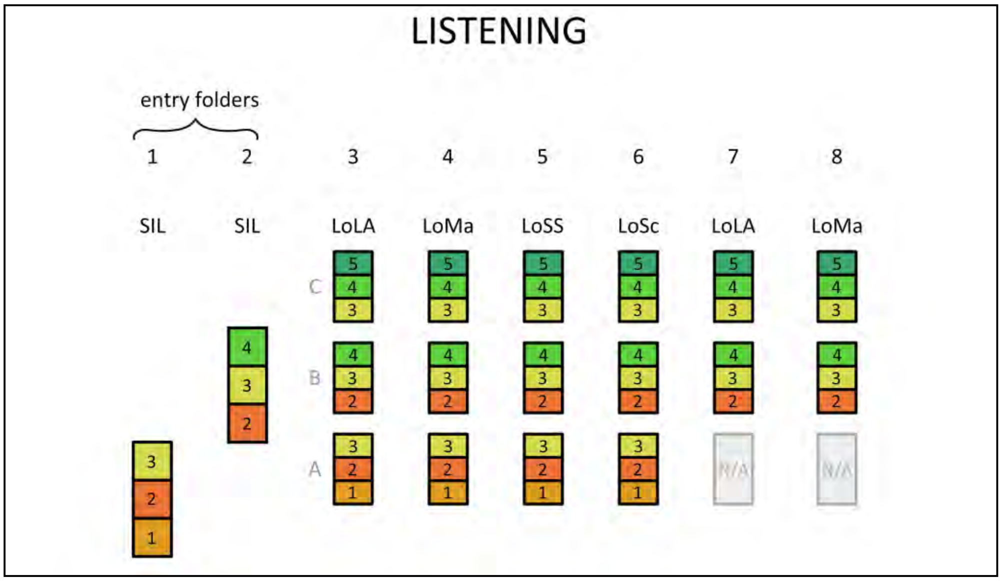

LISTENING
entry folders
1
2
3
4
5
6
7
8
SIL
SIL
LoLA
LoMa
LoSS
LoSc
LoLA
LoMa
C
B
A
5
4
3
2
1
N/A

Figure 2. Format of the Listening test.

2.1.2. Reading

Error! Reference source not found. shows the format of the Reading test. The format and adaptivity are similar to those of the Listening test, but the Reading test consists of 10 stages rather than eight. This reflects the greater weight given to Reading in calculating the composite scores (see Part 2 Chapter 3, "Analyses of Composite Scores"), as well as the view that literacy skills are paramount in developing academic language proficiency. The greater weight afforded to Reading and Writing resulted from a policy decision by the WIDA Board before the first operational administration of ACCESS. Students whose ability estimate falls below PL 2.5 after the eighth stage end the test at this point.

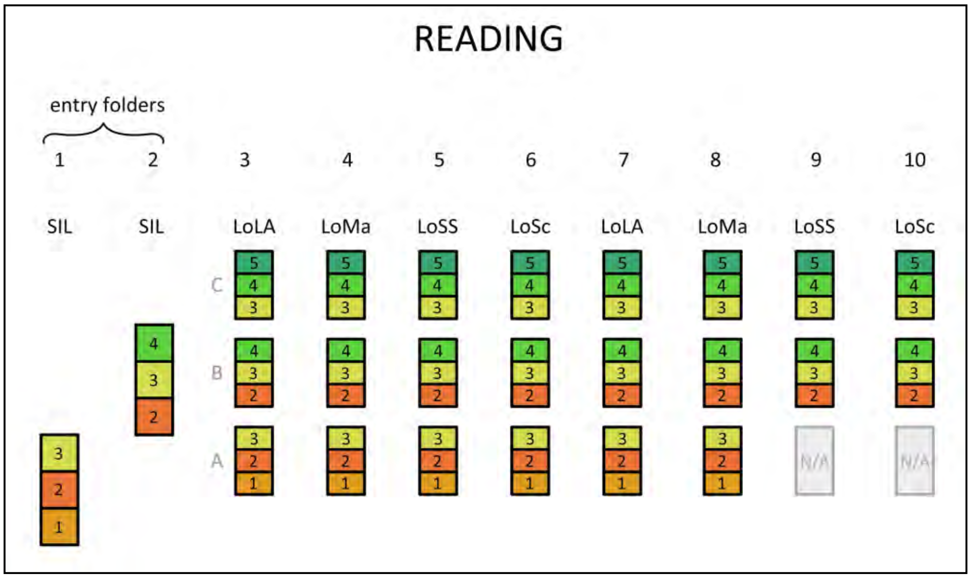

READING
entry folders
1
2
3
4
5
6
7
8
9
10
SIL
SIL
LoLA
LoMa
LoSS
LoSc
LoLA
LoMa
LoSS
LoSc
C
B
A
5
4
3
2
1
N/A

Figure 3. Format of the Reading test.

2.1.3. Writing

As shown in **Error! Reference source not found.**, the format of the Writing test is tiered; in the figure, the targeted proficiency level of each task is indicated by the large number in the colored rectangle, while the small numbers on the right of each rectangle indicate the possible proficiency levels that can be demonstrated on the task. Tier A consists of tasks written to elicit language up to PLs 3, while Tier B/C is designed to elicit language up to PL 4 or 5. However, for both tiers of the test, all tasks are scored using the entire breadth of the scoring scale. Therefore, students can theoretically score anywhere from 0 to 9 on any task (in terms of the raw scores in the scoring scale), although the design of some tasks naturally limits the possible scores. For example, Tier A tasks are not designed to elicit extended responses, so although the tasks are scored using the entire scale, these tasks do not elicit language above PL 4. Likewise, although Tier B/C tasks are designed to elicit extended discourse so that students can display proficiency at PL 4, PL 5, or even PL 6, some students will score throughout the proficiency range. With the exception of Grade 1 Tier A, both tiers consist of three tasks. Both tiers include tasks that target a single standard and tasks that integrate more than one WIDA Standard. For example, in the Tier A forms (except for Grade 1), one task integrates the Language of Math and the Language of Science. On the Tier B/C forms, one task integrates the Language of Math and the Language of Science, while another extended task integrates Social Instructional Language, the Language of Language Arts, and the Language of Social Studies. The ways in which the Standards are targeted by these tasks vary across grade levels and are spelled out in the generative item specifications.

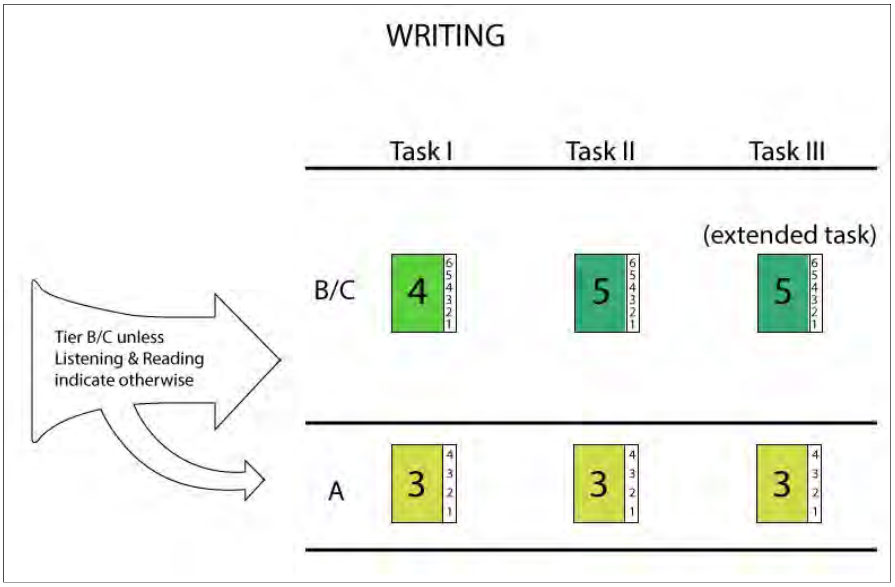


Figure 4. Format of the Writing test. Note: Grade 1 Tier A follows a different model and has four tasks targeting PLs 1, 2, and 3. Numbers inside the boxes represent the targeted proficiency level of the task;

the smaller numbers on the right edge of each box represent the range of proficiency levels that a task may elicit.

Placement into tiers on the Writing test depends on how students perform on the Listening and Reading tests, which receive computerized scores. To determine how to best place students into a tier, test data for all students who were administered the assessment in the 2015–2016 operational year (the first year of the ACCESS Online assessment) were analyzed to examine the relationship between how students perform on Listening and Reading and how they perform on Writing, using logistic regression analyses. This information was used to program an algorithm into the ACCESS Online test that is used by the computer to determine which tier of the Writing test to administer to each student. The purpose of the algorithm is to place students who are predicted to score above PL 3.0, based on their performances in Listening and Reading, into Tier B/C for Writing. All other students are placed into Tier A.

2.1.4. Speaking

Error! Reference source not found. shows the format of the Speaking test. The Speaking test includes tasks that target language elicitation at three PLs: 1, 3, or 5. The tasks are grouped into thematic folders, which are aligned to one or two of the WIDA Standards.

As shown in **Error! Reference source not found.**, the Speaking test includes three tiers: Tier Pre-A, Tier A, and Tier B/C. Tier Pre-A includes tasks that target elicitation of language at PL 1. Tier A includes tasks that target elicitation of language at PLs 1 and 3. Tier B/C includes tasks that target elicitation of language at PLs 3 and 5.

A thematic panel refers to the folders across all tiers within a grade-level cluster that relate to a particular WIDA ELD Standard. For example, the Tier B/C, Tier A, and Tier Pre-A folders that address Social and Instructional Language in a given grade cluster make up a single thematic panel. Within a thematic panel, tasks at PL 1 and PL 3 are the same across tiers. For example, within a Social and Instructional Language panel, the same PL 3 task appears on both the Tier A and the Tier B/C forms of the test.

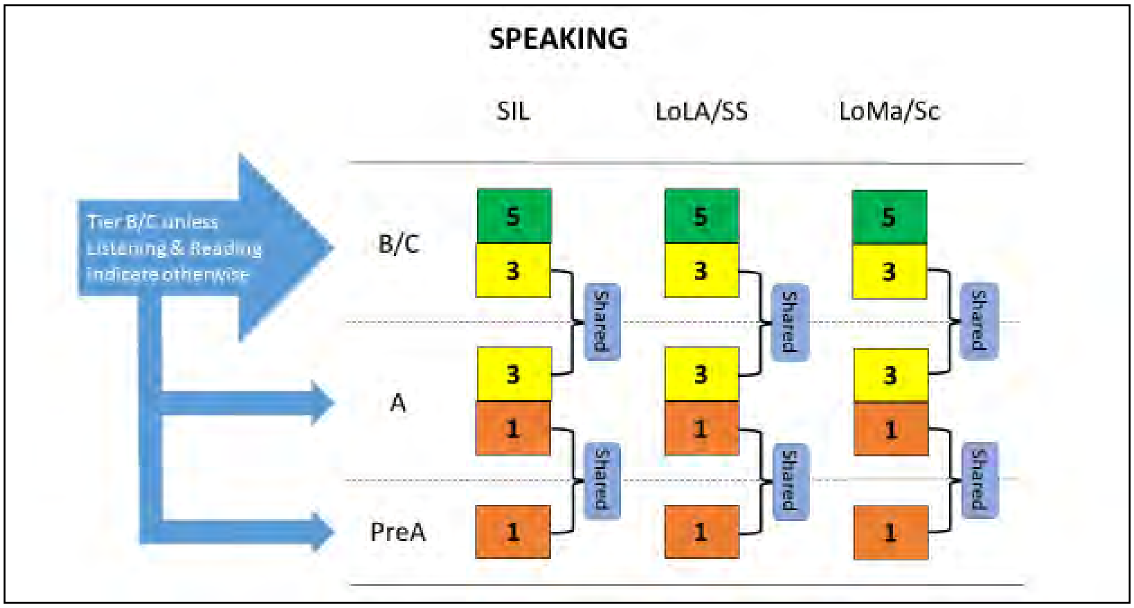


Figure 5. Format of the Speaking test.

As with Writing, placement into the three tiers on the Speaking test shown in **Error! Reference source not found.** depends on performance on the Listening and Reading tests.

Unlike Writing, the Speaking test has one additional tier, Tier Pre-A. Students are placed into Tier Pre-A when their scores on Listening and Reading fall below chance performance. The Speaking Pre-A tier is designed to meet the needs of students in the very early stages of English language development. As noted above, these tasks are targeted to the P1 level and are scored using a modified version of the full Speaking rating scale.

Placement of students into Tiers A and B/C in Writing is analogous to tier placement for Speaking. Test data for all students who were administered the assessment in the 2015–2016 operational year (the first year of the ACCESS Online assessment) were analyzed to examine the relationship between students' performance on Listening and Reading and performance on Speaking, using logistic regression analyses. This information was used to program an algorithm into the ACCESS 2.0 Online test to determine which tier of the Speaking test is administered to each student. The purpose of the algorithm is to place students who are predicted to score above PL 3.0, based on their performances in Listening and Reading, into Tier B/C for Speaking, and to place all other students into Tier A (except for those students, as noted above, who are routed into Tier Pre-A).

2.2. Test Development and Field Testing

The ACCESS item development process spans approximately three years, beginning with the development of the refreshment plan and the updating of item specifications. Trained item writers work from these specifications to draft items within a thematic folder. After initial development, folders are screened at CAL, and those that are approved for further development undergo a rigorous process of internal development and review, including reviews by standards experts and extensive fact checking. During this phase, images and other ancillary materials, such as scripts and directions, are produced.

After items are internally refined, they are reviewed by two panels: a content review panel and a bias and sensitivity review panel. The panels consist of specially trained educators with culturally and linguistically diverse backgrounds from WIDA Consortium states. Items are submitted to the content review panel to ensure that the content is accessible and relevant to students in the targeted grade-level cluster and at the targeted proficiency level and that each item or task matches the MPI from the WIDA ELD Standards that it is intended to assess. The bias and sensitivity review panel ensures that test items are free of material that (1) might favor any subgroup of students over another on the basis on gender, race/ethnicity, home language, religion, culture, region, or socioeconomic status, and (2) might be upsetting to students. Bias and sensitivity panelists are educators with culturally and linguistically diverse backgrounds who have experience interacting with English learners from a range of cultural, regional, religious, linguistic, ethnic, and socioeconomic backgrounds. Based on their recommendations, the items are revised as necessary.

Items that reach this point then go through field testing processes, described by domain below.

2.2.1. Listening

Listening items developed for Series 403 were field tested as embedded items during the operational administration of Series 402. The embedded field test items included innovative item formats, including hot spot items, where the student clicks on an area of the screen, and drag-and-drop items, where the student drags an image/text to a specified screen area to respond.

For Series 403, a total of 120 Listening items (40 folders) were field tested, across all five grade-level clusters, embedded into the Series 402 operational assessment. Each student received one Listening field test folder embedded into the operational test. Field test folders are targeted to refresh a specific operational folder on the test, and field test folder specifications include the stage, standard, and tier pool target (A, B, or C) of the folder. Students are administered the embedded field test folder at the stage targeted for refreshment, with administration randomized so that half of the students see the field test folder before the corresponding operational folder, and half see the operational folder before the field test folder. Field test folders are administered to those students who are routed to take the operational folder that is either at the same tier or adjacent to the tier that the field test folder targets. When field test samples are drawn, the sample includes 50% of students at the tier targeted by the field test folder and 50% at adjacent tiers (if there are adjacent tiers both above and below, 25% from each). In cases where the folder to be field tested is to be placed in one of the entry stages, students who receive that field test folder will receive it directly after the pair of operational entry folders. Entry folders do not have a tier pool designation. Field test sample targets in Listening are set at a minimum of 3,000 responses per folder.

After field test data are drawn, folders of items are analyzed for their psychometric properties, and those that meet established psychometric standards are eligible for selection in the next year's operational test.

Table 1 provides numbers of continuing and new items per grade-level cluster. For further detail on item statistics, including a summary of the number of items used as anchors across years, see Part 2 of this report, Sections 2.1 and 2.7.

Table 1
Number of New and Continuing Items on ACCESS Online Series 403 Listening, by Grade-Level Cluster

Grade-level cluster	Number of new items	Number of continuing items	Total number of items
1	12	42	54
2–3	9	45	54
4–5	12	42	54
6–8	12	42	54
9–12	9	45	54

2.2.2. Reading

For Series 403, a total of 120 Reading items (40 folders) were field tested, across all five grade-level clusters, embedded into the Series 402 operational assessment. The embedded Reading field test is administered in the same way as the embedded Listening field test. Each student receives one Reading field test folder embedded into the operational test. Field test folders are targeted to refresh a specific operational folder on the test, and field test folder specifications include the stage, standard, and tier pool target (A, B, or C) of the folder. Students are administered the embedded field test folder at the stage targeted for refreshment, with administration randomized so that half of the students see the field test folder before the corresponding operational folder, and half see the operational folder before the field test folder. Field test folders are administered to those students who are routed to take the operational folder that is either at the same tier or adjacent to the tier that the field test folder targets. When field test samples are drawn, the sample includes 50% of students at the tier targeted by the field test folder and 50% at adjacent tiers (if there are adjacent tiers both above and below, 25% from each). In cases where the folder to be field tested is to be placed in one of the entry stages, students who receive that field test folder will receive it directly after the pair of operational entry folders. Entry folders do not have a tier pool designation. Field test sample targets in Reading are set at a minimum of 3,000 responses per folder.

After field test data are drawn, folders of items are analyzed for their psychometric properties, and those that meet established psychometric standards are eligible for selection in the next year's operational test.

Table 2 provides numbers of continuing and new items, per grade-level cluster. For further detail on item statistics, including a summary of the number of items used as anchors across years, see Part 2 of this report, Sections 2.1 and 2.7.

Table 2
Number of New and Continuing Items on ACCESS Online Series 403 Reading, by Grade-Level Cluster

Grade-level cluster	Number of new items	Number of continuing items	Total number of items
1	12	60	72
2–3	9	63	72
4–5	6	66	72
6–8	12	60	72
9–12	12	60	72

2.2.3. Writing

The development of Writing tasks is similar to that of Listening and Reading items. Writing tasks, however, do not currently undergo large-scale field testing. Instead, after external bias, sensitivity, and content reviews, they are subject to small-scale tryouts, coordinated by CAL staff and conducted by teachers. In these tryouts, candidate folders for Grades 4–12 are administered to students using the online interface, with all students producing a handwritten response; as noted above, students in Grades 1–3 complete the Writing assessment with a traditional paper-and-pencil administration. Student responses, as well as educator observations and comments, inform further revisions to the folders.

Then, a small-scale stand-alone field test of Writing folders is conducted. For Series 403, a total of 18 Writing tasks were field tested. A sample of 500 students per task was targeted. The field test is administered under standard testing conditions. The field test uses the online interface with keyboarded responses for Grades 4–12 and paper booklets with handwritten responses for Grades 1–3. For the Writing field test, DRC raters score the field test samples. A 20% read behind is instituted as a quality control measure, with the first score as the score of record.

Quantitative and qualitative analyses of the collected responses are conducted. The main purposes of this small-scale field testing are (a) to confirm that the tasks are working as intended, (b) to identify anchor samples for rater training, and (c) to inform the rating of the tasks when they become operational. Note that for the stand-alone Series 403 Writing field test, the sample target was not met for all clusters and tiers. Despite not meeting the sample targets, there were sufficient responses to conduct qualitative analyses, review raw score distributions, and provide evidence for the suitability of tasks for operational testing.

Series 403 incorporates continuing items from Series 402 as well as items newly developed and field tested for Series 403.

Table 3 provides numbers of continuing and new items, per grade-level cluster. For further detail on item statistics, including a summary of the number of items used as anchors across years, see Part 2 of this report, Sections 2.1 and 2.7.

Table 3
Number of New and Continuing Items on ACCESS Online Series 403 Writing, by Grade-Level Cluster

Grade-level cluster	Tier	Number of new items	Number of continuing items	Total number of items
1	A	0	4	4
	B/C	2	1	3
2–3	A	2	1	3
	B/C	2	1	3
4–5	A	2	1	3
	B/C	2	1	3
6–8	A	2	1	3
	B/C	2	1	3
9–12	A	2	1	3
	B/C	2	1	3

2.2.4. Speaking

The development of Speaking tasks is similar to that of Writing tasks, but, as with Listening and Reading, all Speaking tasks undergo large-scale field testing. Speaking tasks undergo both quantitative and qualitative analyses following the field test to determine their appropriateness for inclusion in the following year’s operational test.

Speaking tasks in Series 403 include new tasks for Series 403 and continuing tasks from Series 402.

All Tier A and B/C students are administered a Speaking field test folder appended to their operational Speaking assessment. Tier Pre-A is not included in the field test. A total of 60 tasks (20 panels) were field tested for Series 403, with a target sample size of 500 students per folder. Responses were scored by DRC trained raters with a 20% read-behind as a quality control measure, with the first score as the score of record.

For ACCESS Online, folders are designed to target one or two proficiency levels. Tier Pre-A folders include one task that targets only PL 1. Tier A folders include tasks that target PLs 1 and 3, and Tier B/C folders include tasks that target PLs 3 and 5. Students are routed into a tier based on their performance on the Reading and Listening sections of the test. The content is presented entirely on the computer, and the responses are recorded by the test engine and are transmitted to

DRC for scoring. Students receive a Speaking field test folder in the tier that corresponds to their operational tier.

Table 4 provides numbers of continuing and new tasks, per grade-level cluster. For further detail on item statistics, including a summary of the number of tasks used as anchors across years, see Part 2 of this report, Sections 2.1 and 2.7.

Table 4
Number of New and Continuing Tasks on ACCESS Online Series 403 Speaking, by Grade-Level Cluster

Grade-level cluster	Tier	Number of new tasks	Number of continuing tasks	Total number of tasks
1	Pre-A	2	1	3
1	A	4	2	6
1	B/C	4	2	6
2–3	Pre-A	2	1	3
2–3	A	4	2	6
2–3	B/C	4	2	6
4–5	Pre-A	2	1	3
4–5	A	4	2	6
4–5	B/C	4	2	6
6–8	Pre-A	2	1	3
6–8	A	4	2	6
6–8	B/C	4	2	6
9–12	Pre-A	2	1	3
9–12	A	4	2	6
9–12	B/C	4	2	6

2.3. *Item and Task Design*

This section describes how items and tasks are designed in order to collect the necessary evidence required for the purposes of the assessment. Items and tasks are discussed by language domain.

2.3.1. Listening Items

All Listening items include a prerecorded stimulus passage and question stem. Listening items are selected-response items, with one key and two distractors as answer choices. Answer choices are primarily illustrations; for Grades 2–12, items that test listening proficiency at PLs 3–5 may consist of short written text response options that are written to be about two PLs lower than the targeted PL of the Listening item. Most items on the operational Listening assessment are traditional multiple choice, though some operational items and some items embedded for field-testing purposes may involve enhanced item presentations, including hot spot items, where the student clicks on an area of the screen, and drag-and-drop items, where the student drags an image/text to a specified screen area to respond.

Each item on the Listening test is written to target the language of one of the five WIDA ELD Standards and to test a student's ability to process language at one of the five fully delineated proficiency levels. *Folders* group together three test items that are written around a common theme, with each item targeting a progressively higher proficiency level.

- Tier A folders are constructed to target PLs 1 through 3.
- Tier B folders are constructed to target PLs 2 through 4.
- Tier C folders are constructed to target PLs 3 through 5.

In ACCESS Online Listening, students take a multistage adaptive test form which routes students to Tier A, B, or C folders as appropriate to their ability level.

Listening items are developed so that each item appears on its own screen, with associated graphic support. Scripts containing the item orientation, stimulus, and question stem are audio recorded with professional voice actors and produced by a professional recording studio. Audio playback of test item content is automatic when students advance to the next screen. Listening test content is played one time for students unless the student has a predetermined accommodation allowing for a single repetition of the item stimulus and question stem.

2.3.2. Reading Items

Reading items are similar in format to Listening items. The stimulus for Reading items is written text, and answer choices are also primarily written text, though for Grades 1–12 response options for items targeting PLs 1 and 2 may be illustrations rather than text. As with Listening items, Reading items are grouped into thematic folders of three test items each.

- Tier A folders are constructed to target PLs 1 through 3.
- Tier B folders are constructed to target PLs 2 through 4.
- Tier C folders are constructed to target PLs 3 through 5.

In ACCESS Online Reading, students take a multistage adaptive test form which routes them to Tier A, B, or C folders as appropriate to their ability level.

Most items on the operational Reading assessment are traditional multiple choice, though some operational items and some items embedded for field-testing purposes involve enhanced item presentations, including hot spot and drag-and-drop items, where the student either clicks on an area of the screen or drags an image/text to a specified screen area to respond. Items have one key and either two or three distractors, depending upon grade-level cluster and targeted proficiency level. For Grades 1 and 2–3, all items have a key and two distractors. For Grades 4–5, 6–8, and 9–12, items targeting PLs 1 and 2 have a key and two distractors, and items targeting PLs 3, 4, and 5 have a key and three distractors.

2.3.3. Writing Tasks

Writing tasks are designed to elicit language corresponding to one or more of the WIDA ELD Standards. Tasks appearing on the Tier A test form are designed to give students the opportunity to produce writing samples that fulfill linguistic expectations up to PL 3. As described in Section 2.1.3. above, these tasks are scored using the entire breadth of the scoring scale, therefore students may achieve proficiency levels higher than PL 3, although the tasks are not designed to elicit extended responses, so the scores are limited by task design. Tasks appearing on the Tier B/C form are designed to give students the opportunity to produce writing samples that fulfill linguistic expectations up to PL 4 or 5. Again, although these tasks are designed to elicit extended responses, they are scored on the entire breadth of the scoring scale, so students' actual performances may extend above or below the PL 4-5 range.

For students in Grades 1–3, the test is not administered via computer. For students in these grades, the test administrator reads from a script and the students respond in a printed test booklet.

For students in Grades 4–12, writing prompts appear on the computer screen. In the spirit of providing maximal support and making every provision to ensure that students are given the opportunity to demonstrate the full extent of their English language proficiency, modeling is sometimes used to make task expectations as clear as possible to students. For example, the first of a series of questions may already be partially completed, or a sentence starter may be provided.

Students in Grades 4–5 provide either handwritten or keyboarded responses, with the default response mode determined in advance at the state or district level. For students in Grades 6–12, keyboarding is the default response mode, with a handwriting option offered as an accommodation.

2.3.4. Speaking Tasks

Stimuli on the Speaking test include graphics, audio, and text. All stimuli are presented by a virtual test administrator (VTA). The VTA serves as a narrator who guides students through the test and acts as a virtual interlocutor. The VTA is introduced to students during the test directions in order to establish the testing context.

Task modeling is an essential component of the Speaking test design. In addition to the VTA, students are introduced to a virtual model student during the test directions. Prior to responding to each task, test takers first listen to the model student respond to a parallel task. The purpose of the model is to demonstrate task expectations to both test takers and to DRC raters who score all Speaking task responses.

Students navigate through the Speaking test independently and at their own pace. They must listen to all audio on a screen before the test allows them to advance to the next screen. The amount and complexity of task input varies by grade-level cluster and task level. The purpose of the input is to provide academic content for students to draw on in their responses.

Figure 6 shows the generic screen layout of the Speaking test.

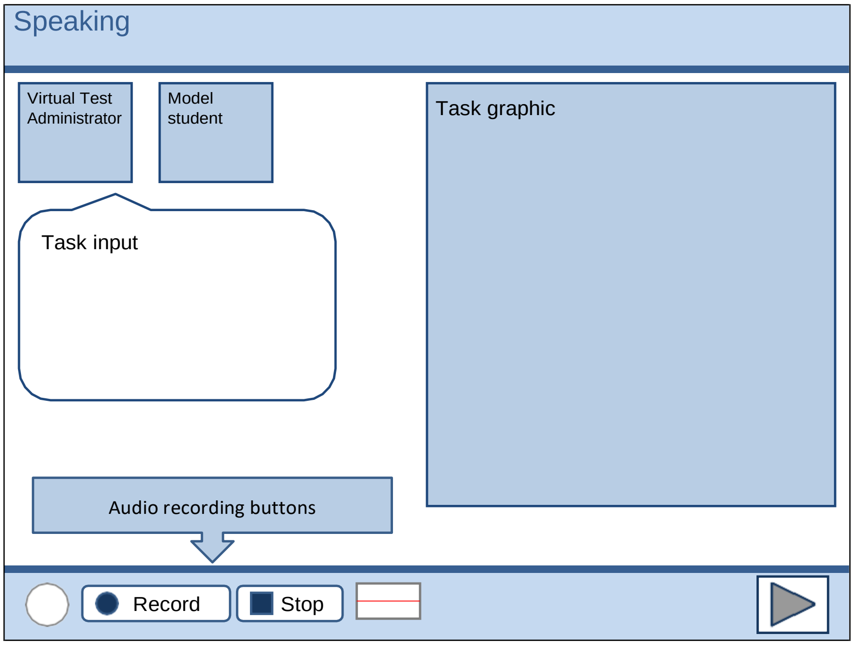


Figure 6. Visualization of the Speaking test screen layout.

Both the VTA and the model student are represented within the testing interface by static images. They are portrayed wearing computer headsets with microphones to reflect the actual testing scenario. Test input and stimuli are presented both aurally and in speech bubbles on the screen. Students respond orally to the tasks, with their responses recorded and transmitted to DRC for later scoring.

All Speaking tasks for a given grade cluster and WIDA Standard are designed in terms of *panels*; a panel is a thematically related set of three tasks, targeting the elicitation of PL 1, PL 3, and PL 5 language. When the tasks are field tested, the panels are split out into folders, with each folder containing one or two tasks. Tier Pre-A folders contain a single task targeting PL 1; Tier A folders contain two tasks targeting PL 1 and PL 3; and Tier C folders contain two tasks targeting PLs 3 and 5. For a given pair of Tier A and Tier C folders based on a single panel, the PL 3 task is identical in both folders.

3. Assessment Performance: The Implementation of ACCESS

3.1. Test Delivery

ACCESS Online is administered between December and April of the academic year, with testing windows determined at the state level. The Reading and Listening tests are administered first (in either order), followed by Writing and Speaking (in either order). The test may be administered in several sessions within a single day or over a series of days.

3.1.1. Listening and Reading

Listening and Reading are the first domains assessed. Students may take these in either order. Students sit at individual computer monitors and are administered the Listening and Reading tests online. They are issued headsets which are used to listen to directions for the Listening and Reading tests, as well as to the Listening items. Students use the computer interface to select or record their answers.

3.1.2. Writing

Writing tasks are delivered on paper to students in Grades 1–3. All students in Grades 1–3 handwrite a response.

Writing tasks are delivered online to students in Grades 4–12. A student may provide handwritten or keyboarded responses, with the choice depending on a combination of local, state, and consortium-wide policies, as follows:

- Grades 4–5: A decision is made at the local or state level as to whether handwriting or keyboarding is the default response mode. In districts where keyboarding is the default, the option exists to use handwriting as an accommodation.
- Grades 6–12: Keyboarding is the default, with the option to use handwriting as an accommodation.

3.1.3. Speaking

Speaking tasks are delivered online. Students listen to prompts via headsets that are equipped with microphones to capture their responses. Extensive support is provided to the student through illustrations and multimodal (text and audio) input designed to provide sufficient content for the response, as well as a model student response that is intended to provide guidance regarding the level of linguistic complexity required to respond adequately (see Section 2.1.4).

3.2. Scoring Procedures

3.2.1. Multiple-Choice Scoring: Listening and Reading

Listening and Reading items are scored dichotomously, as correct or incorrect. Scale scores for each domain are calculated based on the items that are administered to the test taker and the set of those items that the student answers correctly. For details on how scale scores for Listening and Reading are calculated, see Part 2, Chapter 2, "Analysis of Domains."

3.2.2. Scoring Performance-Based Tasks: Writing and Speaking

Performance-based tasks in the domains of Writing and Speaking are scored by trained raters. According to documentation from DRC, raters are well-educated professionals, with at least a 4-year college degree in a relevant field and a demonstrated writing ability. Prior to scoring live student responses, the raters undergo thorough training and qualifying. Training is task-specific in order to ensure that raters understand the nuances of each unique Writing or Speaking task. Team leaders, who are selected based on prior performance as raters and for their leadership skills, are assigned to small groups of raters; there are typically ten raters per team. The team leaders are responsible for monitoring the performance of their team members and providing ongoing feedback to support accurate scoring. Scoring directors are promoted from within DRC and earn their positions by demonstrating quality work as raters and as team leaders on previous projects. Scoring directors are responsible for a specific set of tasks within a single domain. The scoring directors train and oversee the teams of raters assigned to these tasks. What follows are general scoring procedures utilized by DRC.

Rater Training and Qualifying

- Raters are seated at stations and are assigned unique ID numbers and passwords.
- The scoring director provides detailed directions for use of DRC's computerized scoring system.
- The scoring director trains the raters using task-specific anchor sets and training sets.
- Raters must demonstrate scoring proficiency by scoring at least 70% agreement on a qualifying set before scoring live responses.
- Once raters are qualified, they are further trained for their grade-level cluster on the specific tasks for which they will rate responses.
- Once raters have trained, qualified, and begun live scoring, DRC uses calibration sets (of which there are two types, recalibration sets and validation sets, which are explained below) to keep the raters calibrated on the actual tasks they are scoring.

Calculating Score Agreement for Score Monitoring

- For Writing, agreement is defined as two adjacent scores. (See Section 3.2.3 for a description of the Writing Scoring Scale.) For example, using the Writing Scoring Scale,

scores of 2 and 2+ would be considered agreement, as would scores of 2 and 2 or scores of 2+ and 3. Scores of 2 and 3 on the Writing Scoring Scale would be considered adjacent, and scores of 2 and 3+ would be considered nonadjacent.
- For Speaking, agreement is defined as two scores that are exactly the same. (See Section 3.2.4 for a description of the Speaking Scoring Scale.)

Routing Responses to Ensure "Blind" Second Ratings

- The DRC scoring system ensures that responses are routed to qualified raters until the prescribed number of ratings is performed for all responses.
- Raters do not know if they are the first or second rater.

Monitoring Scoring (Quality Control)

- Ongoing quality control checks and procedures help monitor and maintain the quality of the scoring sessions. At least 20% of the responses are independently scored by two raters for the purpose of monitoring interrater reliability. DRC monitors these data daily.
- Responses can be retrieved on demand (e.g., specific grade-level clusters, specific students) should the need arise during or after the scoring process.
- If needed, responses can be rescored based on task- or response-level information, such as task number, date, score value assigned, or rater ID.
- For Writing, DRC uses both recalibration and validation sets. For each of the first 5 days that raters score a task, they take one recalibration set of five responses per task. After the raters take the recalibration sets, the scoring director or team leader reviews them using descriptors from the Writing Scoring Scale and the anchor responses to confirm the rationale behind each response's score. Starting on the sixth day of scoring, DRC uses validity sets to monitor rater performance. These are sets of items seeded into the operational sets that, on a daily basis, monitor how raters are doing when compared to the known ratings of the validity sets. The raters do not know which items are operational and which are from a validation set.
- For Speaking, DRC uses recalibration sets. At the start of the scoring window, raters take these sets every day to ensure that they are calibrated, and raters' performances on recalibration sets are used for monitoring purposes. Later in the scoring window, recalibration sets are used on a weekly basis to monitor scoring.

Handling Unusual Responses

- Raters can forward responses to team leaders for assistance.
- Responses requiring special attention, including nonscorable responses, are routed to scoring directors for review and resolution.

3.2.3. Writing Scoring Scale

The Writing Scoring Scale has six whole score points that range from 1 to 6. For responses that fall in between the whole score points, "plus" score points are available (e.g., a response that falls between 3 and 4 is scored as 3+). The scale descriptors include three different yet interrelated dimensions: discourse, sentence, and word/phrase. These scale descriptors guide raters as they consider all three dimensions in order to make holistic judgments about which score point best suits a response. The dimensions are distinguished as follows:

- The descriptors for the discourse dimension focus on the degree of organization and the extent to which the response is tailored to the context (e.g., purpose, situation, and audience).
- The descriptors for the sentence dimension evaluate the complexity and grammatical accuracy of sentence structures used in the response.
- The descriptors for the word/phrase dimension specify the range and appropriateness of the original vocabulary used (i.e., text other than that copied and adapted from the stimulus and prompt).

When assigning a score, a rater makes an initial judgment about which whole score point (1–6) best describes a response and then determines whether the three descriptors for that whole score point suit that response. If all three descriptors suit the response, a whole score point is awarded. If there is clear evidence that one or two descriptors from an adjacent score point are a better fit, a plus score point between the two applicable whole score points is awarded. In addition to scale descriptors, scoring rules address special cases where responses are nonscorable, completely or partially off task, and completely or partially off topic. Both nonscorable and completely off-task responses are scored as 0. Completely off-topic responses receive a maximum score of 2+. Partially off-topic responses are scored in their entirety, while partially off-task responses are scored by ignoring the off-task portion of the response and scoring only the on-task portion.

To calculate a raw score for the Writing test, raters' scores for each Writing task are converted to whole numbers ranging from 0 to 9, as shown in

Table 5. On Tier A tests, for all grade-level clusters except for Grade 1, the scores from the three tasks are added to calculate a total raw score, which can range from 0 to 27. For the Grade 1 Tier A test, there are four Writing tasks. The first two of these tasks use a modified version of the Scoring Scale and have score ranges of 0–1 and 0–3, respectively. The third and fourth task use the full scoring scale from 0 to 9; additionally, the last task is weighted as 3. Therefore, the possible final raw scores for Grade 1 Tier A range from 0 to 40.

On Tier B/C tests for all grade-level clusters, results from the different tasks are given different weights. These weights are specified to reflect intended amounts of time that a student should spend on each task. The first task is given a weight of 1, the second task is given a weight of 2, and the third task is given a weight of 3. Thus, for example, a student with raw scores of 5, 6, and 7 on the three tasks would have a total raw score of 38 ([1 * 5] + [2 * 6] + [3 * 7]), while a student with raw scores of 7, 6, and 5 on the three tasks would have a total raw score of 34 ([1 * 7] + [2 * 6] + [3 * 5]). Raw scores on the Tier B/C tests can range from 0 to 54.

Table 5
Rating to Raw Score Conversion (Writing)

Rating	Raw score
Nonscorable	0
1	1
1+	2
2	3
2+	4
3	5
3+	6
4	7
4+	8
5	9
5+	9
6	9

The ACCESS Writing Scoring Scale is distinct from the WIDA Writing Rubric, which is a tool for evaluating student writing in classrooms and for interpreting student scores from ACCESS Online. The Writing Scoring Scale was designed specifically as a scoring tool and is not appropriate for any other purposes.

3.2.4. Speaking Scoring Scale

The Speaking Scoring Scale defines five score points: *Exemplary*, *Strong*, *Adequate*, *Attempted*, and *No Response.* (The final score point only applies if the rater uses one of three nonscorable codes: R = dead air or white noise; F = foreign language response; I = nonscorable utterance.) These score points are applied based on the proficiency level expectations of each task, that is, the level of language proficiency that each task is designed to elicit. These expectations are exemplified by the model student response (see Section 2.1.4). In this way, the model response serves as a scoring benchmark. Raters listen to the model response and score test taker responses relative to the model. A score of *Exemplary* means that the student response demonstrates English language use that is equal to or beyond the English language use illustrated by the model student's response.

The Speaking Scoring Scale includes descriptors for overall language use, response sophistication, language delivery, and word choice. As stated above, the scale is applied relative to the proficiency level demands of the task. For tasks targeting language elicitation at PL 1, there are only three possible score points: *No Response*, *Attempted*, and *Adequate and Above*.

This is the case because appropriate responses to PL 1 tasks are single words and short chunks of language, so it is not possible to reliably distinguish between *Adequate*, *Strong*, and *Exemplary* performances.

To calculate a raw score for the Speaking test, the five score points are converted to whole numbers, as shown in

Table 6. To calculate a total raw score, the raw scores for each task are added together; additionally, in Tier B/C, six points are added to the total raw score, representing a score of *Adequate and Above* for three tasks targeting language at PL 1. Though a Tier B/C student would not be administered any tasks targeting the PL 1 level, it is assumed that a score of *Adequate and Above* would be applicable to such tasks. Thus, on the Pre-A test, scores can range from 0 to 6; on the A test, from 0 to 18; and on the B/C test, from 6 to 30.

Table 6

Rating to Raw Score Conversion (Speaking)

Rating	Raw score
No Response (R, F, or I)*	0
Attempted	1
Adequate/Adequate and Above	2
Strong	3
Exemplary	4

*R = Dead air or white noise; F = Foreign language response; I = Nonscorable utterance.

Speaking tasks are scored using the ACCESS Speaking Scoring Scale. The Speaking Scoring Scale is distinct from the WIDA Speaking Rubric, which is a tool for classroom use and score interpretation. The Speaking Scoring Scale was designed specifically for test scoring use and is not intended for classroom purposes.

3.3. Operational Administration

3.3.1. Administering the Test Practice

The administration of the test practice for an individual test domain takes approximately 5–10 minutes, depending on how many questions students have about the directions or practice items. Additional time should be scheduled for students to go through the test practice again if needed. The narration within the test practice is included both as spoken audio and as text captioning displayed directly on the screen, allowing the student to be able to read along as the script is read aloud.

3.3.2. Listening Test Administration

The Listening test (including test practice items) is designed to take approximately 30–40 minutes. Note that the approximate test administration time does not include convening students, taking attendance, or explaining test directions.

3.3.3. Reading Test Administration

The Reading test (including directions and practice items) is designed to take approximately 35 minutes. Note that the approximate test administration time does not include convening students, taking attendance, or explaining test directions.

3.3.3.1. Reading Test Item Types

The Reading test may include three different item types: multiple choice, hotspot, and drag and drop. Although a student may not see all three of these item types, it is important to ensure that students know what to do for these different item types.

- Multiple choice. Students choose an answer from a set of ordered response options under the question. The response options may be images or text. Students select their answer by clicking anywhere within the box that denotes the response options, including inside the circle that appears to the left of the text or image. Students are able to change their answer by clicking on a different response option.
- Hotspot. Students see a large response area under the question. The response area may be an image, a paragraph of text, or some combination of images and text, such as a timeline, a webpage. The answer choices may be pictures or text and are embedded in the response area inside blue boxes. Students answer the question by clicking on one of the boxes in the response area. Each answer choice changes color when selected. Students are able to change their answers by clicking on a different blue box or by clicking on the reset eraser button, which clears the original response and clicking on a different blue box.
- Drag and drop. There are two examples of this item type. Students see one object, either a small image or a line of text, above the response area, which may be an image, a paragraph of text, or some combination of images and text, such as a timeline, a webpage, etc. The response area has three or four blue boxes in it. To show their answer, students click and drag/move the small object into a blue box within the response area. Students do not have to place the object exactly in the blue box; the object snaps into place when students release the mouse button. In this type of drag and drop item, students are able to change their answer by dragging their object into a different blue box in the response area or by clicking on the reset eraser button, which clears the original response, and then dragging the object into a different blue box in the response area. Alternatively,

students may see three small objects above the response area. In this case, students select one object to drag into the single blue box within the response area.

3.3.4. Writing Test Administration

All students in Grades 1–3 complete the ACCESS for ELLs Writing test on paper. The test is group administered. For Grades 6–12, all students view the Writing prompts on the desktop, laptop, or tablet. The default response mode is keyboarding. For Grades 4–5, all students also view the Writing prompts on the device. However, each state determines whether the default response mode for students in Grades 4–5 will be keyboarding or handwriting. If keyboarding is the default response mode, and upon logging in and starting the test a student expresses discomfort, concern, or anxiety about keyboarding, administrators may switch the student to responding to the Writing test on paper.

The Writing test is designed to take approximately 45–60 minutes. For all grade-level clusters, the Tier B/C Writing tests have recommended timing guidelines for Parts A, B, and C of 10, 20, and 30 minutes, respectively. Note that the approximate test administration time does not include convening students, taking attendance, distributing and collecting test materials, or explaining test directions, including the directions and practice that precede the test.

3.3.4.1. Writing Test Tiers

Student performance on the Listening and Reading tests determines the appropriate tier that the student will take in the Writing and Speaking tests. Once the students have completed the Listening and Reading tests, test coordinators run a Tier Placement Report that identifies the tier each student is assigned to take. Test administrators use the report to know which form to administer to which student. The Writing test has two tiers: A and B/C. In Grades 1–3, students must be tested in groups organized by grade-level cluster and tier.

3.3.5. Speaking Test Administration

The Speaking test (including directions and practice) is designed to take approximately 30 minutes. Note that the approximate test administration time does not include convening students, taking attendance, or explaining test directions.

Recording response time on every task on the Speaking test has a preset time limit, which varies depending on the grade-level cluster, tier, and task level. Students learn about the time limits in the test directions and practice. Students see a circle change color and then disappear as the time to respond elapses. While there is a limit to how long students can take to record their response, students can navigate the directions, practice, and test items at their own pace. Students click the Next button when they are ready to move on from a screen, without time limits. The test does not advance automatically.

3.3.5.1. Speaking Test Tiers

For each grade-level cluster, the Speaking test has three different tiered forms, Pre-A, A, and B/C. The tier the student takes is determined by the student's Listening and Reading test results and automatically loads for the student upon logging into the test platform with test ticket information. The Pre-A tier is designed to address the needs of newcomer students and to allow those students at the beginning stages of English language development an opportunity to respond to tasks appropriate to what they are able to do. Tier Pre-A also includes a simplified version of the Speaking test practice to ease the burden of learning how to respond to Speaking tasks on the screen for newcomer students. The majority of students are placed in either Tier A or Tier B/C.

3.3.5.2. Group vs. Individual Delivery

The Speaking test is administered to small groups of students. For students in all grade-level clusters taking the Tier A and Tier B/C forms, it is recommended that the Speaking test be administered to groups of 3 to 5 students.

It is recommended that students taking the Pre-A form be administered the test individually so test administrators can provide additional support during the test. For students in all tiers, the Speaking test may be administered individually or in smaller groups of students than mentioned above if needed. Test administrators use their professional judgment to consider whether students with high test anxiety or students requiring extra support should be given the test individually or in a very small group.

3.3.6. Test Security

Every effort is made to keep the test secure at all levels of development and administration. WIDA, CAL, and DRC (the entity responsible for printing, distributing, collecting, and scoring the printed tests) follow established policies and procedures regarding the security of the test, and every individual involved in the administration of ACCESS, from the district level to the classroom level, is trained in issues of test security.

All materials for ACCESS for ELLs are considered secure test materials. All users of the WIDA website are prompted to read and sign a Non-Disclosure and User Agreement (NDUA) upon their first login. Use of the WIDA Assessment Management System (WIDA AMS) and INSIGHT test engine are also subject to the terms of use outlined in WIDA AMS. Users are prompted to agree with the test security policy upon their first login. The security of all test materials must be maintained before, during, and after the test administration. Under no circumstances are students permitted to handle secure materials before or after test administration. Test materials should never be left unsecured. The test coordinator should track each secure booklet on the ACCESS for ELLs Security Checklist. Individuals are responsible for the secure documents assigned to them. Secure documents should never be destroyed (e.g., shredded, thrown in the trash) except for soiled documents, which must be destroyed in a secure manner. District and school personnel carrying out their roles in the delivery of this assessment must follow ACCESS for ELLs District and School Test Coordinator Manual guidelines to maintain test security.

3.4. Procedures for Including Students with Disabilities

The WIDA Accessibility and Accommodations Framework provides support for all ELLs, as well as targeted accommodations for students with individualized education plans (IEP) or 504 plans. These supports are intended to increase the accessibility for the assessments for all ELLs. (Please see Accessibility and Accommodations Supplement for detailed information.)

3.4.1. Support Provided to All ELLs

Universal design. ACCESS for ELLs incorporates universal design principles in order to provide greater accessibility for all ELLs. The test items are presented using multiple modalities, including supporting prompts with appropriate animations and graphics, embedded scaffolding, tasks broken into chunks, and modeling that uses task prototypes and guides.

Administrative considerations include adaptive and specialized equipment or furniture, alternative microphone, familiar test administrator, frequent or additional supervised breaks, individual or small group setting, monitoring of the placement of responses in the test booklet or on screen, participation in different testing formats (Paper vs Online), reading aloud to self, specific seating, short segments, verbal praise or tangible reinforcement for on-task or appropriate behavior, and verbal redirection of students' attention to the test (in English or native language).

Universal tools are available to all students taking ACCESS for ELLs in order to address their individual accessibility needs. These may either be embedded in the online test or provided by test administrators during testing. Universal tools do not affect the construct being measured on the assessment.

3.4.2. Support Provided to ELLs with IEP or 504 Plans

Accommodations include allowable changes to the test presentation, response method, timing, and setting in which assessments are administered. Accommodations are intended to provide testing conditions that do not result in changes in what the test measures; that provide comparable test results to those of students who do not receive accommodations; and that do not affect the validity and reliability of the interpretation of the scores for their intended purposes. Accommodations are available only to ELLs with disabilities when listed in an approved IEP or 504 plan, and only when the student requires the accommodation(s) to participate in ACCESS for ELLs meaningfully and appropriately. Accommodations are delivered locally by a test administrator.

Accessibility features include tools that are available to all ELLs taking ACCESS for ELLs. Examples of accessibility features include highlighter, line guide, magnification, and color overlay. All accessibility features are available to all ELLs during testing; specific designation is not required prior to testing to make them available to the student during testing. Features available during online-based test administration include the following:

- Audio amplification device (provided by student)
- Highlight tool
- Line guide
- Zoom tool (magnifier)
- Sticky notes—which allow students to take notes to prepare responses to Writing items. This tool is only available in the Writing domain
- Color overlay—which allows students to change the background color that appears behind text, graphics, and response areas. Five colors are available: pink, yellow, blue, green, and orange.
- Color contrast—which allows students to select from a variety of background/text color combinations
- Keyboard shortcuts/equivalents—which are alternatives to using a mouse (for navigating through the test and using online test tools)
- Scratch/blank paper (to be submitted with the test or disposed of according to state policy)

Allowable test administration procedures are variations in standard test administration procedures that provide flexibility to schools and districts in determining the conditions under which ACCESS for ELLs can be administered most effectively. These procedures are available to any student, as needed, at the discretion of the test coordinator (or principal or designee), provided that all security conditions and staffing requirements are met. Examples of allowable test administration procedures include tests administered by familiar school personnel, in an individual or small group setting, in a separate room, with frequent supervised breaks, or in short segments. For detailed information on the allowable test administration procedures, consult the ACCESS for ELLs Test Administration Manual.

Schools and districts should consider how accessibility features and allowable test administration procedures can support accessibility to the test for *all* ELLs. The accommodations, accessibility features, and allowable test administration procedures are based on (1) accepted practices in ELP assessment; (2) existing accommodation policies of WIDA Consortium member states; (3) consultation with representatives of WIDA member states who are experts in the education and assessment of ELLs and students with disabilities; and (4) the expertise of the test developers at the Center for Applied Linguistics.

WIDA also offers *Alternate ACCESS for ELLs.* This test is intended only for those ELLs who have cognitive disabilities that are so significant as to prevent meaningful participation in

ACCESS testing, even with accommodations. The results of the Alternate ACCESS for ELLs operational administration appear in a separate technical report.

4. Summary of Score Reports

4.1. Individual Student Report

The Individual Student Report contains detailed information about the performance of a single student within Grades K–12. Its primary users are students, parents/guardians, teachers, and school teams. It describes one indicator of a student's English language proficiency, the language needed to access content and succeed in school.

In the score report, there are four domain and four composite scores. Each composite score is represented by a label, a breakdown of how individual domains are used to calculate it, and a visual display of the results.

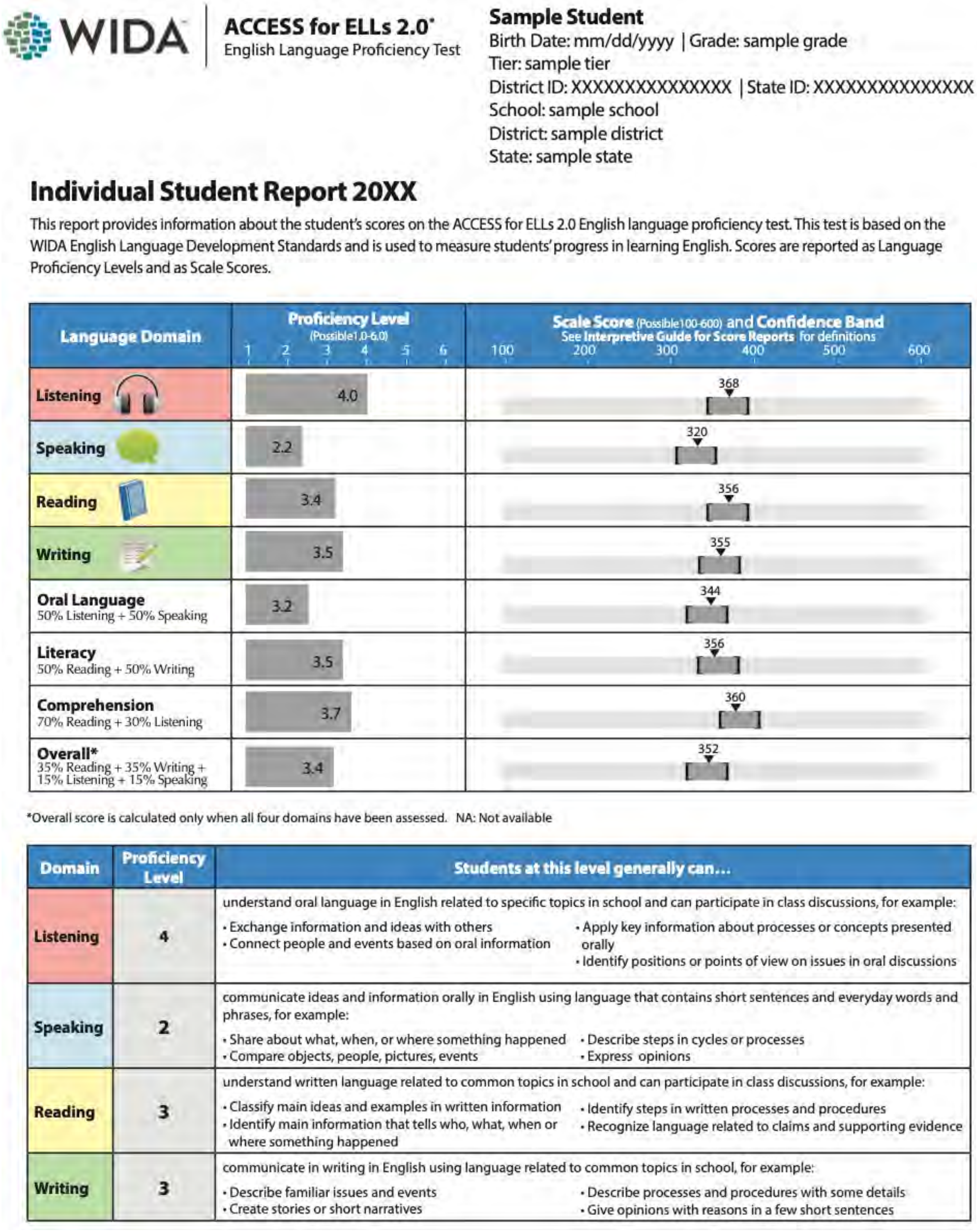
The proficiency level is presented both graphically and as a whole number followed by a decimal. The shaded bar of the graph reflects the exact position of the student's performance on the six-point ELP scale. The whole number reflects a student's ELP level (1–Entering, 2–Emerging, 3–Developing, 4–Expanding, 5–Bridging, and 6–Reaching) in accord with the WIDA ELD Standards. ELLs who attain Level 6, Reaching, have moved through the entire second language continuum, as defined by the test and the WIDA ELD Standards.

The decimal indicates the proportion within the proficiency level range that the student's scale score represents, rounded to the nearest tenth. For example, a proficiency level score of 3.5 is halfway between the 3.0 and 4.0 cut scores.

To the right of the proficiency level is the reported scale score and associated confidence band. The confidence band reflects the standard error of measurement of the scale score, a statistical calculation of a student's likelihood of scoring within a particular range of scores if he or she were to take the same test repeatedly without any change in ability.

If a student does not complete one or more language domains, NA (not available) is inserted in that language domain as well as all applicable composite scores, including the Overall score. Students with identical Overall scores may have very different profiles in terms of their oral language and literacy development.

The Individual Student Report also provides information about the proficiency levels as whole numbers obtained by the student and describes what many students at the reported proficiency level may be expected to be able to do in English. For example, if the student received a proficiency level score of 2 for Speaking, the report will include a description of the type of spoken language the student may be expected to be able to produce. Figure 7 shows a sample Individual Student Report.



WIDA | **ACCESS for ELLs 2.0®**
English Language Proficiency Test

Sample Student
Birth Date: mm/dd/yyyy | Grade: sample grade
Tier: sample tier
District ID: XXXXXXXXXXXXXXX | State ID: XXXXXXXXXXXXXXX
School: sample school
District: sample district
State: sample state

Individual Student Report 20XX

This report provides information about the student's scores on the ACCESS for ELLs 2.0 English language proficiency test. This test is based on the WIDA English Language Development Standards and is used to measure students' progress in learning English. Scores are reported as Language Proficiency Levels and as Scale Scores.

Language Domain	Proficiency Level (Possible 1.0-6.0) 1 2 3 4 5 6	Scale Score (Possible 100-600) and Confidence Band See Interpretive Guide for Score Reports for definitions 100 200 300 400 500 600
Listening	4.0	368
Speaking	2.2	320
Reading	3.4	356
Writing	3.5	355
Oral Language 50% Listening + 50% Speaking	3.2	344
Literacy 50% Reading + 50% Writing	3.5	356
Comprehension 70% Reading + 30% Listening	3.7	360
Overall* 35% Reading + 35% Writing + 15% Listening + 15% Speaking	3.4	352

*Overall score is calculated only when all four domains have been assessed. NA: Not available

Domain	Proficiency Level	Students at this level generally can...
Listening	4	understand oral language in English related to specific topics in school and can participate in class discussions, for example: • Exchange information and ideas with others • Connect people and events based on oral information • Apply key information about processes or concepts presented orally • Identify positions or points of view on issues in oral discussions
Speaking	2	communicate ideas and information orally in English using language that contains short sentences and everyday words and phrases, for example: • Share about what, when, or where something happened • Compare objects, people, pictures, events • Describe steps in cycles or processes • Express opinions
Reading	3	understand written language related to common topics in school and can participate in class discussions, for example: • Classify main ideas and examples in written information • Identify main information that tells who, what, when or where something happened • Identify steps in written processes and procedures • Recognize language related to claims and supporting evidence
Writing	3	communicate in writing in English using language related to common topics in school, for example: • Describe familiar issues and events • Create stories or short narratives • Describe processes and procedures with some details • Give opinions with reasons in a few short sentences

Figure 7. Individual Student Report.

When interpreting scores, the following points should be kept in mind:

- The report provides information on English proficiency. It does not provide information on a student's academic achievement or knowledge of content areas.
- Students do not typically acquire proficiency in Listening, Speaking, Reading, and Writing at the same pace. Generally,
    - Oral language (L+S) is acquired faster than literacy (R+W).
    - Receptive language (L+R) is acquired faster than productive language (S+W).
    - Writing is usually the last domain to be mastered.
- The students' foundation in their home or primary language is a predictor of their English language development. Those who have strong literacy backgrounds in their native language will most likely acquire literacy in English at a quicker pace than students who do not.
- The Overall score is helpful as a summary of other scores and is used because a single number may be needed for reference. However, it is important to remember that it is compensatory; a particularly high score in one domain may effectively raise a low score in another. Similar overall scores can mask very different performances on the test.
- No single score or language proficiency level, including the Overall score (composite), should be used as the sole determiner for making decisions regarding a student's English language proficiency. School work and local assessment throughout the school year also provide evidence of a student's English language development.
- Scale scores from different domains should not be compared. Each domain has its own scale, so scale scores should not be compared, such as comparing Listening to Reading. Proficiency level scores can be used for such comparisons.
- To compare test scores from different years, either scale scores or proficiency level scores can be used, although it is easier to see changes when examining scale scores.

For detailed information about score reports, please refer to the Interpretive Guide.

4.2. Other Reports

Student Roster Report. The Student Roster Report contains information on a group of students within a single school and grade. It provides scale scores for individual students in each language domain and composite, identical to those in the Individual Student Report. Its intended users are teachers, program coordinators/directors, and administrators.

Frequency Reports. The primary audience for frequency reports is typically program coordinators/directors, administrators, and boards of education. There are three types of frequency reports:

- School Frequency Report
- District Frequency Report
- State Frequency Report

They all show the number and percentage of tested students who attain each proficiency level within a given population.

Part 2:
Technical Results

Contents

1. Student Participation and Performance

1.1. Participation 1-3
  1.1.1. Grade-Level Cluster 1-3
  1.1.2. Grade 1-8
1.2. Scale Score Results 1-13
  1.2.1. Mean Scale Score Across Domain and Composite Score by Cluster 1-13
  1.2.2. Mean Scale Score Across Domain and Composite Score by Grade 1-18
  1.2.3. Correlations 1-27
1.3. Proficiency Level Results 1-29
  1.3.1. Domains 1-29
  1.3.2. Composites 1-37

2. Analysis of Domains

2.1 Complete Item or Task Analysis and Summary 2-4
  2.1.1 Listening 2-6
  2.1.2 Reading 2-16
  2.1.3 Writing 2-31
  2.1.4 Speaking 2-41
2.2 DIF Analysis and Summary 2-46
  2.2.1 Listening 2-49
  2.2.2 Reading 2-51
  2.2.3 Writing 2-54
  2.2.4 Speaking 2-57
2.3 Raw Score Distribution for Speaking and Writing 2-61
  2.3.1 Listening 2-62
  2.3.2 Reading 2-62
  2.3.3 Writing 2-62
  2.3.4 Speaking 2-72
2.4 Scale Score Distribution 2-87
  2.4.1 Listening 2-88
  2.4.2 Reading 2-93
  2.4.3 Writing 2-98
  2.4.4 Speaking 2-113
2.5 Proficiency Level Distribution 2-133
  2.5.1 Listening 2-134
  2.5.2 Reading 2-139
  2.5.3 Writing 2-144
  2.5.4 Speaking 2-159
2.6 Raw Score to Scale Score to Proficiency Level Conversion for Speaking and Writing 2-179
  2.6.1 Listening 2-179
  2.6.2 Reading 2-179

2.6.3 Writing 2-180
2.6.4 Speaking 2-192
2.7 Equating Summary 2-202
2.7.1 Listening 2-204
2.7.2 Reading 2-214
2.7.3 Writing 2-229
2.7.4 Speaking 2-239
2.8 Test Characteristic Curve 2-244
2.8.1 Listening 2-244
2.8.2 Reading 2-244
2.8.3 Writing 2-245
2.8.4 Speaking 2-253
2.9 Test Information Function 2-263
2.9.1 Listening 2-264
2.9.2 Reading 2-266
2.9.3 Writing 2-269
2.9.4 Speaking 2-277
2.10 Reliability 2-287
2.10.1 Listening 2-289
2.10.2 Reading 2-290
2.10.3 Writing 2-292
2.10.4 Speaking 2-294
2.11 Interrater Agreement 2-296
2.11.3 Writing 2-296
2.11.4 Speaking 2-299
2.12 Conditional Standard Errors of Measurement at Cut Score 2-303
2.12.1 Listening 2-304
2.12.2 Reading 2-306
2.12.3 Writing 2-310
2.12.4 Speaking 2-313
2.13 Accuracy and Consistency 2-316
2.13.1 Listening 2-318
2.13.2 Reading 2-324
2.13.3 Writing 2-330
2.13.4 Speaking 2-336

3. Analyses of Composite Scores

3.1 Scale Score Distribution for Composites 3-1
3.1.1 Oral 3-2
3.1.2 Literacy 3-7
3.1.3 Comprehension 3-12
3.1.4 Overall 3-17
3.2 Proficiency Level Distribution for Composites 3-22
3.2.1 Oral 3-23

3.2.2 Literacy 3-28
3.2.3 Comprehension 3-33
3.2.4 Overall 3-38
3.3 Reliability 3-43
3.3.1 Oral 3-44
3.3.2 Literacy 3-46
3.3.3 Comprehension 3-48
3.3.4 Overall 3-50
3.4 Accuracy and Consistency of Composites 3-54
3.4.1 Oral 3-54
3.4.2 Literacy 3-60
3.4.3 Comprehension 3-66
3.4.4 Overall 3-72
3.5 CSEM for Composites 3-79
3.5.1 Oral 3-81
3.5.2 Literacy 3-84
3.5.3 Comprehension 3-87
3.5.4 Overall 3-90

4. Annual Updates of Validity Evidence

4.1. Standards 4-1
4.1.1. Test content 4-1
4.1.2. Response processes 4-1
4.1.3. Internal structure 4-1
4.1.4. Relations to other structure 4-2
4.2. Annual validity studies 4-2
4.2.1. ACCESS for ELLs 2.0 Series 400 Construct Validity Study 4-2
4.2.2. WIDA Screener Online and WIDA ACCESS for ELLs: Examining the Relationship between Student Scores on Two Assessments 4-3
4.2.3. 2018 ACCESS for ELLs 2.0 Speaking Study 4-3

1. Student Participation and Performance

This section of the report provides an overview of students' participation, the distribution of students' scale scores, and the distribution of students' proficiency levels. Results are presented, where appropriate, by grade-level cluster, grade, and tier (for Writing and Speaking), and also by state, by gender, and by race and ethnicity.

Following the approach of the Census Bureau, ethnicity is a binary category (Hispanic or non-Hispanic), with five categories for race (American Indian/Alaskan Native, Asian, Black/African American, Pacific Islander/Hawaiian, and White) that are not mutually exclusive. Thus, for example, Student A may be labeled as Hispanic for ethnicity and Asian for race, while Student B may be labeled as non-Hispanic for ethnicity and both American Indian/Alaskan Native and Black/African American for race. Students who are labeled as Hispanic are included in the Hispanic (of any race) category, regardless of how many racial categories they are included in. Students who are identified in one racial category (e.g., Asian) who have not been identified as Hispanic are identified in only one racial category; if they are identified in more than one racial category and have not been identified as Hispanic, they are labeled non-Hispanic multiracial.

There is a subset of students who are included in the descriptions of student participation and performance but are excluded from subsequent analyses, namely, students who were flagged as potentially having experienced test interruptions. Using telemetry data, three variables were selected that might potentially indicate interruption (that is, testing experiences that are outside of regular testing experiences). The interruption indicators WIDA used are (1) longer than expected testing time, (2) number of appearances (i.e., more than 1) of test items, and (3) number of log-ins. Records are flagged if they fall outside of established criteria for any of these three indicators. Students whose records are flagged as interrupted are included in the tables that describe participation in the assessment but are excluded from all subsequent analyses. Table 1.1 summarizes the numbers of students who are excluded from these analyses.

Table 1.1

Students Excluded from Analysis due to Test Interruptions by Domain and Cluster

Domain	Cluster	No. of Students	Percent
Listening	1	8,568	10.17%
	2–3	17,706	21.02%
	4–5	16,164	19.19%
	6–8	20,052	23.80%
	9–12	21,745	25.81%
	Total	84,235	100.00%
Reading	1	6,215	6.98%
	2–3	16,906	18.99%
	4–5	20,378	22.89%
	6–8	19,862	22.31%
	9–12	25,661	28.83%
	Total	89,022	100.00%
Writing	1	0	0.00%
	2–3	0	0.00%
	4–5	32,893	35.85%
	6–8	29,046	31.66%
	9–12	29,813	32.49%
	Total	91,752	100.00%
Speaking	1	12,186	11.12%
	2–3	24,556	22.41%
	4–5	25,486	23.25%
	6–8	23,084	21.06%
	9–12	24,288	22.16%
	Total	109,600	100.00%

1.1. Participation

Participation in ACCESS Online is shown in three ways: by grade-level cluster, by grade, and, for Writing and Speaking only, by tier.

1.1.1. Grade-Level Cluster

Table 1.1.1.1 shows participation across the 38 WIDA states and U.S. territories that participated in the ACCESS Online operational testing program in 2018–2019 by grade-level cluster. The 38 rows show the number of students in that grade-level cluster who took the test by state, and the final row shows the total number of participants across all 38 states and U.S. territories.

Table 1.1.1.2 shows participation by grade-level cluster by gender across all 38 states and U.S. territories combined, while Table 1.1.1.3 shows participation by grade-level cluster by ethnicity across all 38 states and U.S. territories.

Table 1.1.1.4 shows participation by grade-level cluster and tier for all Writing and Speaking forms.

Table 1.1.1.1

Participation by Cluster by State, S403 Online

State	**Cluster**					**Total**
	1	**2–3**	**4–5**	**6–8**	**9–12**	
AK	1,148	2,482	2,589	2,685	2,530	11,434
AL	3,758	7,154	6,017	4,091	4,179	25,199
BI	153	312	374	538	233	1,610
CO	9,826	20,699	17,355	18,292	17,801	83,973
DC	1,062	1,904	1,460	1,152	1,526	7,104
DE	1,718	3,468	3,026	1,892	1,930	12,034
GA	14,470	28,081	23,923	17,087	16,371	99,932
HI	1,977	3,941	3,203	2,817	2,932	14,870
ID	2,179	4,839	3,532	3,562	2,857	16,969
IL	17,374	43,738	44,395	33,149	29,060	167,716
IN	7,572	15,226	12,939	9,290	12,441	57,468
KY	3,267	6,673	4,975	4,213	5,433	24,561
MA	11,160	20,490	15,600	15,930	18,647	81,827
MD	10,139	19,953	15,488	12,252	16,312	74,144
ME	488	955	846	959	1,061	4,309
MI	9,051	19,651	17,177	19,542	22,858	88,279
MN	8,026	15,770	12,822	11,939	13,195	61,752
MO	4,146	7,944	6,255	5,700	5,271	29,316
MP	95	218	192	432	423	1,360
MT	243	534	787	765	494	2,823
NC	12,458	25,940	25,526	17,395	19,966	101,285
ND	440	787	692	643	834	3,396
NH	469	1,026	939	932	1,016	4,382
NJ	11,705	19,691	13,498	12,973	17,909	75,776
NM	4,524	9,795	10,594	10,597	10,672	46,182
NV	6,753	15,120	13,524	13,928	15,366	64,691
OK	6,332	12,723	10,520	7,918	7,943	45,436
PA	5,754	12,777	11,661	13,463	15,946	59,601
RI	1,252	2,840	2,610	2,587	3,346	12,635
SC	2,593	6,287	6,212	7,036	8,121	30,249
SD	659	1,278	1,114	851	1,001	4,903
TN	5,759	10,488	8,105	7,791	8,493	40,636
UT	4,968	11,185	11,225	9,254	7,982	44,614
VA	12,087	24,241	18,447	14,910	19,333	89,018
VI	33	42	54	66	71	266
VT	208	392	323	247	340	1,510
WI	5,677	11,348	10,892	10,862	9,595	48,374
WY	287	666	463	423	488	2,327
Total	189,810	390,658	339,354	298,163	323,976	1,541,961

Table 1.1.1.2

Participation by Cluster by Gender, S403 Online

Cluster		Gender			Total
		F	M	Missing	
1	Count	88,567	99,927	1,316	189,810
	% within Cluster	46.7%	52.6%	0.7%	100.0%
2–3	Count	182,013	206,395	2,250	390,658
	% within Cluster	46.6%	52.8%	0.6%	100.0%
4–5	Count	153,692	183,553	2,109	339,354
	% within Cluster	45.3%	54.1%	0.6%	100.0%
6–8	Count	128,731	167,102	2,330	298,163
	% within Cluster	43.2%	56.0%	0.8%	100.0%
9–12	Count	140,030	180,610	3,336	323,976
	% within Cluster	43.2%	55.7%	1.0%	100.0%
Total	Count	693,033	837,587	11,341	1,541,961
	% within Cluster	44.9%	54.3%	0.7%	100.0%

Table 1.1.1.3

Participation by Cluster by Ethnicity, S403 Online

Cluster		Hispanic/Non-Hispanic			Total
		Hispanic	Other	Unknown	
1	Count	117,646	60,825	11,339	189,810
	% within Cluster	62.0%	32.0%	6.0%	100.0%
2–3	Count	251,337	115,929	23,392	390,658
	% within Cluster	64.3%	29.7%	6.0%	100.0%
4–5	Count	226,260	88,231	24,863	339,354
	% within Cluster	66.7%	26.0%	7.3%	100.0%
6–8	Count	191,629	78,478	28,056	298,163
	% within Cluster	64.3%	26.3%	9.4%	100.0%
9–12	Count	201,159	91,756	31,061	323,976
	% within Cluster	62.1%	28.3%	9.6%	100.0%
Total	Count	988,031	435,219	118,711	1,541,961
	% within Cluster	64.1%	28.2%	7.7%	100.0%

Table 1.1.1.4

Participation by Cluster by Tier by Domain, S403 Online

Cluster			Domain	
			Writing	Speaking
1	Tier	Pre-A	-	5,271
		A	154,669	64,971
		BC	35,116	119,567
	Total		189,785	189,809
2–3	Tier	Pre-A	-	16,757
		A	93,720	83,643
		BC	296,872	290,246
	Total		390,592	390,646
4–5	Tier	Pre-A	-	6,176
		A	47,718	29,919
		BC	291,618	303,246
	Total		339,336	339,341
6–8	Tier	Pre-A	-	11,773
		A	109,741	61,672
		BC	188,416	224,711
	Total		298,157	298,156
9–12	Tier	Pre-A	-	23,732
		A	119,084	132,058
		BC	204,865	168,166
	Total		323,949	323,956

1.1.2. Grade

This section provides tables parallel to those in the previous section, but broken out by grade rather than by grade-level cluster.

Table 1.1.2.1

Participation by Grade by State, S403 Online

State	Grade												Total
	1	2	3	4	5	6	7	8	9	10	11	12	
AK	1,148	1,217	1,265	1,334	1,255	1,035	843	807	780	676	600	474	11,434
AL	3,758	3,579	3,575	3,372	2,645	1,784	1,199	1,108	1,455	1,080	886	758	25,199
BI	153	169	143	211	163	189	177	172	64	68	57	44	1,610
CO	9,826	10,382	10,317	9,362	7,993	6,434	6,225	5,633	5,632	4,825	4,075	3,269	83,973
DC	1,062	1,037	867	771	689	462	344	346	562	371	385	208	7,104
DE	1,718	1,661	1,807	1,716	1,310	796	600	496	709	516	366	339	12,034
GA	14,470	14,422	13,659	14,085	9,838	6,318	5,633	5,136	6,407	4,370	3,324	2,270	99,932
HI	1,977	1,924	2,017	1,602	1,601	1,010	929	878	1,013	785	706	428	14,870
ID	2,179	2,381	2,458	1,852	1,680	1,328	1,115	1,119	896	720	681	560	16,969
IL	17,374	20,053	23,685	24,605	19,790	13,429	10,425	9,295	9,336	7,749	6,953	5,022	167,716
IN	7,572	7,566	7,660	7,295	5,644	3,634	2,925	2,731	3,612	3,180	3,501	2,148	57,468
KY	3,267	3,483	3,190	2,918	2,057	1,490	1,381	1,342	1,789	1,421	1,283	940	24,561
MA	11,160	10,693	9,797	8,756	6,844	5,392	5,348	5,190	6,345	4,868	4,175	3,259	81,827
MD	10,139	10,267	9,686	9,118	6,370	4,380	4,023	3,849	5,943	4,298	3,410	2,661	74,144
ME	488	473	482	485	361	321	313	325	299	272	263	227	4,309
MI	9,051	9,892	9,759	9,241	7,936	6,623	6,660	6,259	6,855	6,188	5,149	4,666	88,279
MN	8,026	7,941	7,829	7,392	5,430	4,011	4,063	3,865	4,077	3,452	3,223	2,443	61,752
MO	4,146	4,107	3,837	3,555	2,700	1,982	1,951	1,767	1,721	1,418	1,310	822	29,316
MP	95	127	91	96	96	133	141	158	151	106	105	61	1,360
MT	243	249	285	396	391	333	237	195	173	135	103	83	2,823
NC	12,458	12,742	13,198	13,421	12,105	6,863	5,414	5,118	6,718	5,578	4,762	2,908	101,285
ND	440	392	395	375	317	210	236	197	212	227	229	166	3,396
NH	469	487	539	498	441	354	304	274	327	258	248	183	4,382
NJ	11,705	10,440	9,251	7,838	5,660	4,385	4,310	4,278	5,456	4,854	4,233	3,366	75,776
NM	4,524	4,576	5,219	5,489	5,105	4,052	3,522	3,023	3,462	2,842	2,504	1,864	46,182
NV	6,753	7,478	7,642	7,653	5,871	4,427	4,684	4,817	4,484	4,200	3,930	2,752	64,691
OK	6,332	6,407	6,316	5,812	4,708	3,391	2,604	1,923	2,452	2,223	1,960	1,308	45,436
PA	5,754	6,228	6,549	6,193	5,468	4,583	4,614	4,266	4,552	4,185	3,741	3,468	59,601
RI	1,252	1,372	1,468	1,390	1,220	912	852	823	970	866	865	645	12,635
SC	2,593	2,956	3,331	3,112	3,100	2,328	2,292	2,416	2,667	2,189	1,987	1,278	30,249
SD	659	663	615	583	531	342	259	250	310	308	236	147	4,903
TN	5,759	5,579	4,909	4,487	3,618	2,705	2,482	2,604	3,235	2,100	1,851	1,307	40,636
UT	4,968	5,495	5,690	6,025	5,200	3,917	3,034	2,303	2,392	2,273	1,987	1,330	44,614
VA	12,087	12,451	11,790	10,832	7,615	5,251	4,920	4,739	6,602	4,964	4,992	2,775	89,018
VI	33	15	27	26	28	27	17	22	26	22	15	8	266
VT	208	194	198	173	150	100	76	71	95	74	105	66	1,510
WI	5,677	5,635	5,713	5,840	5,052	4,122	3,691	3,049	2,918	2,526	2,326	1,825	48,374
WY	287	341	325	285	178	137	137	149	129	121	123	115	2,327
Total	189,810	195,074	195,584	188,194	151,160	109,190	97,980	90,993	104,826	86,308	76,649	56,193	1,541,961

Table 1.1.2.2

Participation by Grade by Gender, S403 Online

Grade		Gender			Total
		F	M	Missing	
1	Count	88,567	99,927	1,316	189,810
	% within Grade	46.7%	52.6%	0.7%	100.0%
2	Count	91,212	102,691	1,171	195,074
	% within Grade	46.8%	52.6%	0.6%	100.0%
3	Count	90,801	103,704	1,079	195,584
	% within Grade	46.4%	53.0	0.6%	100.0%
4	Count	86,353	100,714	1,127	188,194
	% within Grade	45.9%	53.5%	0.6%	100.0%
5	Count	67,339	82,839	982	151,160
	% within Grade	44.5%	54.8%	0.6%	100.0%
6	Count	46,801	61,495	894	109,190
	% within Grade	42.9%	56.3%	0.8%	100.0%
7	Count	42,365	54,862	753	97,980
	% within Grade	43.2%	56.0%	0.8%	100.0%
8	Count	39,565	50,745	683	90,993
	% within Grade	43.5%	55.8%	0.8%	100.0%
9	Count	43,668	59,772	1,386	104,826
	% within Grade	41.7%	57.0%	1.3%	100.0%
10	Count	37,170	48,410	728	86,308
	% within Grade	43.1%	56.1%	0.8%	100.0%
11	Count	33,618	42,240	791	76,649
	% within Grade	43.9%	55.1%	1.0%	100.0%
12	Count	25,574	30,188	431	56,193
	% within Grade	45.5%	53.7%	0.8%	100.0%
Total	Count	693,033	837,587	11,341	1,541,961
	% within Grade	44.9%	54.3%	0.7%	100.0%

Table 1.1.2.3

Participation by Grade by Ethnicity, S403 Online

Grade		Hispanic/Non-Hispanic			Total
		Hispanic	Other	Unknown	
1	Count	117,646	60,825	11,339	189,810
	% within Grade	62.0%	32.0%	6.0%	100.0%
2	Count	123,723	59,838	11,513	195,074
	% within Grade	63.4%	30.7%	5.9%	100.0%
3	Count	127,614	56,091	11,879	195,584
	% within Grade	65.2%	28.7%	6.1%	100.0%
4	Count	124,555	50,254	13,385	188,194
	% within Grade	66.2%	26.7%	7.1%	100.0%
5	Count	101,705	37,977	11,478	151,160
	% within Grade	67.3%	25.1%	7.6%	100.0%
6	Count	71,752	27,345	10,093	109,190
	% within Grade	65.7%	25.0%	9.2%	100.0%
7	Count	62,160	26,447	9,373	97,980
	% within Grade	63.4%	27.0%	9.6%	100.0%
8	Count	57,717	24,686	8,590	90,993
	% within Grade	63.4%	27.1%	9.4%	100.0%
9	Count	66,714	27,309	10,803	104,826
	% within Grade	63.6%	26.1%	10.3%	100.0%
10	Count	54,471	24,041	7,796	86,308
	% within Grade	63.1%	27.9%	9.0%	100.0%
11	Count	47,077	22,605	6,967	76,649
	% within Grade	61.4%	29.5%	9.1%	100.0%
12	Count	32,897	17,801	5,495	56,193
	% within Grade	58.5%	31.7%	9.8%	100.0%
Total	Count	988,031	435,219	118,711	1,541,961
	% within Grade	64.1%	28.2%	7.7%	100.0%

Table 1.1.2.4

Participation by Grade by Tier by Domain, S403 Online

Grade			Domain	
			Writing	Speaking
1	Tier	Pre-A	-	5,271
		A	154,669	64,971
		BC	35,116	119,567
	Total		189,785	189,809
2	Tier	Pre-A	-	7,077
		A	56,228	44,059
		BC	138,798	143,934
	Total		195,026	195,070
3	Tier	Pre-A	-	9,680
		A	37,492	39,584
		BC	158,074	146,312
	Total		195,566	195,576
4	Tier	Pre-A	-	2,455
		A	24,067	16,869
		BC	164,115	168,861
	Total		188,182	188,185
5	Tier	Pre-A	-	3,721
		A	23,651	13,050
		BC	127,503	134,385
	Total		151,154	151,156
6	Tier	Pre-A	-	2,891
		A	34,070	19,079
		BC	75,118	87,218
	Total		109,188	109,188
7	Tier	Pre-A	-	3,939
		A	38,144	16,047
		BC	59,834	77,991
	Total		97,978	97,977
8	Tier	Pre-A	-	4,943
		A	37,527	26,546
		BC	53,464	59,502
	Total		90,991	90,991

Grade			Domain	
			Writing	Speaking
9	Tier	Pre-A	-	5,593
		A	45,234	58,243
		BC	59,586	40,985
	Total		104,820	104,821
10	Tier	Pre-A	-	6,645
		A	31,560	34,003
		BC	54,743	45,656
	Total		86,303	86,304
11	Tier	Pre-A	-	6,377
		A	25,523	15,089
		BC	51,120	55,177
	Total		76,643	76,643
12	Tier	Pre-A	-	5,117
		A	16,767	24,723
		BC	39,416	26,348
	Total		56,183	56,188

1.2. Scale Score Results

This section provides information on students' scale score results.

1.2.1. Mean Scale Score Across Domain and Composite Score by Cluster

This section shows mean (average) scale scores by grade-level cluster across the eight scores awarded, first for the four domains (Listening, Reading, Writing, and Speaking) and then for the four composites (Oral Language, Literacy, Comprehension, and Overall Composite). The mean scale scores are expected to increase as grade increases, as ACCESS is vertically scaled, however there is also an intersection between this principle and the population of test takers. In this section, under each average, the number of students in each group is also given.

Additional tables show this information by gender, and by race and ethnicity.

Table 1.2.1.1
Mean Scale Scores by Cluster, S403 Online

Cluster		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
1	Mean	325.28	291.83	257.60	251.47	288.50	274.84	301.87	278.65
	N	181,107	183,421	189,643	177,524	169,891	183,361	175,698	165,022
2–3	Mean	337.28	326.16	306.94	265.92	301.72	316.71	329.54	311.97
	N	372,664	373,371	390,277	365,867	350,367	373,230	357,977	337,194
4–5	Mean	415.77	357.15	340.26	308.77	362.53	348.80	374.85	352.67
	N	322,964	318,641	306,009	313,658	300,016	291,006	305,345	262,341
6–8	Mean	396.98	348.79	329.09	312.92	355.14	339.04	363.44	343.63
	N	277,787	277,957	268,694	274,839	258,336	255,314	262,840	228,280
9–12	Mean	392.26	378.41	359.07	310.40	351.49	368.90	382.83	363.45
	N	301,765	297,819	293,627	299,309	281,269	275,972	282,140	247,592

Table 1.2.1.2

Mean Scale Scores by Gender, S403 Online

Cluster	Gender		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
1	F	Mean	329.05	294.05	261.61	257.70	293.49	277.94	304.51	282.27
		N	84,584	85,320	88,494	82,942	79,431	85,301	81,797	76,928
	M	Mean	322.20	289.96	254.25	246.17	284.31	272.25	299.67	275.63
		N	95,256	96,829	99,834	93,348	89,267	96,788	92,674	86,937
	Missing	Mean	305.00	286.31	241.98	233.57	269.45	264.18	291.90	265.75
		N	1,267	1,272	1,315	1,234	1,193	1,272	1,227	1,157
2–3	F	Mean	339.09	328.49	312.75	271.87	305.58	320.79	331.69	315.96
		N	173,798	173,462	181,866	170,570	163,474	173,414	166,527	156,985
	M	Mean	335.91	324.19	301.98	260.87	298.52	313.25	327.76	308.62
		N	196,719	197,749	206,162	193,186	184,867	197,657	189,381	178,251
	Missing	Mean	316.82	319.12	291.92	248.31	282.28	305.77	318.58	298.44
		N	2,147	2,160	2,249	2,111	2,026	2,159	2,069	1,958
4–5	F	Mean	415.72	358.89	345.55	311.29	363.76	352.35	376.03	355.47
		N	146,422	143,800	138,078	142,005	135,974	131,023	138,007	118,362
	M	Mean	416.11	355.83	336.12	306.90	361.79	346.07	374.05	350.58
		N	174,527	172,892	166,013	169,699	162,157	158,177	165,455	142,331
	Missing	Mean	389.43	346.12	317.38	287.77	338.28	331.24	358.87	332.47
		N	2,015	1,949	1,918	1,954	1,885	1,806	1,883	1,648
6–8	F	Mean	396.99	352.23	334.87	314.40	355.86	343.66	365.80	347.04
		N	120,267	119,706	115,677	118,374	111,569	109,867	113,552	98,390
	M	Mean	397.25	346.32	324.96	312.17	354.92	335.74	361.83	341.28
		N	155,344	156,029	150,919	154,301	144,732	143,417	147,198	128,060
	Missing	Mean	377.22	336.70	307.10	286.56	331.64	322.13	349.13	324.90
		N	2,176	2,222	2,098	2,164	2,035	2,030	2,090	1,830
9–12	F	Mean	393.22	382.10	365.30	313.21	353.37	373.94	385.71	367.50
		N	130,894	128,400	126,323	129,154	121,779	118,621	122,097	106,730
	M	Mean	391.89	375.80	354.76	308.69	350.46	365.41	380.88	360.72
		N	167,760	166,373	164,273	167,052	156,583	154,507	157,163	138,314
	Missing	Mean	371.88	365.01	332.57	285.84	328.52	349.07	367.24	342.01
		N	3,111	3,046	3,031	3,103	2,907	2,844	2,880	2,548

Table 1.2.1.3

Mean Scale Scores by Ethnicity, S403 Online

Cluster	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
1	Non-Hispanic Asian	Mean	334.64	307.56	270.61	257.44	296.03	289.21	315.67	290.87
		N	25,087	25,373	26,289	24,608	23,560	25,370	24,301	22,861
	Non-Hispanic Pacific Islander	Mean	306.70	286.89	257.14	243.67	275.41	272.20	292.71	272.77
		N	1,554	1,589	1,648	1,554	1,474	1,588	1,507	1,432
	Non-Hispanic Black	Mean	323.64	293.89	258.81	261.96	293.13	276.54	302.88	281.34
		N	9,769	9,948	10,311	9,598	9,123	9,943	9,471	8,849
	Hispanic (Of Any Race)	Mean	323.06	287.60	254.16	248.61	285.98	271.02	298.27	275.27
		N	112,427	113,988	117,550	110,027	105,409	113,945	109,305	102,591
	Non-Hispanic American Indian	Mean	324.31	284.66	254.36	249.40	287.04	269.67	296.52	274.47
		N	1,545	1,552	1,621	1,518	1,451	1,552	1,481	1,400
	Non-Hispanic Multiracial	Mean	341.25	303.77	266.58	263.50	302.48	285.21	315.13	290.28
		N	832	837	871	831	793	837	801	764
	Non-Hispanic White	Mean	333.12	297.04	263.55	260.03	296.69	280.35	307.79	284.92
		N	19,030	19,234	20,025	18,779	17,887	19,231	18,348	17,275
	Unknown	Mean	315.93	289.27	251.28	243.13	279.62	270.37	297.23	272.70
		N	10,863	10,900	11,328	10,609	10,194	10,895	10,484	9,850
2–3	Non-Hispanic Asian	Mean	350.70	338.45	319.10	272.23	311.47	328.98	342.14	323.40
		N	44,742	44,845	46,797	43,862	42,074	44,830	43,073	40,615
	Non-Hispanic Pacific Islander	Mean	320.29	318.74	306.43	256.85	288.47	312.69	319.42	305.24
		N	3,034	3,063	3,194	3,025	2,882	3,063	2,927	2,785
	Non-Hispanic Black	Mean	335.30	325.68	305.15	272.31	303.96	315.62	328.65	311.91
		N	20,048	20,113	21,165	19,668	18,705	20,108	19,181	17,935
	Hispanic (Of Any Race)	Mean	334.48	323.56	304.76	263.98	299.36	314.32	326.89	309.62
		N	240,018	240,528	251,114	235,569	225,780	240,428	230,721	217,414
	Non-Hispanic American Indian	Mean	333.49	318.30	300.23	260.53	296.93	309.28	322.80	305.28
		N	3,209	3,261	3,396	3,195	3,027	3,260	3,088	2,912
	Non-Hispanic Multiracial	Mean	351.90	335.03	312.47	273.24	313.13	324.04	340.46	321.02
		N	1,613	1,613	1,675	1,585	1,532	1,613	1,558	1,483
	Non-Hispanic White	Mean	347.13	331.60	312.04	273.48	310.50	321.97	336.25	318.29
		N	37,605	37,638	39,567	37,096	35,364	37,625	35,930	33,871
	Unknown	Mean	327.62	322.19	299.69	257.15	292.44	311.03	323.86	305.07
		N	22,395	22,310	23,369	21,867	21,003	22,303	21,499	20,179

Cluster	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
4–5	Non-Hispanic Asian	Mean	424.67	368.42	348.99	315.57	370.35	358.86	385.40	361.99
		N	30,845	30,552	29,391	29,979	28,797	28,150	29,408	25,611
	Non-Hispanic Pacific Islander	Mean	403.21	350.47	339.49	301.87	352.89	344.84	366.23	346.90
		N	2,500	2,495	2,334	2,491	2,358	2,235	2,364	2,022
	Non-Hispanic Black	Mean	413.34	354.11	333.68	314.63	364.12	344.00	372.05	349.94
		N	17,037	16,853	16,201	16,541	15,683	15,311	16,037	13,612
	Hispanic (Of Any Race)	Mean	415.74	356.21	340.66	307.83	362.08	348.57	374.20	352.41
		N	215,657	212,768	204,020	209,414	200,471	194,286	204,028	175,399
	Non-Hispanic American Indian	Mean	412.82	352.36	336.25	306.49	360.10	344.67	370.58	349.11
		N	3,430	3,433	3,254	3,300	3,161	3,116	3,293	2,787
	Non-Hispanic Multiracial	Mean	424.60	364.11	341.81	315.96	370.02	352.94	382.29	357.22
		N	1,079	1,047	1,022	1,050	1,003	963	1,006	873
	Non-Hispanic White	Mean	420.83	360.82	343.75	315.81	368.64	352.40	378.91	357.09
		N	28,842	28,265	26,457	27,954	26,644	24,938	26,972	22,302
	Unknown	Mean	401.32	349.86	326.90	296.35	348.86	338.10	365.25	340.72
		N	23,574	23,228	23,330	22,929	21,899	22,007	22,237	19,735
6–8	Non-Hispanic Asian	Mean	409.46	362.66	342.27	326.13	368.04	352.74	376.88	357.15
		N	24,885	24,700	23,840	24,361	23,115	22,665	23,502	20,417
	Non-Hispanic Pacific Islander	Mean	386.60	342.06	325.39	306.26	346.42	334.04	355.62	337.87
		N	2,320	2,325	2,183	2,353	2,146	2,020	2,171	1,772
	Non-Hispanic Black	Mean	396.53	346.50	323.27	322.28	359.54	335.02	361.74	342.25
		N	17,530	17,353	16,749	17,269	16,155	15,755	16,364	13,952
	Hispanic (Of Any Race)	Mean	395.85	347.78	329.24	311.37	353.81	338.60	362.39	342.91
		N	178,861	179,405	173,576	177,200	166,759	165,401	169,849	148,120
	Non-Hispanic American Indian	Mean	397.62	347.12	329.67	316.30	357.04	338.56	362.42	344.18
		N	3,549	3,580	3,506	3,542	3,317	3,338	3,398	2,988
	Non-Hispanic Multiracial	Mean	408.76	357.30	335.36	323.95	366.39	346.76	372.95	352.20
		N	831	824	816	820	785	774	793	712
	Non-Hispanic White	Mean	405.00	354.15	334.13	321.79	363.71	344.38	369.67	350.09
		N	23,931	23,763	23,076	23,581	22,091	21,732	22,379	19,325
	Unknown	Mean	386.18	339.78	314.72	296.49	341.22	327.05	353.75	330.67
		N	25,880	26,007	24,948	25,713	23,968	23,629	24,384	20,994

Cluster	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
9–12	Non-Hispanic Asian	Mean	403.76	392.30	374.68	329.65	366.74	383.64	395.94	378.29
		N	29,644	28,889	28,340	29,018	27,455	26,592	27,556	23,973
	Non-Hispanic Pacific Islander	Mean	389.50	377.34	364.80	305.88	348.08	371.05	381.22	363.98
		N	2,326	2,297	2,213	2,323	2,128	2,025	2,116	1,760
	Non-Hispanic Black	Mean	388.21	378.88	358.37	318.30	353.18	368.85	381.86	363.82
		N	23,150	22,420	22,159	22,913	21,432	20,462	21,188	18,232
	Hispanic (Of Any Race)	Mean	391.25	376.75	358.37	307.33	349.51	367.74	381.35	362.07
		N	187,961	186,545	184,006	186,465	175,635	173,841	176,954	156,275
	Non-Hispanic American Indian	Mean	400.93	381.82	363.39	314.52	358.00	372.62	387.55	367.96
		N	3,885	3,898	3,856	3,847	3,641	3,670	3,696	3,294
	Non-Hispanic Multiracial	Mean	407.04	388.48	369.09	325.47	366.81	378.95	394.35	375.11
		N	825	811	811	815	775	765	777	698
	Non-Hispanic White	Mean	399.93	383.05	362.50	318.65	359.40	373.11	388.57	368.85
		N	25,572	25,078	24,558	25,513	23,911	23,075	23,717	20,759
	Unknown	Mean	381.97	369.87	343.84	296.53	339.15	356.74	373.71	350.93
		N	28,402	27,881	27,684	28,415	26,292	25,542	26,136	22,601

1.2.2. Mean Scale Score Across Domain and Composite Score by Grade

This section provides parallel information to the prior section, with mean scale scores broken down by grade rather than by grade-level cluster.

Table 1.2.2.1

Mean Scale Scores by Grade, S403 Online

Grade		Listening	Reading	Writing	Speaking	Oral	Literacy	Comprehension	Overall
1	Mean	325.28	291.83	257.60	251.47	288.50	274.84	301.87	278.65
	N	181,107	183,421	189,643	177,524	169,891	183,361	175,698	165,022
2	Mean	325.54	318.59	296.62	259.67	292.72	307.72	320.68	302.95
	N	185,843	186,415	194,851	182,248	174,336	186,324	178,466	167,720
3	Mean	348.97	333.70	317.23	272.13	310.63	325.67	338.35	320.90
	N	186,821	186,956	195,426	183,619	176,031	186,906	179,511	169,474
4	Mean	411.73	354.98	336.22	307.80	360.04	345.67	372.11	349.74
	N	178,833	176,492	168,788	173,762	165,948	160,386	168,857	144,258
5	Mean	420.77	359.86	345.22	309.97	365.62	352.65	378.25	356.26
	N	144,131	142,149	137,221	139,896	134,068	130,620	136,488	118,083
6	Mean	392.60	343.10	325.33	312.55	352.79	334.34	358.12	339.68
	N	102,133	102,022	98,248	100,736	94,973	93,378	96,582	83,413
7	Mean	396.97	348.97	329.18	312.52	354.97	339.18	363.59	343.68
	N	90,916	91,277	88,369	90,104	84,402	83,957	86,076	74,786
8	Mean	402.27	355.45	333.49	313.82	358.16	344.53	369.70	348.29
	N	84,738	84,658	82,077	83,999	78,961	77,979	80,182	70,081
9	Mean	387.25	370.19	350.34	304.33	346.02	360.30	375.56	355.79
	N	97,431	96,429	95,109	97,071	91,005	89,312	91,152	80,043
10	Mean	391.72	378.29	359.39	310.91	351.49	369.11	382.56	363.60
	N	80,344	79,330	78,172	79,802	74,941	73,549	75,172	66,038
11	Mean	396.59	384.67	365.61	314.77	355.79	375.36	388.55	369.27
	N	71,529	70,173	69,211	70,198	66,181	64,874	66,590	57,988
12	Mean	396.51	385.38	365.95	315.04	355.86	375.84	388.97	369.58
	N	52,461	51,887	51,135	52,238	49,142	48,237	49,226	43,523

Table 1.2.2.2

Mean Scale Scores by Grade by Gender, S403 Online

Grade	Gender		Listening	Reading	Writing	Speaking	Oral	Literacy	Comprehension	Overall
1	F	Mean	329.05	294.05	261.61	257.70	293.49	277.94	304.51	282.27
		N	84,584	85,320	88,494	82,942	79,431	85,301	81,797	76,928
	M	Mean	322.20	289.96	254.25	246.17	284.31	272.25	299.67	275.63
		N	95,256	96,829	99,834	93,348	89,267	96,788	92,674	86,937
	Missing	Mean	305.00	286.31	241.98	233.57	269.45	264.18	291.90	265.75
		N	1,267	1,272	1,315	1,234	1,193	1,272	1,227	1,157
2	F	Mean	327.93	320.52	302.05	265.61	296.86	311.39	322.72	306.72
		N	87,052	86,861	91,129	85,340	81,760	86,826	83,326	78,432
	M	Mean	323.60	316.95	291.94	254.54	289.20	304.58	318.97	299.73
		N	97,688	98,427	102,551	95,808	91,533	98,371	94,072	88,276
	Missing	Mean	308.25	313.22	284.03	244.61	276.70	298.77	311.90	292.17
		N	1,103	1,127	1,171	1,100	1,043	1,127	1,068	1,012
3	F	Mean	350.28	336.48	323.50	278.13	314.30	330.22	340.68	325.17
		N	86,746	86,601	90,737	85,230	81,714	86,588	83,201	78,553
	M	Mean	348.06	331.36	311.91	267.09	307.66	321.84	336.44	317.34
		N	99,031	99,322	103,611	97,378	93,334	99,286	95,309	89,975
	Missing	Mean	325.87	325.56	300.49	252.34	288.21	313.42	325.72	305.15
		N	1,044	1,033	1,078	1,011	983	1,032	1,001	946
4	F	Mean	411.90	356.54	341.39	310.80	361.60	349.04	373.21	352.48
		N	82,087	80,533	77,060	79,675	76,142	72,954	77,098	65,680
	M	Mean	411.85	353.75	332.07	305.44	358.92	343.00	371.32	347.61
		N	95,678	94,924	90,689	93,062	88,822	86,456	90,766	77,702
	Missing	Mean	388.55	346.37	315.70	289.85	339.08	330.55	358.93	332.76
		N	1,068	1,035	1,039	1,025	984	976	993	876
5	F	Mean	420.59	361.89	350.79	311.92	366.50	356.51	379.61	359.20
		N	64,335	63,267	61,018	62,330	59,832	58,069	60,909	52,682
	M	Mean	421.29	358.37	341.00	308.67	365.26	349.77	377.37	354.15
		N	78,849	77,968	75,324	76,637	73,335	71,721	74,689	64,629
	Missing	Mean	390.42	345.85	319.36	285.47	337.41	332.05	358.81	332.13
		N	947	914	879	929	901	830	890	772

Grade	Gender		Listening	Reading	Writing	Speaking	Oral	Literacy	Comprehension	Overall
6	F	Mean	391.99	346.21	330.96	313.39	352.89	338.72	360.06	342.73
		N	43,894	43,632	41,979	43,060	40,694	39,896	41,428	35,656
	M	Mean	393.31	340.88	321.38	312.24	353.00	331.25	356.80	337.61
		N	57,396	57,534	55,454	56,843	53,495	52,693	54,345	47,053
	Missing	Mean	376.23	334.29	304.56	290.13	333.33	319.52	347.16	323.71
		N	843	856	815	833	784	789	809	704
7	F	Mean	397.28	352.58	335.14	314.03	355.83	343.99	366.14	347.26
		N	39,433	39,447	38,104	38,842	36,497	36,238	37,316	32,322
	M	Mean	397.02	346.36	324.89	311.71	354.63	335.70	361.81	341.18
		N	50,789	51,117	49,592	50,569	47,255	47,073	48,094	41,882
	Missing	Mean	375.43	336.65	307.15	286.43	330.95	322.23	348.52	324.64
		N	694	713	673	693	650	646	666	582
8	F	Mean	402.62	359.04	339.20	315.98	359.42	349.16	372.27	351.87
		N	36,940	36,627	35,594	36,472	34,378	33,733	34,808	30,412
	M	Mean	402.30	352.89	329.36	312.57	357.56	341.22	367.93	345.81
		N	47,159	47,378	45,873	46,889	43,982	43,651	44,759	39,125
	Missing	Mean	380.46	339.92	310.45	282.03	330.17	325.47	352.40	326.71
		N	639	653	610	638	601	595	615	544
9	F	Mean	388.92	374.48	357.73	307.63	348.43	366.18	379.05	360.54
		N	40,706	40,003	39,474	40,370	37,933	36,975	37,935	33,205
	M	Mean	386.61	367.51	345.80	302.74	344.97	356.68	373.51	353.03
		N	55,438	55,137	54,385	55,406	51,859	51,146	52,003	45,764
	Missing	Mean	361.48	351.59	314.96	269.40	315.12	333.31	354.49	326.65
		N	1,287	1,289	1,250	1,295	1,213	1,191	1,214	1,074
10	F	Mean	392.33	381.57	365.05	313.04	352.88	373.65	385.05	367.17
		N	34,741	34,091	33,476	34,323	32,361	31,495	32,453	28,408
	M	Mean	391.50	375.94	355.35	309.56	350.67	365.87	380.81	361.09
		N	44,927	44,581	44,034	44,805	41,952	41,443	42,098	37,089
	Missing	Mean	375.92	368.06	341.33	292.54	334.12	355.36	370.62	347.83
		N	676	658	662	674	628	611	621	541
11	F	Mean	396.83	387.68	370.85	317.26	357.17	379.58	390.73	372.65
		N	31,511	30,732	30,281	30,725	29,087	28,362	29,284	25,399
	M	Mean	396.75	382.45	361.91	313.16	355.05	372.33	387.03	366.91
		N	39,276	38,724	38,203	38,740	36,407	35,832	36,624	31,980
	Missing	Mean	377.79	375.63	341.45	295.73	336.18	358.93	376.70	351.81
		N	742	717	727	733	687	680	682	609
12	F	Mean	397.10	388.51	371.33	317.70	357.52	380.18	391.40	373.08
		N	23,936	23,574	23,092	23,736	22,398	21,789	22,425	19,718
	M	Mean	396.15	382.74	361.57	312.86	354.56	372.27	386.94	366.71
		N	28,119	27,931	27,651	28,101	26,365	26,086	26,438	23,481
	Missing	Mean	387.29	385.08	357.45	309.56	348.25	371.75	386.29	364.80
		N	406	382	392	401	379	362	363	324

Table 1.2.2.3

Mean Scale Scores by Grade by Ethnicity, S403 Online

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Comprehension	Overall
1	Non-Hispanic Asian	Mean	334.64	307.56	270.61	257.44	296.03	289.21	315.67	290.87
		N	25,087	25,373	26,289	24,608	23,560	25,370	24,301	22,861
	Non-Hispanic Pacific Islander	Mean	306.70	286.89	257.14	243.67	275.41	272.20	292.71	272.77
		N	1,554	1,589	1,648	1,554	1,474	1,588	1,507	1,432
	Non-Hispanic Black	Mean	323.64	293.89	258.81	261.96	293.13	276.54	302.88	281.34
		N	9,769	9,948	10,311	9,598	9,123	9,943	9,471	8,849
	Hispanic (Of Any Race)	Mean	323.06	287.60	254.16	248.61	285.98	271.02	298.27	275.27
		N	112,427	113,988	117,550	110,027	105,409	113,945	109,305	102,591
	Non-Hispanic American Indian	Mean	324.31	284.66	254.36	249.40	287.04	269.67	296.52	274.47
		N	1,545	1,552	1,621	1,518	1,451	1,552	1,481	1,400
	Non-Hispanic Multiracial	Mean	341.25	303.77	266.58	263.50	302.48	285.21	315.13	290.28
		N	832	837	871	831	793	837	801	764
	Non-Hispanic White	Mean	333.12	297.04	263.55	260.03	296.69	280.35	307.79	284.92
		N	19,030	19,234	20,025	18,779	17,887	19,231	18,348	17,275
	Unknown	Mean	315.93	289.27	251.28	243.13	279.62	270.37	297.23	272.70
		N	10,863	10,900	11,328	10,609	10,194	10,895	10,484	9,850
2	Non-Hispanic Asian	Mean	340.58	330.86	309.94	266.61	303.65	320.55	333.77	315.18
		N	23,430	23,459	24,522	22,950	22,004	23,448	22,512	21,204
	Non-Hispanic Pacific Islander	Mean	306.93	311.37	295.56	250.19	278.42	303.49	310.25	295.73
		N	1,524	1,536	1,604	1,525	1,453	1,536	1,470	1,403
	Non-Hispanic Black	Mean	325.45	318.94	296.21	267.76	296.95	307.72	320.92	304.44
		N	10,136	10,167	10,709	9,949	9,455	10,163	9,676	9,044
	Hispanic (Of Any Race)	Mean	321.70	315.71	293.71	257.22	289.57	304.83	317.51	299.98
		N	117,962	118,424	123,591	115,675	110,722	118,362	113,400	106,609
	Non-Hispanic American Indian	Mean	321.53	311.99	290.14	254.03	287.74	301.20	314.81	296.77
		N	1,603	1,649	1,715	1,611	1,510	1,648	1,545	1,455
	Non-Hispanic Multiracial	Mean	340.66	327.08	302.28	266.41	303.92	314.93	331.25	311.49
		N	819	824	853	802	773	824	794	752
	Non-Hispanic White	Mean	336.66	323.78	302.49	267.75	302.37	313.22	327.59	309.70
		N	19,387	19,376	20,358	19,018	18,163	19,366	18,522	17,392
	Unknown	Mean	317.12	315.34	290.19	251.03	284.16	302.84	315.91	296.84
		N	10,982	10,980	11,499	10,718	10,256	10,977	10,547	9,861

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
3	Non-Hispanic Asian	Mean	361.82	346.76	329.18	278.41	320.04	338.22	351.30	332.39
		N	21,312	21,386	22,275	20,912	20,070	21,382	20,561	19,411
	Non-Hispanic Pacific Islander	Mean	333.77	326.15	317.39	263.62	298.68	321.93	328.68	314.90
		N	1,510	1,527	1,590	1,500	1,429	1,527	1,457	1,382
	Non-Hispanic Black	Mean	345.38	332.56	314.31	276.97	311.13	323.69	336.53	319.51
		N	9,912	9,946	10,456	9,719	9,250	9,945	9,505	8,891
	Hispanic (Of Any Race)	Mean	346.82	331.18	315.46	270.50	308.78	323.53	335.95	318.89
		N	122,056	122,104	127,523	119,894	115,058	122,066	117,321	110,805
	Non-Hispanic American Indian	Mean	345.43	324.75	310.52	267.14	306.09	317.54	330.81	313.78
		N	1,606	1,612	1,681	1,584	1,517	1,612	1,543	1,457
	Non-Hispanic Multiracial	Mean	363.50	343.34	323.04	280.24	322.51	333.54	350.03	330.83
		N	794	789	822	783	759	789	764	731
	Non-Hispanic White	Mean	358.27	339.89	322.16	279.50	319.07	331.24	345.46	327.36
		N	18,218	18,262	19,209	18,078	17,201	18,259	17,408	16,479
	Unknown	Mean	337.71	328.82	308.89	263.03	300.34	318.96	331.51	312.94
		N	11,413	11,330	11,870	11,149	10,747	11,326	10,952	10,318
4	Non-Hispanic Asian	Mean	422.27	366.77	346.15	315.00	368.85	356.60	383.54	359.92
		N	18,184	18,008	17,256	17,638	16,929	16,522	17,313	14,984
	Non-Hispanic Pacific Islander	Mean	398.68	347.77	333.72	300.36	349.95	340.40	363.00	342.78
		N	1,276	1,279	1,200	1,274	1,198	1,148	1,199	1,027
	Non-Hispanic Black	Mean	410.05	352.50	330.66	313.73	362.03	341.66	369.89	347.77
		N	9,317	9,207	8,857	9,039	8,549	8,324	8,739	7,371
	Hispanic (Of Any Race)	Mean	411.10	353.68	336.13	306.71	359.19	345.00	371.02	349.04
		N	118,526	116,996	111,693	115,240	110,137	106,313	112,010	95,812
	Non-Hispanic American Indian	Mean	407.23	348.45	329.22	304.67	356.54	339.18	366.35	344.34
		N	1,747	1,748	1,637	1,677	1,606	1,566	1,672	1,389
	Non-Hispanic Multiracial	Mean	419.78	359.88	337.39	312.51	365.69	348.87	377.87	353.19
		N	605	584	572	595	562	534	557	482
	Non-Hispanic White	Mean	417.48	358.89	340.64	314.60	366.32	349.85	376.54	354.62
		N	16,526	16,196	15,002	16,032	15,272	14,144	15,450	12,635
	Unknown	Mean	397.82	348.27	323.17	295.49	346.79	335.45	363.13	338.30
		N	12,652	12,474	12,571	12,267	11,695	11,835	11,917	10,558

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Comprehension	Overall
5	Non-Hispanic Asian	Mean	428.13	370.79	353.03	316.39	372.48	362.07	388.06	364.91
		N	12,661	12,544	12,135	12,341	11,868	11,628	12,095	10,627
	Non-Hispanic Pacific Islander	Mean	407.94	353.31	345.60	303.46	355.92	349.52	369.56	351.15
		N	1,224	1,216	1,134	1,217	1,160	1,087	1,165	995
	Non-Hispanic Black	Mean	417.32	356.04	337.31	315.71	366.63	346.79	374.63	352.49
		N	7,720	7,646	7,344	7,502	7,134	6,987	7,298	6,241
	Hispanic (Of Any Race)	Mean	421.39	359.29	346.13	309.21	365.60	352.88	378.07	356.46
		N	97,131	95,772	92,327	94,174	90,334	87,973	92,018	79,587
	Non-Hispanic American Indian	Mean	418.62	356.42	343.37	308.37	363.78	350.20	374.94	353.85
		N	1,683	1,685	1,617	1,623	1,555	1,550	1,621	1,398
	Non-Hispanic Multiracial	Mean	430.76	369.44	347.43	320.46	375.53	358.01	387.78	362.20
		N	474	463	450	455	441	429	449	391
	Non-Hispanic White	Mean	425.34	363.40	347.82	317.44	371.74	355.74	382.08	360.32
		N	12,316	12,069	11,455	11,922	11,372	10,794	11,522	9,667
	Unknown	Mean	405.36	351.71	331.27	297.34	351.24	341.19	367.71	343.51
		N	10,922	10,754	10,759	10,662	10,204	10,172	10,320	9,177
6	Non-Hispanic Asian	Mean	401.72	353.74	335.01	320.31	361.24	344.69	368.39	349.55
		N	8,584	8,559	8,273	8,390	7,959	7,899	8,134	7,078
	Non-Hispanic Pacific Islander	Mean	382.72	337.64	322.48	306.98	344.51	330.51	351.01	334.87
		N	864	858	789	876	802	730	806	643
	Non-Hispanic Black	Mean	389.90	339.06	316.83	319.05	354.66	328.04	354.47	335.99
		N	5,921	5,852	5,647	5,780	5,433	5,298	5,522	4,672
	Hispanic (Of Any Race)	Mean	391.95	342.56	326.00	311.82	352.10	334.36	357.54	339.46
		N	67,244	67,333	64,750	66,414	62,692	61,717	63,794	55,216
	Non-Hispanic American Indian	Mean	391.84	340.67	325.78	312.64	352.31	333.35	355.99	338.63
		N	1,314	1,304	1,274	1,273	1,209	1,211	1,252	1,070
	Non-Hispanic Multiracial	Mean	405.68	349.96	330.21	324.43	364.96	340.98	367.14	347.93
		N	310	301	304	302	290	283	289	259
	Non-Hispanic White	Mean	399.47	347.60	329.66	320.47	360.49	338.93	363.34	345.38
		N	8,586	8,464	8,253	8,449	7,953	7,748	8,007	6,916
	Unknown	Mean	384.88	336.38	313.01	299.54	342.16	324.57	350.99	329.41
		N	9,310	9,351	8,958	9,252	8,635	8,492	8,778	7,559

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Comprehension	Overall
7	Non-Hispanic Asian	Mean	410.95	364.26	343.71	327.23	369.40	354.21	378.40	358.47
		N	8,413	8,342	8,076	8,266	7,815	7,639	7,923	6,864
	Non-Hispanic Pacific Islander	Mean	384.25	340.35	324.07	303.55	344.16	333.00	353.93	336.66
		N	775	770	748	782	717	682	717	599
	Non-Hispanic Black	Mean	396.90	346.66	323.41	322.20	359.72	335.17	361.90	342.36
		N	5,802	5,786	5,581	5,755	5,353	5,253	5,408	4,615
	Hispanic (Of Any Race)	Mean	395.48	347.63	329.03	310.54	353.25	338.42	362.21	342.64
		N	57,736	58,164	56,360	57,345	53,751	53,753	54,916	47,981
	Non-Hispanic American Indian	Mean	399.65	348.63	330.87	319.04	359.59	340.09	364.37	346.44
		N	1,157	1,188	1,143	1,162	1,084	1,103	1,126	994
	Non-Hispanic Multiracial	Mean	411.20	360.57	338.90	323.44	367.69	349.69	375.80	354.28
		N	262	265	258	263	248	247	254	226
	Non-Hispanic White	Mean	405.51	354.41	334.43	321.91	363.95	344.64	370.05	350.41
		N	8,128	8,041	7,824	7,990	7,483	7,348	7,578	6,523
	Unknown	Mean	385.64	340.25	315.00	295.80	340.67	327.50	353.93	330.73
		N	8,643	8,721	8,379	8,541	7,951	7,932	8,154	6,984
8	Non-Hispanic Asian	Mean	416.30	370.75	348.74	331.30	373.97	360.08	384.55	364.07
		N	7,888	7,799	7,491	7,705	7,341	7,127	7,445	6,475
	Non-Hispanic Pacific Islander	Mean	394.21	349.39	330.48	308.42	351.45	339.46	363.22	342.87
		N	681	697	646	695	627	608	648	530
	Non-Hispanic Black	Mean	402.92	353.96	329.71	325.60	364.29	341.97	368.97	348.41
		N	5,807	5,715	5,521	5,734	5,369	5,204	5,434	4,665
	Hispanic (Of Any Race)	Mean	401.10	354.47	333.46	311.70	356.55	344.03	368.64	347.45
		N	53,881	53,908	52,466	53,441	50,316	49,931	51,139	44,923
	Non-Hispanic American Indian	Mean	402.51	353.20	332.95	317.64	359.91	343.07	368.16	348.17
		N	1,078	1,088	1,089	1,107	1,024	1,024	1,020	924
	Non-Hispanic Multiracial	Mean	409.97	362.51	337.93	323.91	366.76	350.52	376.76	355.01
		N	259	258	254	255	247	244	250	227
	Non-Hispanic White	Mean	411.00	361.50	339.08	323.22	367.29	350.44	376.70	355.27
		N	7,217	7,258	6,999	7,142	6,655	6,636	6,794	5,886
	Unknown	Mean	388.28	343.27	316.43	293.68	340.70	329.48	356.79	332.09
		N	7,927	7,935	7,611	7,920	7,382	7,205	7,452	6,451

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
9	Non-Hispanic Asian	Mean	403.20	388.11	370.21	327.15	365.02	379.31	392.89	374.72
		N	8,439	8,195	8,026	8,230	7,802	7,510	7,831	6,779
	Non-Hispanic Pacific Islander	Mean	383.05	368.75	357.73	300.56	341.85	363.09	373.37	356.50
		N	859	849	799	848	785	737	786	644
	Non-Hispanic Black	Mean	386.45	372.67	350.90	314.53	350.56	361.97	377.01	358.14
		N	6,609	6,398	6,398	6,562	6,149	5,901	6,047	5,254
	Hispanic (Of Any Race)	Mean	386.25	368.53	349.95	301.92	344.40	359.32	374.10	354.69
		N	62,141	61,685	60,985	61,937	58,167	57,430	58,384	51,574
	Non-Hispanic American Indian	Mean	395.47	373.13	356.21	308.76	352.00	363.98	379.60	359.69
		N	1,277	1,245	1,255	1,270	1,207	1,163	1,184	1,055
	Non-Hispanic Multiracial	Mean	399.81	380.51	362.18	317.22	359.14	371.20	386.33	367.62
		N	284	279	271	271	262	257	269	235
	Non-Hispanic White	Mean	395.24	376.51	356.67	314.84	355.08	366.80	382.34	362.92
		N	7,959	7,840	7,626	7,984	7,450	7,176	7,384	6,437
	Unknown	Mean	372.89	358.56	329.45	284.74	328.91	343.66	363.14	338.75
		N	9,863	9,938	9,749	9,969	9,183	9,138	9,267	8,065
10	Non-Hispanic Asian	Mean	402.60	391.19	373.43	328.98	365.92	382.41	394.78	377.09
		N	7,813	7,598	7,516	7,663	7,208	7,013	7,226	6,271
	Non-Hispanic Pacific Islander	Mean	389.97	378.97	365.62	307.32	349.70	372.34	381.97	364.96
		N	578	559	556	581	529	500	516	436
	Non-Hispanic Black	Mean	387.78	378.28	358.25	318.17	352.97	368.85	381.39	364.04
		N	6,004	5,827	5,744	5,933	5,558	5,327	5,523	4,753
	Hispanic (Of Any Race)	Mean	390.36	376.62	358.51	307.60	349.22	367.85	380.97	362.09
		N	50,985	50,632	49,813	50,578	47,695	47,189	48,086	42,523
	Non-Hispanic American Indian	Mean	402.23	382.38	365.17	317.50	360.18	373.73	388.26	369.71
		N	1,004	1,014	1,008	998	940	962	959	862
	Non-Hispanic Multiracial	Mean	410.85	389.64	367.72	327.77	369.52	378.72	396.23	375.63
		N	208	210	213	214	195	199	194	173
	Non-Hispanic White	Mean	400.02	382.84	362.77	319.78	360.04	373.20	388.58	368.99
		N	6,644	6,557	6,412	6,724	6,243	6,024	6,178	5,419
	Unknown	Mean	383.21	371.06	346.62	299.48	341.07	358.89	374.79	352.93
		N	7,108	6,933	6,910	7,111	6,573	6,335	6,490	5,601

Grade	Ethnicity		Listening	Reading	Writing	Speaking	Oral	Literacy	Compre-hension	Overall
11	Non-Hispanic Asian	Mean	406.81	396.52	378.54	332.63	369.88	387.72	399.89	382.26
		N	7,402	7,189	7,083	7,199	6,830	6,644	6,863	5,981
	Non-Hispanic Pacific Islander	Mean	396.02	385.89	372.11	310.90	353.92	378.80	389.18	371.32
		N	534	524	498	515	480	457	485	396
	Non-Hispanic Black	Mean	389.52	382.00	362.78	319.83	354.50	372.57	384.43	366.72
		N	5,913	5,687	5,601	5,799	5,422	5,144	5,368	4,563
	Hispanic (Of Any Race)	Mean	396.06	383.46	365.27	311.84	354.09	374.54	387.51	368.16
		N	44,015	43,513	42,893	43,269	40,837	40,423	41,301	36,151
	Non-Hispanic American Indian	Mean	404.13	388.42	368.57	317.47	361.55	378.89	393.21	373.26
		N	914	941	915	903	848	888	885	786
	Non-Hispanic Multiracial	Mean	413.68	395.06	374.17	330.67	373.25	385.18	401.50	381.60
		N	204	196	200	201	195	188	191	178
	Non-Hispanic White	Mean	401.65	387.03	366.17	320.29	361.04	377.02	391.90	372.27
		N	6,151	5,960	5,879	6,011	5,697	5,515	5,676	4,950
	Unknown	Mean	388.51	378.58	353.87	303.94	346.01	366.35	381.90	359.83
		N	6,396	6,163	6,142	6,301	5,872	5,615	5,821	4,983
12	Non-Hispanic Asian	Mean	402.32	394.40	377.82	330.37	366.36	386.26	396.87	379.88
		N	5,990	5,907	5,715	5,926	5,615	5,425	5,636	4,942
	Non-Hispanic Pacific Islander	Mean	394.51	382.57	369.14	308.74	351.77	376.15	387.07	369.23
		N	355	365	360	379	334	331	329	284
	Non-Hispanic Black	Mean	389.61	384.51	363.76	321.89	355.52	374.09	386.12	368.07
		N	4,624	4,508	4,416	4,619	4,303	4,090	4,250	3,662
	Hispanic (Of Any Race)	Mean	395.95	383.95	365.34	311.43	353.81	374.79	387.77	368.24
		N	30,820	30,715	30,315	30,681	28,936	28,799	29,183	26,027
	Non-Hispanic American Indian	Mean	404.93	387.60	367.04	317.01	361.39	377.83	393.14	373.14
		N	690	698	678	676	646	657	668	591
	Non-Hispanic Multiracial	Mean	406.33	393.96	378.09	330.88	368.63	386.11	397.80	379.74
		N	129	126	127	129	123	121	123	112
	Non-Hispanic White	Mean	405.35	389.18	367.06	321.37	363.57	378.44	394.63	374.01
		N	4,818	4,721	4,641	4,794	4,521	4,360	4,479	3,953
	Unknown	Mean	389.69	380.26	356.05	306.43	347.96	368.41	383.20	361.74
		N	5,035	4,847	4,883	5,034	4,664	4,454	4,558	3,952

1.2.3. Correlations

Tables in this section show Pearson correlations among the four domain scale scores by grade-level cluster across all tiers, as well as the number of students included in each correlation.

Table 1.2.3.1
Correlations Among Scale Scores: Grade 1, S403 Online

		Listening	**Reading**	**Writing**	**Speaking**
Listening	Pearson Correlation	1	.395	.514	.484
	N	181,107	175,698	181,045	169,891
Reading	Pearson Correlation		1	.509	.290
	N			183,361	172,010
Writing	Pearson Correlation			1	.400
	N				177,466
Speaking	Pearson Correlation				1
	N				177,524

Table 1.2.3.2
Correlations Among Scale Scores: Grades 2–3, S403 Online

		Listening	**Reading**	**Writing**	**Speaking**
Listening	Pearson Correlation	1	.595	.625	.550
	N	372,664	357,977	372,521	350,367
Reading	Pearson Correlation		1	.647	.462
	N		373,371	373,230	350,946
Writing	Pearson Correlation			1	.543
	N				365,728
Speaking	Pearson Correlation				1
	N				365,867

Table 1.2.3.3
Correlations Among Scale Scores: Grades 4–5, S403 Online

		Listening	**Reading**	**Writing**	**Speaking**
Listening	Pearson Correlation	1	.656	.633	.553
	N	322,964	305,345	293,093	300,016
Reading	Pearson Correlation		1	.636	.480
	N			291,006	296,270
Writing	Pearson Correlation			1	.550
	N				285,110
Speaking	Pearson Correlation				1
	N				313,658

Table 1.2.3.4

Correlations Among Scale Scores: Grades 6–8, S403 Online

		Listening	**Reading**	**Writing**	**Speaking**
Listening	Pearson Correlation	1	.681	.648	.578
	N	277,787	262,840	254,117	258,336
Reading	Pearson Correlation		1	.758	.538
	N			255,314	258,714
Writing	Pearson Correlation			1	.610
	N				250,833
Speaking	Pearson Correlation				1
	N				274,839

Table 1.2.3.5

Correlations Among Scale Scores: Grades 9–12, S403 Online

		Listening	**Reading**	**Writing**	**Speaking**
Listening	Pearson Correlation	1	.690	.524	.547
	N	301,765	282,140	277,444	281,269
Reading	Pearson Correlation		1	.634	.586
	N			275,972	278,198
Writing	Pearson Correlation			1	.608
	N				274,232
Speaking	Pearson Correlation				1
	N				299,309

1.3. Proficiency Level Results

1.3.1. Domains

1.3.1.1. Listening

1.3.1.1.1. By Cluster

Table 1.3.1.1.1.1

Proficiency Level by Cluster (Count): Listening, S403 Online

Cluster	Listening Proficiency Range						Total
	1	2	3	4	5	6	
1	16,004	8,117	16,930	10,511	25,558	103,987	181,107
2–3	35,537	33,822	69,544	25,884	66,331	141,546	372,664
4–5	8,807	5,306	16,593	13,910	36,476	241,872	322,964
6–8	11,785	16,185	49,379	46,169	49,646	104,623	277,787
9–12	35,658	38,341	73,155	57,976	41,236	55,399	301,765

Table 1.3.1.1.1.2

Proficiency Level by Cluster (Percent): Listening, S403 Online

Cluster	Listening Proficiency Range						Total
	1	2	3	4	5	6	
1	8.8%	4.5%	9.3%	5.8%	14.1%	57.4%	100.0%
2–3	9.5%	9.1%	18.7%	6.9%	17.8%	38.0%	100.0%
4–5	2.7%	1.6%	5.1%	4.3%	11.3%	74.9%	100.0%
6–8	4.2%	5.8%	17.8%	16.6%	17.9%	37.7%	100.0%
9–12	11.8%	12.7%	24.2%	19.2%	13.7%	18.4%	100.0%

1.3.1.1.2. By Grade

Table 1.3.1.1.2.1

Proficiency Level by Grade (Count): Listening, S403 Online

Grade	Listening Proficiency Range						Total
	1	2	3	4	5	6	
1	16,004	8,117	16,930	10,511	25,558	103,987	181,107
2	19,280	17,281	39,292	14,007	34,505	61,478	185,843
3	16,257	16,541	30,252	11,877	31,826	80,068	186,821
4	4,149	2,530	9,527	7,668	20,003	134,956	178,833
5	4,658	2,776	7,066	6,242	16,473	106,916	144,131
6	2,982	4,475	16,732	16,370	22,048	39,526	102,133
7	3,934	5,261	17,538	16,225	15,114	32,844	90,916
8	4,869	6,449	15,109	13,574	12,484	32,253	84,738
9	7,598	14,155	24,186	19,509	13,233	18,750	97,431
10	9,665	10,525	18,780	15,406	10,491	15,477	80,344
11	10,424	6,931	18,215	12,864	10,669	12,426	71,529
12	7,971	6,730	11,974	10,197	6,843	8,746	52,461

Table 1.3.1.1.1.2

Proficiency Level by Grade (Percent): Listening, S403 Online

Grade	Listening Proficiency Range						Total
	1	2	3	4	5	6	
1	8.8%	4.5%	9.3%	5.8%	14.1%	57.4%	100.0%
2	10.4%	9.3%	21.1%	7.5%	18.6%	33.1%	100.0%
3	8.7%	8.9%	16.2%	6.4%	17.0%	42.9%	100.0%
4	2.3%	1.4%	5.3%	4.3%	11.2%	75.5%	100.0%
5	3.2%	1.9%	4.9%	4.3%	11.4%	74.2%	100.0%
6	2.9%	4.4%	16.4%	16.0%	21.6%	38.7%	100.0%
7	4.3%	5.8%	19.3%	17.8%	16.6%	36.1%	100.0%
8	5.7%	7.6%	17.8%	16.0%	14.7%	38.1%	100.0%
9	7.8%	14.5%	24.8%	20.0%	13.6%	19.2%	100.0%
10	12.0%	13.1%	23.4%	19.2%	13.1%	19.3%	100.0%
11	14.6%	9.7%	25.5%	18.0%	14.9%	17.4%	100.0%
12	15.2%	12.8%	22.8%	19.4%	13.0%	16.7%	100.0%

1.3.1.2. Reading

1.3.1.2.1. By Cluster

Table 1.3.1.2.1.1

Proficiency Level by Cluster (Count): Reading, S403 Online

Cluster	Reading Proficiency Range						Total
	1	2	3	4	5	6	
1	33,816	49,853	38,074	21,768	19,998	19,912	183,421
2–3	50,565	87,129	81,442	38,959	69,149	46,127	373,371
4–5	30,983	59,147	61,767	33,323	65,877	67,544	318,641
6–8	88,895	62,784	62,601	19,649	27,489	16,539	277,957
9–12	65,145	76,809	58,712	22,702	37,728	36,723	297,819

Table 1.3.1.2.1.2

Proficiency Level by Cluster (Percent): Reading, S403 Online

Cluster	Reading Proficiency Range						Total
	1	2	3	4	5	6	
1	18.4%	27.2%	20.8%	11.9%	10.9%	10.9%	100.0%
2–3	13.5%	23.3%	21.8%	10.4%	18.5%	12.4%	100.0%
4–5	9.7%	18.6%	19.4%	10.5%	20.7%	21.2%	100.0%
6–8	32.0%	22.6%	22.5%	7.1%	9.9%	6.0%	100.0%
9–12	21.9%	25.8%	19.7%	7.6%	12.7%	12.3%	100.0%

1.3.1.2.2. By Grade

Table 1.3.1.2.2.1

Proficiency Level by Grade (Count): Reading, S403 Online

Grade	Reading Proficiency Range						Total
	1	2	3	4	5	6	
1	33,816	49,853	38,074	21,768	19,998	19,912	183,421
2	22,892	42,882	45,239	21,205	34,136	20,061	186,415
3	27,673	44,247	36,203	17,754	35,013	26,066	186,956
4	14,447	30,894	33,552	21,406	38,129	38,064	176,492
5	16,536	28,253	28,215	11,917	27,748	29,480	142,149
6	31,825	23,449	24,836	7,856	9,764	4,292	102,022
7	29,924	19,833	21,103	6,360	7,596	6,461	91,277
8	27,146	19,502	16,662	5,433	10,129	5,786	84,658
9	23,743	24,163	19,044	6,243	11,403	11,833	96,429
10	16,769	19,604	16,490	6,432	11,284	8,751	79,330
11	13,656	18,573	12,974	5,751	8,817	10,402	70,173
12	10,977	14,469	10,204	4,276	6,224	5,737	51,887

Table 1.3.1.2.2.2

Proficiency Level by Grade (Percent): Reading, S403 Online

Grade	Reading Proficiency Range						Total
	1	2	3	4	5	6	
1	18.4%	27.2%	20.8%	11.9%	10.9%	10.9%	100.0%
2	12.3%	23.0%	24.3%	11.4%	18.3%	10.8%	100.0%
3	14.8%	23.7%	19.4%	9.5%	18.7%	13.9%	100.0%
4	8.2%	17.5%	19.0%	12.1%	21.6%	21.6%	100.0%
5	11.6%	19.9%	19.8%	8.4%	19.5%	20.7%	100.0%
6	31.2%	23.0%	24.3%	7.7%	9.6%	4.2%	100.0%
7	32.8%	21.7%	23.1%	7.0%	8.3%	7.1%	100.0%
8	32.1%	23.0%	19.7%	6.4%	12.0%	6.8%	100.0%
9	24.6%	25.1%	19.7%	6.5%	11.8%	12.3%	100.0%
10	21.1%	24.7%	20.8%	8.1%	14.2%	11.0%	100.0%
11	19.5%	26.5%	18.5%	8.2%	12.6%	14.8%	100.0%
12	21.2%	27.9%	19.7%	8.2%	12.0%	11.1%	100.0%

1.3.1.3. Writing

1.3.1.3.1. By Cluster

Table 1.3.1.3.1.1

Proficiency Level by Cluster (Count): Writing, S403 Online

	Writing Proficiency Range						
Cluster	**1**	**2**	**3**	**4**	**5**	**6**	**Total**
1	36,560	112,912	38,660	1,506	4	1	189,643
2–3	23,372	59,606	243,975	61,999	1,255	70	390,277
4–5	12,733	18,693	136,693	132,042	5,625	223	306,009
6–8	30,296	44,238	131,148	62,455	526	31	268,694
9–12	28,498	39,382	137,791	85,543	2,402	11	293,627

Table 1.3.1.3.1.2

Proficiency Level by Cluster (Percent): Writing, S403 Online

	Writing Proficiency Range						
Cluster	**1**	**2**	**3**	**4**	**5**	**6**	**Total**
1	19.3%	59.5%	20.4%	0.8%	0.0%	0.0%	100.0%
2–3	6.0%	15.3%	62.5%	15.9%	0.3%	0.0%	100.0%
4–5	4.2%	6.1%	44.7%	43.1%	1.8%	0.1%	100.0%
6–8	11.3%	16.5%	48.8%	23.2%	0.2%	0.0%	100.0%
9–12	9.7%	13.4%	46.9%	29.1%	0.8%	0.0%	100.0%

1.3.1.3.2. By Grade

Table 1.3.1.3.2.1

Proficiency Level by Grade (Count): Writing, S403 Online

Grade	Writing Proficiency Range						Total
	1	2	3	4	5	6	
1	36,560	112,912	38,660	1,506	4	1	189,643
2	14,021	35,231	126,900	18,503	189	7	194,851
3	9,351	24,375	117,075	43,496	1,066	63	195,426
4	7,489	9,171	79,945	69,676	2,377	130	168,788
5	5,244	9,522	56,748	62,366	3,248	93	137,221
6	8,977	15,302	51,861	21,970	132	6	98,248
7	10,063	16,580	39,902	21,651	162	11	88,369
8	11,256	12,356	39,385	18,834	232	14	82,077
9	9,222	13,620	43,100	28,138	1,020	9	95,109
10	6,123	10,043	37,800	23,357	847	2	78,172
11	6,623	8,327	33,252	20,638	371	0	69,211
12	6,530	7,392	23,639	13,410	164	0	51,135

Table 1.3.1.3.2.2

Proficiency Level by Grade (Percent): Writing, S403 Online

Grade	Writing Proficiency Range						Total
	1	2	3	4	5	6	
1	19.3%	59.5%	20.4%	0.8%	0.0%	0.0%	100.0%
2	7.2%	18.1%	65.1%	9.5%	0.1%	0.0%	100.0%
3	4.8%	12.5%	59.9%	22.3%	0.5%	0.0%	100.0%
4	4.4%	5.4%	47.4%	41.3%	1.4%	0.1%	100.0%
5	3.8%	6.9%	41.4%	45.4%	2.4%	0.1%	100.0%
6	9.1%	15.6%	52.8%	22.4%	0.1%	0.0%	100.0%
7	11.4%	18.8%	45.2%	24.5%	0.2%	0.0%	100.0%
8	13.7%	15.1%	48.0%	22.9%	0.3%	0.0%	100.0%
9	9.7%	14.3%	45.3%	29.6%	1.1%	0.0%	100.0%
10	7.8%	12.8%	48.4%	29.9%	1.1%	0.0%	100.0%
11	9.6%	12.0%	48.0%	29.8%	0.5%	0.0%	100.0%
12	12.8%	14.5%	46.2%	26.2%	0.3%	0.0%	100.0%

1.3.1.4. Speaking

1.3.1.4.1. By Cluster

Table 1.3.1.4.1.1

Proficiency Level by Cluster (Count): Speaking, S403 Online

Cluster	Speaking Proficiency Range						Total
	1	**2**	**3**	**4**	**5**	**6**	
1	26,620	73,520	59,638	17,186	541	19	177,524
2–3	58,072	158,461	127,360	21,221	645	108	365,867
4–5	27,831	72,872	140,802	68,506	3,501	146	313,658
6–8	50,672	73,011	114,987	35,337	774	58	274,839
9–12	91,818	98,331	102,557	6,503	86	14	299,309

Table 1.3.1.4.1.2

Proficiency Level by Cluster (Percent): Speaking, S403 Online

Cluster	Speaking Proficiency Range						Total
	1	**2**	**3**	**4**	**5**	**6**	
1	15.0%	41.4%	33.6%	9.7%	0.3%	0.0%	100.0%
2–3	15.9%	43.3%	34.8%	5.8%	0.2%	0.0%	100.0%
4–5	8.9%	23.2%	44.9%	21.8%	1.1%	0.0%	100.0%
6–8	18.4%	26.6%	41.8%	12.9%	0.3%	0.0%	100.0%
9–12	30.7%	32.9%	34.3%	2.2%	0.0%	0.0%	100.0%

1.3.1.4.2. By Grade

Table 1.3.1.4.2.1

Proficiency Level by Grade (Count): Speaking, S403 Online

Grade	Speaking Proficiency Range						Total
	1	2	3	4	5	6	
1	26,620	73,520	59,638	17,186	541	19	177,524
2	26,963	80,187	64,375	10,298	395	30	182,248
3	31,109	78,274	62,985	10,923	250	78	183,619
4	12,558	36,423	80,068	42,259	2,341	113	173,762
5	15,273	36,449	60,734	26,247	1,160	33	139,896
6	13,808	26,880	43,681	16,139	201	27	100,736
7	15,902	27,993	35,012	10,854	335	8	90,104
8	20,962	18,138	36,294	8,344	238	23	83,999
9	30,243	33,120	30,542	3,128	37	1	97,071
10	22,764	29,952	25,407	1,661	16	2	79,802
11	22,060	20,875	26,285	955	19	4	70,198
12	16,751	14,384	20,323	759	14	7	52,238

Table 1.3.1.2.1.2

Proficiency Level by Grade (Percent): Speaking, S403 Online

Grade	Speaking Proficiency Range						Total
	1	2	3	4	5	6	
1	15.0%	41.4%	33.6%	9.7%	0.3%	0.0%	100.0%
2	14.8%	44.0%	35.3%	5.7%	0.2%	0.0%	100.0%
3	16.9%	42.6%	34.3%	5.9%	0.1%	0.0%	100.0%
4	7.2%	21.0%	46.1%	24.3%	1.3%	0.1%	100.0%
5	10.9%	26.1%	43.4%	18.8%	0.8%	0.0%	100.0%
6	13.7%	26.7%	43.4%	16.0%	0.2%	0.0%	100.0%
7	17.6%	31.1%	38.9%	12.0%	0.4%	0.0%	100.0%
8	25.0%	21.6%	43.2%	9.9%	0.3%	0.0%	100.0%
9	31.2%	34.1%	31.5%	3.2%	0.0%	0.0%	100.0%
10	28.5%	37.5%	31.8%	2.1%	0.0%	0.0%	100.0%
11	31.4%	29.7%	37.4%	1.4%	0.0%	0.0%	100.0%
12	32.1%	27.5%	38.9%	1.5%	0.0%	0.0%	100.0%

1.3.2. Composites

1.3.2.1. Oral Composite

1.3.2.1.1. By Cluster

Table 1.3.2.1.1.1

Proficiency Level by Cluster (Count): Oral, S403 Online

Cluster	Oral Language Proficiency Range						Total
	1	2	3	4	5	6	
1	16,268	20,630	56,868	48,739	24,819	2,567	169,891
2–3	33,932	66,879	131,052	98,836	19,089	579	350,367
4–5	11,453	14,509	58,346	116,889	80,314	18,505	300,016
6–8	21,611	32,862	94,752	87,676	18,726	2,709	258,336
9–12	51,701	62,090	112,882	47,743	6,177	676	281,269

Table 1.3.2.1.1.2

Proficiency Level by Cluster (Percent): Oral, S403 Online

Cluster	Oral Language Proficiency Range						Total
	1	2	3	4	5	6	
1	9.6%	12.1%	33.5%	28.7%	14.6%	1.5%	100.0%
2–3	9.7%	19.1%	37.4%	28.2%	5.4%	0.2%	100.0%
4–5	3.8%	4.8%	19.4%	39.0%	26.8%	6.2%	100.0%
6–8	8.4%	12.7%	36.7%	33.9%	7.2%	1.0%	100.0%
9–12	18.4%	22.1%	40.1%	17.0%	2.2%	0.2%	100.0%

1.3.2.1.2. By Grade

Table 1.3.2.1.2.1

Proficiency Level by Grade (Count): Oral, S403 Online

Grade	Oral Language Proficiency Range						Total
	1	2	3	4	5	6	
1	16,268	20,630	56,868	48,739	24,819	2,567	169,891
2	17,801	37,002	65,597	43,223	10,297	416	174,336
3	16,131	29,877	65,455	55,613	8,792	163	176,031
4	5,431	7,476	31,198	63,240	47,417	11,186	165,948
5	6,022	7,033	27,148	53,649	32,897	7,319	134,068
6	5,443	10,010	36,443	35,147	7,150	780	94,973
7	7,387	11,217	31,350	27,528	5,924	996	84,402
8	8,781	11,635	26,959	25,001	5,652	933	78,961
9	15,570	21,438	35,073	16,559	2,123	242	91,005
10	13,268	16,281	30,334	13,051	1,798	209	74,941
11	12,537	14,164	26,569	11,222	1,535	154	66,181
12	10,326	10,207	20,906	6,911	721	71	49,142

Table 1.3.2.1.1.2

Proficiency Level by Grade (Percent): Oral, S403 Online

Grade	Oral Language Proficiency Range						Total
	1	2	3	4	5	6	
1	9.6%	12.1%	33.5%	28.7%	14.6%	1.5%	100.0%
2	10.2%	21.2%	37.6%	24.8%	5.9%	0.2%	100.0%
3	9.2%	17.0%	37.2%	31.6%	5.0%	0.1%	100.0%
4	3.3%	4.5%	18.8%	38.1%	28.6%	6.7%	100.0%
5	4.5%	5.2%	20.2%	40.0%	24.5%	5.5%	100.0%
6	5.7%	10.5%	38.4%	37.0%	7.5%	0.8%	100.0%
7	8.8%	13.3%	37.1%	32.6%	7.0%	1.2%	100.0%
8	11.1%	14.7%	34.1%	31.7%	7.2%	1.2%	100.0%
9	17.1%	23.6%	38.5%	18.2%	2.3%	0.3%	100.0%
10	17.7%	21.7%	40.5%	17.4%	2.4%	0.3%	100.0%
11	18.9%	21.4%	40.1%	17.0%	2.3%	0.2%	100.0%
12	21.0%	20.8%	42.5%	14.1%	1.5%	0.1%	100.0%

1.3.2.2. Literacy Composite

1.3.2.2.1. By Cluster

Table 1.3.2.2.1.1

Proficiency Level by Cluster (Count): Literacy, S403 Online

Cluster	Literacy Proficiency Range						Total
	1	2	3	4	5	6	
1	32,238	82,333	54,653	12,156	1,815	166	183,361
2–3	27,507	69,785	176,879	84,437	13,102	1,520	373,230
4–5	17,330	26,139	106,194	112,087	25,746	3,510	291,006
6–8	48,321	50,506	105,223	47,174	3,835	255	255,314
9–12	33,533	55,593	107,452	65,616	13,275	503	275,972

Table 1.3.2.2.1.2

Proficiency Level by Cluster (Percent): Literacy, S403 Online

Cluster	Literacy Proficiency Range						Total
	1	2	3	4	5	6	
1	17.6%	44.9%	29.8%	6.6%	1.0%	0.1%	100.0%
2–3	7.4%	18.7%	47.4%	22.6%	3.5%	0.4%	100.0%
4–5	6.0%	9.0%	36.5%	38.5%	8.8%	1.2%	100.0%
6–8	18.9%	19.8%	41.2%	18.5%	1.5%	0.1%	100.0%
9–12	12.2%	20.1%	38.9%	23.8%	4.8%	0.2%	100.0%

1.3.2.2.2. By Grade

Table 1.3.2.2.2.1

Proficiency Level by Grade (Count): Literacy, S403 Online

Grade	Literacy Proficiency Range						Total
	1	2	3	4	5	6	
1	32,238	82,333	54,653	12,156	1,815	166	183,361
2	14,033	40,724	93,705	33,312	4,130	420	186,324
3	13,474	29,061	83,174	51,125	8,972	1,100	186,906
4	8,774	13,902	60,909	61,580	13,384	1,837	160,386
5	8,556	12,237	45,285	50,507	12,362	1,673	130,620
6	16,261	17,134	42,989	16,026	901	67	93,378
7	15,858	17,725	33,661	15,275	1,338	100	83,957
8	16,202	15,647	28,573	15,873	1,596	88	77,979
9	13,130	16,653	33,649	21,104	4,517	259	89,312
10	7,728	14,445	29,139	18,328	3,749	160	73,549
11	6,654	13,284	25,120	16,266	3,479	71	64,874
12	6,021	11,211	19,544	9,918	1,530	13	48,237

Table 1.3.2.2.2.2

Proficiency Level by Grade (Percent): Literacy, S403 Online

Grade	Literacy Proficiency Range						Total
	1	2	3	4	5	6	
1	17.6%	44.9%	29.8%	6.6%	1.0%	0.1%	100.0%
2	7.5%	21.9%	50.3%	17.9%	2.2%	0.2%	100.0%
3	7.2%	15.5%	44.5%	27.4%	4.8%	0.6%	100.0%
4	5.5%	8.7%	38.0%	38.4%	8.3%	1.1%	100.0%
5	6.6%	9.4%	34.7%	38.7%	9.5%	1.3%	100.0%
6	17.4%	18.3%	46.0%	17.2%	1.0%	0.1%	100.0%
7	18.9%	21.1%	40.1%	18.2%	1.6%	0.1%	100.0%
8	20.8%	20.1%	36.6%	20.4%	2.0%	0.1%	100.0%
9	14.7%	18.6%	37.7%	23.6%	5.1%	0.3%	100.0%
10	10.5%	19.6%	39.6%	24.9%	5.1%	0.2%	100.0%
11	10.3%	20.5%	38.7%	25.1%	5.4%	0.1%	100.0%
12	12.5%	23.2%	40.5%	20.6%	3.2%	0.0%	100.0%

1.3.2.3. Comprehension Composite

1.3.2.3.1. By Cluster

Table 1.3.2.3.1.1

Proficiency Level by Cluster (Count): Comprehension, S403 Online

Cluster	Comprehension Proficiency Range						Total
	1	2	3	4	5	6	
1	13,495	27,464	42,973	24,810	35,904	31,052	175,698
2–3	31,117	62,760	81,288	48,459	68,168	66,185	357,977
4–5	11,462	23,957	40,470	39,504	74,633	115,319	305,345
6–8	39,810	54,673	60,416	39,584	40,341	28,016	262,840
9–12	41,855	69,840	60,192	34,835	41,357	34,061	282,140

Table 1.3.2.3.1.2

Proficiency Level by Cluster (Percent): Comprehension, S403 Online

Cluster	Comprehension Proficiency Range						Total
	1	2	3	4	5	6	
1	7.7%	15.6%	24.5%	14.1%	20.4%	17.7%	100.0%
2–3	8.7%	17.5%	22.7%	13.5%	19.0%	18.5%	100.0%
4–5	3.8%	7.8%	13.3%	12.9%	24.4%	37.8%	100.0%
6–8	15.1%	20.8%	23.0%	15.1%	15.3%	10.7%	100.0%
9–12	14.8%	24.8%	21.3%	12.3%	14.7%	12.1%	100.0%

1.3.2.3.2. By Grade

Table 1.3.2.3.2.1

Proficiency Level by Grade (Count): Comprehension, S403 Online

Grade	Comprehension Proficiency Range						Total
	1	2	3	4	5	6	
1	13,495	27,464	42,973	24,810	35,904	31,052	175,698
2	13,969	33,380	43,050	25,596	33,885	28,586	178,466
3	17,148	29,380	38,238	22,863	34,283	37,599	179,511
4	4,667	12,384	21,932	21,539	42,432	65,903	168,857
5	6,795	11,573	18,538	17,965	32,201	49,416	136,488
6	11,983	20,786	23,988	15,617	16,304	7,904	96,582
7	13,690	18,150	19,709	12,848	12,035	9,644	86,076
8	14,137	15,737	16,719	11,119	12,002	10,468	80,182
9	13,056	24,046	19,428	10,980	12,822	10,820	91,152
10	10,940	17,835	16,506	9,408	10,921	9,562	75,172
11	9,819	15,689	13,609	7,962	10,681	8,830	66,590
12	8,040	12,270	10,649	6,485	6,933	4,849	49,226

Table 1.3.2.3.2.2

Proficiency Level by Grade (Percent): Comprehension, S403 Online

Grade	Comprehension Proficiency Range						Total
	1	2	3	4	5	6	
1	7.7%	15.6%	24.5%	14.1%	20.4%	17.7%	100.0%
2	7.8%	18.7%	24.1%	14.3%	19.0%	16.0%	100.0%
3	9.6%	16.4%	21.3%	12.7%	19.1%	20.9%	100.0%
4	2.8%	7.3%	13.0%	12.8%	25.1%	39.0%	100.0%
5	5.0%	8.5%	13.6%	13.2%	23.6%	36.2%	100.0%
6	12.4%	21.5%	24.8%	16.2%	16.9%	8.2%	100.0%
7	15.9%	21.1%	22.9%	14.9%	14.0%	11.2%	100.0%
8	17.6%	19.6%	20.9%	13.9%	15.0%	13.1%	100.0%
9	14.3%	26.4%	21.3%	12.0%	14.1%	11.9%	100.0%
10	14.6%	23.7%	22.0%	12.5%	14.5%	12.7%	100.0%
11	14.7%	23.6%	20.4%	12.0%	16.0%	13.3%	100.0%
12	16.3%	24.9%	21.6%	13.2%	14.1%	9.9%	100.0%

1.3.2.4. Overall Composite

1.3.2.4.1. By Cluster

Table 1.3.2.4.1.1

Proficiency Level by Cluster (Count): Overall, S403 Online

Cluster	Overall Proficiency Range						Total
	1	2	3	4	5	6	
1	16,588	50,813	78,148	16,688	2,640	145	165,022
2–3	25,566	63,080	156,906	81,987	9,397	258	337,194
4–5	11,700	19,322	78,782	116,417	32,854	3,266	262,341
6–8	28,917	46,543	93,138	54,919	4,506	257	228,280
9–12	34,707	49,860	104,471	51,698	6,656	200	247,592

Table 1.3.2.4.1.2

Proficiency Level by Cluster (Percent): Overall, S403 Online

Cluster	Overall Proficiency Range						Total
	1	2	3	4	5	6	
1	10.1%	30.8%	47.4%	10.1%	1.6%	0.1%	100.0%
2–3	7.6%	18.7%	46.5%	24.3%	2.8%	0.1%	100.0%
4–5	4.5%	7.4%	30.0%	44.4%	12.5%	1.2%	100.0%
6–8	12.7%	20.4%	40.8%	24.1%	2.0%	0.1%	100.0%
9–12	14.0%	20.1%	42.2%	20.9%	2.7%	0.1%	100.0%

1.3.2.4.2. By Grade

Table 1.3.2.4.2.1

Proficiency Level by Grade (Count): Overall, S403 Online

Grade	Overall Proficiency Range						Total
	1	2	3	4	5	6	
1	16,588	50,813	78,148	16,688	2,640	145	165,022
2	12,831	37,285	80,984	33,045	3,443	132	167,720
3	12,735	25,795	75,922	48,942	5,954	126	169,474
4	5,651	9,959	44,590	64,411	17,767	1,880	144,258
5	6,049	9,363	34,192	52,006	15,087	1,386	118,083
6	8,081	16,653	37,614	19,892	1,103	70	83,413
7	9,910	15,734	29,784	17,643	1,621	94	74,786
8	10,926	14,156	25,740	17,384	1,782	93	70,081
9	12,074	15,329	33,203	16,888	2,452	97	80,043
10	8,554	12,867	28,125	14,626	1,812	54	66,038
11	7,584	11,627	24,379	12,686	1,670	42	57,988
12	6,495	10,037	18,764	7,498	722	7	43,523

Table 1.3.2.4.2.2

Proficiency Level by Grade (Percent): Overall, S403 Online

Grade	Overall Proficiency Range						Total
	1	2	3	4	5	6	
1	10.1%	30.8%	47.4%	10.1%	1.6%	0.1%	100.0%
2	7.7%	22.2%	48.3%	19.7%	2.1%	0.1%	100.0%
3	7.5%	15.2%	44.8%	28.9%	3.5%	0.1%	100.0%
4	3.9%	6.9%	30.9%	44.6%	12.3%	1.3%	100.0%
5	5.1%	7.9%	29.0%	44.0%	12.8%	1.2%	100.0%
6	9.7%	20.0%	45.1%	23.8%	1.3%	0.1%	100.0%
7	13.3%	21.0%	39.8%	23.6%	2.2%	0.1%	100.0%
8	15.6%	20.2%	36.7%	24.8%	2.5%	0.1%	100.0%
9	15.1%	19.2%	41.5%	21.1%	3.1%	0.1%	100.0%
10	13.0%	19.5%	42.6%	22.1%	2.7%	0.1%	100.0%
11	13.1%	20.1%	42.0%	21.9%	2.9%	0.1%	100.0%
12	14.9%	23.1%	43.1%	17.2%	1.7%	0.0%	100.0%

2 Analysis of Domains

The measurement model that forms the basis of the analysis for the development of ACCESS is the Rasch measurement model (Wright & Stone, 1979). Additional information on its use in the development of the ACCESS for ELLs assessment program is available in WIDA Consortium Technical Report No. 1, *Development and Field Test of ACCESS for ELLs* (Kenyon, 2006). The original ACCESS 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. Thus, 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. All Rasch analyses are conducted using the Rasch measurement software program *Winsteps* (Linacre, 2006).

Rasch Model for Dichotomous Scoring

For Listening and Reading, the dichotomous Rasch model is used as the measurement model. Mathematically, the measurement model may be presented as

$$\log\left(\frac{P_{ni1}}{P_{ni0}}\right) = B_n - D_i$$

where

P_{ni1} = probability of a correct response “1” by person “n” on item “i”
P_{ni0} = probability of an incorrect response “0” by person “n” on item “i”
B_n = ability of person “n”
D_i = difficulty of item “i”

When the probability of a person getting a correct answer equals the probability of a person getting an incorrect answer (i.e., 50% probability of getting it right and 50% probability of getting it wrong), P_{ni1}/P_{ni0} is equal to 1. The log of 1 is 0. This is the point at which a person’s ability equals the difficulty of an item. For example, a person whose ability is 1.56 on the Rasch logit scale encountering an item whose difficulty is 1.56 on the Rasch logit scale would have a 50% probability of answering that question correctly.

Rasch Model for Polytomous Scoring

For the Writing and Speaking tasks, a Rasch-grouped rating scale model is used. Mathematically, this can be represented as

$$\log\left(\frac{P_{ngik}}{1 - P_{ngi(k-1)}}\right) = \beta_n - D_{gi} - F_{gk}$$

where

P_{ngik} = probability of person “n” on task “i” receiving a rating at level “k” on rating scale “g”
$P_{ngi(k-1)}$ = probability of person “n” on task “i” receiving a rating at level “k – 1” on rating scale “g” (i.e., the next lowest rating)
β_n = ability of person “n”
D_{gi} = difficulty of task “i” specific to rating scale “g”
F_{gk} = calibration of step “k” on rating scale “g”

The subscript “g” is a group index specifying the group of tasks to which task “i” belongs. It also identifies the scoring scale that was used for the group of tasks.

As described in Part 1 Section 3.2.3, ratings on the ACCESS Writing Scoring Scale range from 0, 1, 1+,…, 6 and the possible raw scores range from 0 to 9. All Writing tasks are scored using this scoring scale except for Grade 1 Tier A Tasks 1 and 2. The profiles of the responses to these two tasks do not fit the generic scoring scale well, so additional task-specific instructions are provided to raters. These instructions guide raters in applying a limited number of score points on the scoring scale to responses elicited by these two tasks. The possible ratings for Grade 1 Tier A Task 1 are 0 or 1, and the possible ratings for Grade 1 Tier A Task 2 are 0, 1, 1+, or 2. To simplify the year-to-year linking process, the Grade 1 Writing Tier A Task 1 is treated as a dichotomously scored task. The Grade 1 Writing Tier A Task 2 is modeled using a rating scale with a possible raw score of 0 to 3. All other Writing tasks are modeled using a rating scale with possible raw scores of 0 to 9. Thus, a total of two rating scales are being modeled for ACCESS Writing. One rating scale is associated with the Grade 1 Writing Tier A Task 2, and the other rating scale is associated with all Writing tasks that are scored using the rating scale with raw score values 0–9.

For Speaking, Proficiency Level (PL) 1 tasks are modeled as a group on a 0–2 scale and PL 3 and PL 5 tasks are modeled as a group on a 0–4 scale (see Part 1 Section 3.2.4).

Scale Scores and Proficiency Level Scores

Scale scores are calculated by transforming the person ability estimate via a scaling equation.

The following scaling equations are used to convert ability measures in logits to scale scores:

- L: (Ability Measure in Logits * 37.571) + 316.637
- R: (Ability Measure in Logits * 26.000) + 323.272
- W: (Ability Measure in Logits * 26.851) + 303.332
- S: (Ability Measure in Logits * 29.248) + 265.076

In the domains of Listening and Reading, the current ACCESS scale was established for the original paper-only version of the test and was maintained through the transition to an online and paper delivered test in the 2015–16 school year (Series 400). Evidence for scale maintenance in the transitional year is described elsewhere (Center for Applied Linguistics, 2016). In the domains of Writing and Speaking, a study was conducted in the summer of 2016 to reconstruct the logit scale (see Center for Applied Linguistics, 2017).

Proficiency level scores are interpretations of these scale scores in terms of the proficiency levels described in the WIDA ELD Standards. These interpretations derive from a series of standard setting studies, in which educators reviewed evidence from the test, either in the form of items for the selected response sections (Listening and Reading) or student portfolios for the constructed response sections (Writing and Speaking), to establish cut scores between the proficiency levels. The first standard setting study for ACCESS took place in 2005; it established cut scores for all four domains by grade-level cluster (Kenyon, 2006). The second cut score study took place in 2007; it established cut scores for all four domains by grade level (Kenyon, Ryu, & MacGregor, 2013). These cut scores were used to derive proficiency level scores through the 2015–16 administration (Series 400) of ACCESS for ELLs. A third cut score study was conducted in summer 2016 (Cook & MacGregor, 2017). The purpose of this study was to re-examine cut scores for each of the proficiency levels in light of the migration from the paper-and-pencil–only assessment to both online and paper delivery, the revision of the Speaking test, and the influence of college- and career-ready standards. These new cut scores were first used for ACCESS Series 401 (2016–17 school year).

A proficiency level score consists of a two-digit decimal number (e.g., 4.5). The first digit represents the student's overall proficiency level range based on the student's scale score. The number to the right of the decimal is an indication of the proportion of the range between cut scores that the student's scale score represents. A score of 4.5, for example, tells us that the student is in PL 4 and that his or her scale score is halfway between the cut scores for PLs 4 and 5.

Unlike the scale scores, which form an interval scale and are continuous across grades from Kindergarten to Grade 12, proficiency level scores are dependent upon the grade a student was in when the assessment was administered. For example, a score of 350 in Listening would be interpreted as a PL score of 5.8 for a Grade 2 student, a 3.8 for a Grade 5 student, a 3.1 for a Grade 8 student, and a 2.3 for a Grade 12 student.

Because the bands between cut scores on the score scale vary in width, proficiency level scores should not be considered to form an interval scale. That is, the distance between PL scores 1.5 and 2.5 cannot be assumed to be equal to the distance between PL scores 2.5 and 3.5. Only scale scores should be used as interval measures. Proficiency level scores are at even intervals within a grade and proficiency level (e.g., in Grade 3, the distance between 3.1 and 3.2 is the same as the distance between 3.7 and 3.8), but they do not form an interval scale across proficiency levels.

2.1 Complete Item or Task Analysis and Summary

The tables in this section provide a summary of the analyses of the items (for Listening and Reading) or tasks (for Writing and Speaking), along with analyses of each item or task. These tables have either two parts (in the case of Listening and Reading) or three parts (in the case of Writing and Speaking). The first part of the table gives a summary of the total set of items or tasks on the form. The second part provides statistics pertaining to the individual items or tasks, and the third part (for Writing and Speaking only) expresses raw score distributions by task.

Statistics included across these three parts include item or task difficulties in logits, the number of items or tasks on the form, the average *p* value (for forms with selected-response items), and the Rasch model fit statistics.

For Listening and Reading, items form a pool for the multistage adaptive tests, and tables in this section provide information on every item in the grade-level cluster. For Writing, separate tables are provided for Tier A and Tier B/C forms, by grade-level cluster. For Speaking, which has tasks that are shared between Tier A and Tier B/C, there is one table for each grade-level cluster, which provides information on every task in the grade-level cluster.

All Rasch analyses were conducted using the Rasch measurement software program *Winsteps* (Linacre, 2006). When speaking of the measure of person ability, we use the term *ability measure* (rather than *theta,* used commonly when discussing models based on item response theory). When speaking of the measure of how hard an item is, we use the term *item difficulty measure* (rather than *b parameter,* used commonly when discussing models based on item response theory). *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 ACCESS score scale for reporting purposes.

Fit statistics for the Rasch model are calculated by comparing the observed empirical data with the data that would be expected to be produced by the Rasch model if the data fit the model perfectly. Outfit mean square statistics are influenced by outliers. For example, a difficult item that some low-ability examinees get correct—for reasons unknown—will have a high outfit mean square statistic. Infit mean square statistics are influenced by unexpected patterns of observations by persons on items that are roughly targeted for them and generally indicate a more serious measurement problem. The expectation for both of these statistics is 1.00, and values near 1.00 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 of greater concern.

Linacre (2002) provided more guidance on how to interpret these statistics for dichotomous items. He writes:

- 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 “less productive for measurement, but not degrading.”

Linacre also stated in his guidance that infit problems are more serious to the construction of measurement than are outfit problems.

Because conservative guidelines were followed in the development of ACCESS for ELLs, the vast majority of dichotomous items on the test forms have mean square fit statistics in the range of 0.5 and 1.5; thus, they fit the range that is “productive for measurement” according to the guidelines above.

Since performance tasks are constructed and scored very differently from dichotomous items, it is not as straightforward to apply this same guidance to interpret these fit statistics for performance tasks that were scored polytomously. Some performance tasks that were designed to elicit a restricted range of performances (for example, very easy tasks where it is expected that most students will get the highest score) can cause the model to predict the data too well (overfitting). Conversely, when performance tasks are scored using a very wide rubric scale such as the case with ACCESS for ELLs Writing tasks, sometimes unmodeled noise or other sources of variance in the data will cause the model to underpredict the data (underfitting). Overall, for ACCESS for ELLs performance tasks, overfitting is more common than underfitting. Underfitting indicates that the task is less productive for measurement, but it is not degrading to the measurement of student performance.

The first section of the Complete Item/Task Analysis and Summary table provides information about the total set of items or tasks and includes the item type (selected response or constructed response), the average item difficulty (in logits), the number of items, the average *p* value (for Listening and Reading only), the average infit mean square, and the average outfit mean square.

The second section of these tables presents results of the analyses of all of the items or tasks on the test form. The first column provides the unique item name. The second column in this section presents the item difficulty in logits. The third column provides information on whether the item or task served as an anchor item or task. For dichotomously scored items (Listening and Reading), the fourth column shows the *p* value (percentage of correct answers on that item). The next two columns show the Rasch fit statistics for the item or task.

In addition, Writing and Speaking tables have a final section which provides raw score distributions by task.

2.1.1 Listening

2.1.1.1 Grade 1

Table 2.1.1.1

Complete Item Analysis and Summary: List 1 S403 Online

Item Type	Average Item Difficulty (in logits)	No. of Items	Average P-value	Average Infit Mean Square	Average Outfit Mean Square
Selected Response	-1.16	54	0.63	0.97	0.97

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics	
				Infit Mnsq	Outfit Mnsq
1.L01A_SI_GoingHome_P100_Screen_2_12445	-4.38	Yes	0.97	0.90	0.82
2.L01A_SI_GoingHome_P100_Screen_3_12446	-2.54	Yes	0.87	1.18	1.74
3.L01A_SI_GoingHome_P100_Screen_4_12447	1.15	Yes	0.37	1.26	2.01
4.L01B_SI_MorningMeeting_402_Screen_2_14897	-1.30		0.76	1.18	1.39
5.L01B_SI_MorningMeeting_402_Screen_3_14898	-1.93		0.82	1.17	1.18
6.L01B_SI_MorningMeeting_402_Screen_4_14899	-1.50	Yes	0.78	0.94	0.83
7.L01A_LA_OutdoorFun_P100_Screen_2_12690	-3.39	Yes	0.61	0.88	0.83
8.L01A_LA_OutdoorFun_P100_Screen_3_12691	-2.96	Yes	0.57	1.00	1.00
9.L01A_LA_OutdoorFun_P100_Screen_4_12692	-3.17	Yes	0.63	0.92	0.87
10.L01A_MA_DrawingaRobot_P100_A202_Screen_2_13889	-3.27	Yes	0.63	0.92	0.88
11.L01A_MA_DrawingaRobot_P100_A202_Screen_3_13890	-2.23	Yes	0.41	0.97	0.97
12.L01A_MA_DrawingaRobot_P100_A202_Screen_4_13891	-2.55	Yes	0.50	1.00	1.00
13.L01C_SS_CampingTrip_P100_A301FT_Screen_2_13808	-2.41	Yes	0.46	1.01	1.02
14.L01C_SS_CampingTrip_P100_A301FT_Screen_3_13809	-1.90	Yes	0.33	0.97	1.00
15.L01C_SS_CampingTrip_P100_A301FT_Screen_4_13810	-2.97	Yes	0.57	0.94	0.92
16.L01B_SC_PropertiesOfObjects_402_Screen_2_14951	-2.03		0.36	0.92	0.92
17.L01B_SC_PropertiesOfObjects_402_Screen_3_14952	-3.03	Yes	0.65	0.92	0.88
18.L01B_SC_PropertiesOfObjects_402_Screen_4_14953	-1.43		0.28	0.95	0.98
19.L01C_LA_WindyDay_402_Screen_2_14836	-1.14		0.61	0.90	0.83
20.L01C_LA_WindyDay_402_Screen_3_14837	-2.11	Yes	0.79	0.95	0.91
21.L01C_LA_WindyDay_402_Screen_4_14838	-1.32		0.65	0.96	0.93
22.L01C_MA_RainyDay_P100_A202_Screen_2_13898	-0.94	Yes	0.56	0.99	0.97
23.L01C_MA_RainyDay_P100_A202_Screen_3_13899	-1.15	Yes	0.59	0.94	0.90
24.L01C_MA_RainyDay_P100_A202_Screen_4_13900	-1.82	Yes	0.69	0.94	0.89
25.L01B_SS_ConstructionWorker_P100_A301FT_alt1_Screen_2_13802	-2.40	Yes	0.79	0.92	0.86

				Fit Statistics	
Name	Item Difficulty (in logits)	Anchored?	P-value	Infit Mnsq	Outfit Mnsq
26.L01B_SS_ConstructionWorker_P100_A301FT_alt1_Screen_3_13803	-1.67	Yes	0.74	0.91	0.87
27.L01B_SS_ConstructionWorker_P100_A301FT_alt1_Screen_4_13804	-1.03	Yes	0.51	0.91	0.90
28.L01C_SC_ForestHabitat_P100_Screen_2_11667	0.44	Yes	0.27	0.94	0.92
29.L01C_SC_ForestHabitat_P100_Screen_3_11668	-0.81	Yes	0.50	0.96	0.95
30.L01C_SC_ForestHabitat_P100_Screen_4_11671	-2.02	Yes	0.77	0.95	0.91
31.L01B_LA_Chef_403_Screen_2_16531	-1.79		0.77	0.94	0.90
32.L01B_LA_Chef_403_Screen_3_16533	0.25		0.44	0.95	0.95
33.L01B_LA_Chef_403_Screen_4_16535	-0.41		0.61	0.96	0.94
34.L01B_MA_OnesTensBlocks_403_Screen_2_16640	0.67		0.37	0.94	0.93
35.L01B_MA_OnesTensBlocks_403_Screen_3_16641	0.16		0.52	0.97	0.96
36.L01B_MA_OnesTensBlocks_403_Screen_4_16642	0.10		0.49	0.94	0.94
37.L01B_LA_ASpecialDay_P100_alt1_Screen_2_13814	-1.12		0.86	0.96	0.88
38.L01B_LA_ASpecialDay_P100_alt1_Screen_3_13815	-0.38	Yes	0.74	1.02	1.03
39.L01B_LA_ASpecialDay_P100_alt1_Screen_4_13816	-1.89	Yes	0.94	0.95	0.80
40.L01C_MA_ShapeRiddles_P100_Screen_2_12846	0.58	Yes	0.58	1.05	1.05
41.L01C_MA_ShapeRiddles_P100_Screen_3_12847	0.02	Yes	0.69	0.98	0.97
42.L01C_MA_ShapeRiddles_P100_Screen_4_12848	0.10	Yes	0.68	0.96	0.93
43.L01B_SS_SchoolCommunity_P100_Screen_2_12745	-1.35	Yes	0.88	0.98	0.96
44.L01B_SS_SchoolCommunity_P100_Screen_3_12746	-0.86	Yes	0.85	0.94	0.86
45.L01B_SS_SchoolCommunity_P100_Screen_4_12754	-0.47	Yes	0.80	0.94	0.91
46.L01B_SC_Solids_P100_alt1_Screen_2_13877	1.34	Yes	0.46	0.98	0.97
47.L01B_SC_Solids_P100_alt1_Screen_3_13878	0.31	Yes	0.69	0.93	0.89
48.L01B_SC_Solids_P100_alt1_Screen_4_13879	-0.64	Yes	0.84	0.95	0.89
49.L01C_LA_MeetTheScientist_403_Screen_2_16558	-0.15		0.72	0.96	0.94
50.L01C_LA_MeetTheScientist_403_Screen_3_16559	0.50		0.58	0.99	0.99
51.L01C_LA_MeetTheScientist_403_Screen_4_16560	-0.02		0.71	0.93	0.88
52.L01C_MA_FestivalFun_403_Screen_2_16648	-1.24		0.87	0.97	0.93
53.L01C_MA_FestivalFun_403_Screen_3_16649	1.13		0.47	1.01	1.01
54.L01C_MA_FestivalFun_403_Screen_4_16650	0.63		0.54	1.04	1.05

2.1.1.2 *Grades 2–3*

Table 2.1.1.2

Complete Item Analysis and Summary: List 2-3 S403 Online

[redacted]

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	Outfit Mnsq
26.L23C_SS_AMapOfOakValley_P 100_Screen_3_12814	-1.26	Yes	0.70	0.92	0.87
27.L23C_SS_AMapOfOakValley_P 100_Screen_4_12815	-0.47	Yes	0.54	0.93	0.89
28.L23B_SC_WaterCycle_P 100_A202_Screen_2_13910	-2.33	Yes	0.86	0.97	0.91
29.L23B_SC_WaterCycle_P 100_A202_Screen_3_13911	-0.87	Yes	0.55	1.00	0.98
30.L23B_SC_WaterCycle_P 100_A202_Screen_4_13912	-0.24	Yes	0.46	0.94	0.92
31.L23B_LA_VisitingAuthor_403_Screen_2_16602	-2.13		0.90	0.98	0.91
32.L23B_LA_VisitingAuthor_403_Screen_3_16603	0.18		0.52	0.94	0.92
33.L23B_LA_VisitingAuthor_403_Screen_4_16604	0.62		0.45	0.94	0.92
34.L23B_MA_LinearMeasurement_P 100_Screen_2_12988	-1.07	Yes	0.69	0.96	0.95
35.L23B_MA_LinearMeasurement_P 100_Screen_3_12990	0.29	Yes	0.31	1.00	0.99
36.L23B_MA_LinearMeasurement_P 100_Screen_4_12991	-0.07	Yes	0.39	0.98	0.97
37.L23C_LA_CrownAndTheCoins_P 100_Screen_2_12705	-0.17	Yes	0.82	0.99	1.01
38.L23C_LA_CrownAndTheCoins_P 100_Screen_3_12706	-1.64	Yes	0.94	0.97	0.92
39.L23C_LA_CrownAndTheCoins_P 100_Screen_4_12707	-0.30	Yes	0.82	0.94	0.86
40.L23B_MA_GettingToSchool_P 100_Screen_2_12953	0.13	Yes	0.75	1.11	1.21
41.L23B_MA_GettingToSchool_P 100_Screen_3_12954	-1.97	Yes	0.96	1.00	1.03
42.L23B_MA_GettingToSchool_P 100_Screen_4_12955	1.23	Yes	0.54	1.06	1.07
43.L23B_SS_AMapOfOakValley_P 100_Screen_2_12733	-1.69	Yes	0.94	0.97	0.83
44.L23B_SS_AMapOfOakValley_P 100_Screen_3_12734	0.26	Yes	0.79	1.01	1.03
45.L23B_SS_AMapOfOakValley_P 100_Screen_4_12735	1.28	Yes	0.53	1.03	1.03
46.L23B_SC_Habitats_P 100_Screen_2_12785	0.00	Yes	0.82	0.97	0.93
47.L23B_SC_Habitats_P 100_Screen_3_12786	1.56	Yes	0.53	0.92	0.90
48.L23B_SC_Habitats_P 100_Screen_4_12787	-0.11	Yes	0.83	0.95	0.90
49.L23C_LA_Visitor_403_Screen_2_16652	-1.50		0.93	0.93	0.73
50.L23C_LA_Visitor_403_Screen_3_16653	0.43		0.66	1.03	1.05
51.L23C_LA_Visitor_403_Screen_4_16654	0.76		0.63	0.94	0.93
52.L23C_MA_WritingWall_403_Screen_2_16684	1.64		0.45	0.98	0.97
53.L23C_MA_WritingWall_403_Screen_3_16685	0.04		0.75	0.97	0.95
54.L23C_MA_WritingWall_403_Screen_4_16686	-0.93		0.88	0.96	0.88

2.1.1.3 *Grades 4–5*

Table 2.1.1.3

Complete Item Analysis and Summary: List 4-5 S403 Online

Item Type	Average Item Difficulty (in logits)	No. of Items	Average P-value	Average Infit Mean Square	Average Outfit Mean Square
Selected Response	0.41	54	0.62	0.97	1.00

				Fit Statistics	
Name	Item Difficulty (in logits)	Anchored?	P-value	Infit Mnsq	Outfit Mnsq
1.L45A_SI_MorningArrivalAtClass_P100_Screen_2_12413	-1.62		0.96	1.26	2.31
2.L45A_SI_MorningArrivalAtClass_P100_Screen_3_12415	-2.46		0.98	0.95	1.51
3.L45A_SI_MorningArrivalAtClass_P100_Screen_4_12416	-1.40	Yes	0.95	0.93	1.04
4.L45B_SI_SchoolClubs_P100_Screen_2_12386	-1.11		0.95	0.98	1.09
5.L45B_SI_SchoolClubs_P100_Screen_3_12394	-0.89	Yes	0.93	1.22	1.73
6.L45B_SI_SchoolClubs_P100_Screen_4_12396	1.26	Yes	0.76	0.97	0.94
7.L45B_LA_BookDiscussion_P100_Screen_2_12571	-1.61	Yes	0.54	0.88	0.84
8.L45B_LA_BookDiscussion_P100_Screen_3_12572	-1.63		0.56	0.88	0.84
9.L45B_LA_BookDiscussion_P100_Screen_4_12573	-0.54		0.38	1.07	1.08
10.L45B_MA_TheSchoolMascot_P100_Screen_2_12917	-0.39	Yes	0.29	0.90	0.89
11.L45B_MA_TheSchoolMascot_P100_Screen_3_12918	-0.47	Yes	0.29	0.98	1.00
12.L45B_MA_TheSchoolMascot_P100_Screen_4_12919	-2.65	Yes	0.75	0.94	0.90
13.L45A_SS_RomanToolsArtifacts_P100_Screen_2_13024	-1.78	Yes	0.57	0.99	0.99
14.L45A_SS_RomanToolsArtifacts_P100_Screen_3_13025	-0.38	Yes	0.24	0.96	0.96
15.L45A_SS_RomanToolsArtifacts_P100_Screen_4_13026	-0.76	Yes	0.33	0.94	0.94
16.L45A_SC_WindSpeed_P100_Screen_2_12792	-1.97	Yes	0.60	0.97	0.95
17.L45A_SC_WindSpeed_P100_Screen_3_12793	-2.36	Yes	0.70	0.99	0.98
18.L45A_SC_WindSpeed_P100_Screen_4_12794	-0.22	Yes	0.26	0.97	0.97
19.L45C_LA_BookDiscussion_P100_Screen_2_12581	-1.63		0.87	0.88	0.72
20.L45C_LA_BookDiscussion_P100_Screen_3_12583	1.00		0.45	0.96	0.94
21.L45C_LA_BookDiscussion_P100_Screen_4_12585	0.50		0.55	0.99	0.94
22.L45B_MA_Garden_P100_Screen_2_12596	-0.86		0.78	0.93	0.88
23.L45B_MA_Garden_P100_Screen_3_12597	0.43	Yes	0.53	0.99	0.96
24.L45B_MA_Garden_P100_Screen_4_12598	1.43		0.34	1.08	1.12
25.L45B_SS_FarmingInAncientEgypt_402_Screen_2_14945	0.10	Yes	0.66	0.90	0.86

	Item			Fit Statistics	
Name	Difficulty (in logits)	Anchored?	P-value	Infit Mnsq	Outfit Mnsq
26.L45B_SS_FarmingInAncientEgypt_402_Screen_3_14946	0.95		0.44	0.95	0.93
27.L45B_SS_FarmingInAncientEgypt_402_Screen_4_14947	-0.04		0.64	0.90	0.86
28.L45B_SC_DynamicDunes_402_Screen_2_14940	0.48		0.53	0.96	0.94
29.L45B_SC_DynamicDunes_402_Screen_3_14939	2.28		0.18	0.99	1.02
30.L45B_SC_DynamicDunes_402_Screen_4_14941	0.32		0.55	0.92	0.90
31.L45B_LA_LostBottle_403_Screen_2_16613	-0.38		0.79	0.90	0.82
32.L45B_LA_LostBottle_403_Screen_3_16615	0.10		0.62	0.90	0.87
33.L45B_LA_LostBottle_403_Screen_4_16616	1.07		0.58	0.94	0.92
34.L45B_MA_FractionBars_403_Screen_2_16708	0.13		0.79	0.95	0.92
35.L45B_MA_FractionBars_403_Screen_3_16709	2.61		0.37	0.99	0.99
36.L45B_MA_FractionBars_403_Screen_4_16710	0.53		0.75	0.96	0.94
37.L45C_LA_GardenMystery_402_Screen_2_14833	-0.21	Yes	0.94	0.99	0.96
38.L45C_LA_GardenMystery_402_Screen_3_14845	0.17	Yes	0.92	0.91	0.81
39.L45C_LA_GardenMystery_402_Screen_4_14846	3.22	Yes	0.46	1.00	1.04
40.L45C_MA_TheSchoolMascot_P100_Screen_2_12942	1.44	Yes	0.75	1.02	1.07
41.L45C_MA_TheSchoolMascot_P100_Screen_3_12943	3.60	Yes	0.33	0.92	0.90
42.L45C_MA_TheSchoolMascot_P100_Screen_4_12946	1.46	Yes	0.78	0.96	0.92
43.L45B_SS_RomanToolsArtifacts_P100_Screen_2_13027	3.33	Yes	0.49	1.05	1.06
44.L45B_SS_RomanToolsArtifacts_P100_Screen_3_13028	3.20	Yes	0.51	1.12	1.16
45.L45B_SS_RomanToolsArtifacts_P100_Screen_4_13029	1.83	Yes	0.74	0.95	0.92
46.L45C_SC_FromRocksToSoil_P100_Screen_2_14212	2.49	Yes	0.61	1.02	1.01
47.L45C_SC_FromRocksToSoil_P100_Screen_3_14213	1.13	Yes	0.83	0.94	0.86
48.L45C_SC_FromRocksToSoil_P100_Screen_4_14214	4.50	Yes	0.27	1.04	1.09
49.L45B_LA_RoleModels_403_Screen_2_16618	-1.76		0.99	0.90	0.62
50.L45B_LA_RoleModels_403_Screen_3_16619	2.15		0.67	0.95	0.92
51.L45B_LA_RoleModels_403_Screen_4_16620	2.50		0.58	0.97	0.95
52.L45B_MA_SchoolSupplies_403_Screen_2_16712	0.50		0.88	0.99	0.99
53.L45B_MA_SchoolSupplies_403_Screen_3_16713	1.64		0.75	1.02	1.03
54.L45B_MA_SchoolSupplies_403_Screen_4_16714	2.68		0.54	0.99	0.99

2.1.1.4 *Grades 6–8*

Table 2.1.1.4

Complete Item Analysis and Summary: List 6-8 S403 Online

Item Type	Average Item Difficulty (in logits)	No. of Items	Average P-value	Average Infit Mean Square	Average Outfit Mean Square
Selected Response	1.22	54	0.56	0.98	0.99

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	Outfit Mnsq
1.L68A_SL_TalentShow_P100_Screen_2_12454	-0.49	Yes	0.88	0.98	1.03
2.L68A_SL_TalentShow_P100_Screen_3_12455	-0.01	Yes	0.81	1.02	1.03
3.L68A_SL_TalentShow_P100_Screen_4_12456	1.26	Yes	0.63	1.02	1.01
4.L68B_SL_AuthorProject_P100_Screen_2_12380	-0.74	Yes	0.89	0.96	1.02
5.L68B_SL_AuthorProject_P100_Screen_3_12381	1.30	Yes	0.62	1.12	1.23
6.L68B_SL_AuthorProject_P100_Screen_4_12382	2.70	Yes	0.39	1.16	1.80
7.L68A_LA_RemoteControl_P100_Screen_2_12849	-0.68	Yes	0.57	0.87	0.82
8.L68A_LA_RemoteControl_P100_Screen_3_12850	0.90	Yes	0.28	1.06	1.12
9.L68A_LA_RemoteControl_P100_Screen_4_12922	-0.51	Yes	0.51	1.05	1.05
10.L68B_MA_ClassPresidentSurvey_P100_Screen_2_13039	-0.52		0.47	0.95	0.94
11.L68B_MA_ClassPresidentSurvey_P100_Screen_3_13040	-0.83		0.53	0.90	0.87
12.L68B_MA_ClassPresidentSurvey_P100_Screen_4_13041	-0.29		0.42	1.07	1.08
13.L68A_SS_TypesOfGovernment_P100_Screen_2_13060	-0.69	Yes	0.53	0.89	0.86
14.L68A_SS_TypesOfGovernment_P100_Screen_3_13061	-0.79	Yes	0.52	0.88	0.86
15.L68A_SS_TypesOfGovernment_P100_Screen_4_13062	-0.13	Yes	0.41	1.02	1.03
16.L68A_SC_NatureTrail_P100_Screen_2_12884	-2.40	Yes	0.84	0.97	0.90
17.L68A_SC_NatureTrail_P100_Screen_3_12885	-0.67	Yes	0.49	0.91	0.90
18.L68A_SC_NatureTrail_P100_Screen_4_12886	0.11	Yes	0.34	0.91	0.89
19.L68B_LA_Eastwood_402_Screen_2_14850	0.34		0.70	1.06	1.12
20.L68B_LA_Eastwood_402_Screen_3_14851	1.65		0.44	1.04	1.05
21.L68B_LA_Eastwood_402_Screen_4_14852	1.81		0.40	1.03	1.03
22.L68B_MA_WaterCooler_402_Screen_2_14915	1.08	Yes	0.59	0.88	0.85
23.L68B_MA_WaterCooler_402_Screen_3_14916	1.73		0.39	1.01	1.01
24.L68B_MA_WaterCooler_402_Screen_4_14917	1.05	Yes	0.58	1.00	0.98
25.L68B_SS_BillsOfRights_402_Screen_2_14858	1.54		0.45	0.94	0.93

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	 Outfit Mnsq
26.L68B_SS_BillsOfRights_402_Screen_3_14859	1.51		0.46	0.98	0.97
27.L68B_SS_BillsOfRights_402_Screen_4_14863	1.43	Yes	0.52	0.93	0.92
28.L68B_SC_LeafClassification_P100_Screen_2_11467	0.84	Yes	0.54	0.98	0.97
29.L68B_SC_LeafClassification_P100_Screen_3_11472	1.61	Yes	0.39	0.97	0.96
30.L68B_SC_LeafClassification_P100_Screen_4_11474	1.71	Yes	0.46	1.03	1.03
31.L68B_LA_Heroes_403_Screen_2_16562	0.70		0.73	0.95	0.92
32.L68B_LA_Heroes_403_Screen_3_16563	0.06		0.80	0.94	0.88
33.L68B_LA_Heroes_403_Screen_4_16564	2.03		0.42	0.93	0.92
34.L68B_MA_Forest_403_Screen_2_16664	0.58		0.75	0.94	0.91
35.L68B_MA_Forest_403_Screen_3_16665	1.70		0.54	0.97	0.96
36.L68B_MA_Forest_403_Screen_4_16666	1.98		0.47	0.97	0.97
37.L68B_LA_Steinman_402_Screen_2_14855	2.29		0.66	0.97	0.95
38.L68B_LA_Steinman_402_Screen_3_14856	3.31		0.43	1.02	1.03
39.L68B_LA_Steinman_402_Screen_4_14857	3.55		0.38	1.10	1.15
40.L68C_MA_MediaGrowth_402_Screen_2_14921	2.48		0.60	1.05	1.07
41.L68C_MA_MediaGrowth_402_Screen_3_14922	2.74		0.55	1.05	1.07
42.L68C_MA_MediaGrowth_402_Screen_4_14923	3.37		0.42	1.03	1.04
43.L68C_SS_ThailandMap_P100_A301FT_Screen_2_13838	2.51	Yes	0.63	1.01	1.00
44.L68C_SS_ThailandMap_P100_A301FT_Screen_3_13839	2.15	Yes	0.69	0.99	1.01
45.L68C_SS_ThailandMap_P100_A301FT_Screen_4_13840	2.02	Yes	0.71	0.93	0.88
46.L68C_SC_NatureTrail_P100_Screen_2_12899	2.09	Yes	0.72	0.95	0.93
47.L68C_SC_NatureTrail_P100_Screen_3_12901	2.25	Yes	0.69	0.94	0.91
48.L68C_SC_NatureTrail_P100_Screen_4_12905	2.50	Yes	0.64	0.93	0.91
49.L68B_LA_Themes_403_Screen_2_16566	1.36		0.75	0.98	1.00
50.L68B_LA_Themes_403_Screen_3_16567	3.07		0.44	0.97	0.96
51.L68B_LA_Themes_403_Screen_4_16568	1.72		0.72	0.95	0.93
52.L68C_MA_Regions_403_Screen_2_16672	2.78		0.54	0.95	0.93
53.L68C_MA_Regions_403_Screen_3_16673	2.13		0.67	0.95	0.92
54.L68C_MA_Regions_403_Screen_4_16674	2.97		0.42	1.00	1.00

2.1.1.5 Grades 9–12

Table 2.1.1.5

Complete Item Analysis and Summary: List 9-12 S403 Online

Item Type	Average Item Difficulty (in logits)	No. of Items	Average P-value	Average Infit Mean Square	Average Outfit Mean Square
Selected Response	1.43	54	0.60	0.97	0.97

				Fit Statistics	
Name	**Item Difficulty (in logits)**	**Anchored?**	**P-value**	**Infit Mnsq**	**Outfit Mnsq**
1.L91A_SI_ArtContest_P100_Screen_2_12360	-0.42	Yes	0.87	1.01	1.15
2.L91A_SI_ArtContest_P100_Screen_3_12361	2.15	Yes	0.46	1.13	1.26
3.L91A_SI_ArtContest_P100_Screen_4_12441	1.47	Yes	0.56	1.18	1.33
4.L91B_SI_RenewableEnergy_P100_Screen_2_12442	1.10	Yes	0.65	1.15	1.31
5.L91B_SI_RenewableEnergy_P100_Screen_3_12443	2.22	Yes	0.48	1.29	1.53
6.L91B_SI_RenewableEnergy_P100_Screen_4_12444	0.11	Yes	0.79	1.00	1.01
7.L91B_LA_FreeVerse_P100_Screen_2_12887	-0.80		0.72	0.88	0.79
8.L91B_LA_FreeVerse_P100_Screen_3_12889	0.12		0.53	0.92	0.89
9.L91B_LA_FreeVerse_P100_Screen_4_12890	0.60		0.44	0.95	0.94
10.L91B_MA_WaterSale_P100_Screen_2_12714	0.00		0.51	0.97	0.96
11.L91B_MA_WaterSale_P100_Screen_3_12715	1.01		0.30	0.97	0.98
12.L91B_MA_WaterSale_P100_Screen_4_12716	-0.45		0.61	1.00	0.99
13.L91B_SS_GlobalCarProduction_P100_Screen_2_12869	-0.48	Yes	0.63	0.93	0.91
14.L91B_SS_GlobalCarProduction_P100_Screen_3_12870	0.37	Yes	0.46	0.95	0.94
15.L91B_SS_GlobalCarProduction_P100_Screen_4_12871	1.19	Yes	0.27	1.00	1.04
16.L91A_SC_DesertPlants_P100_Screen_2_11034	-0.40	Yes	0.61	0.93	0.91
17.L91A_SC_DesertPlants_P100_Screen_3_11035	-0.52	Yes	0.63	0.93	0.91
18.L91A_SC_DesertPlants_P100_Screen_4_11036	0.51	Yes	0.41	0.96	0.95
19.L91C_LA_SoniaAtWork_402_Screen_2_14885	2.16	Yes	0.45	1.00	0.99
20.L91C_LA_SoniaAtWork_402_Screen_3_14886	1.96	Yes	0.48	0.91	0.87
21.L91C_LA_SoniaAtWork_402_Screen_4_14888	1.28	Yes	0.64	0.99	0.99
22.L91B_MA_SchoolMarket_402_Screen_2_15649	0.54	Yes	0.70	0.94	0.90
23.L91B_MA_SchoolMarket_402_Screen_3_15650	1.23	Yes	0.57	0.94	0.92
24.L91B_MA_SchoolMarket_402_Screen_4_15651	2.46	Yes	0.36	1.04	1.04
25.L91B_SS_CostaRica_402_Screen_2_14954	0.53	Yes	0.71	0.90	0.86

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	Fit Statistics Outfit Mnsq
26.L91B_SS_Costa Rica_402_Screen_3_14955	1.98	Yes	0.44	0.95	0.93
27.L91B_SS_Costa Rica_402_Screen_4_14956	1.86	Yes	0.47	0.97	0.95
28.L91B_SC_GrasshopperLifeCycle_P100_Screen_2_11198	1.98	Yes	0.45	1.05	1.06
29.L91B_SC_GrasshopperLifeCycle_P100_Screen_3_11199	0.40	Yes	0.76	0.86	0.78
30.L91B_SC_GrasshopperLifeCycle_P100_Screen_4_11215	1.44	Yes	0.55	0.95	0.94
31.L91B_LA_CommunityGarden_403_Screen_2_16586	1.08		0.68	0.92	0.89
32.L91B_LA_CommunityGarden_403_Screen_3_16587	2.22		0.45	0.94	0.93
33.L91B_LA_CommunityGarden_403_Screen_4_16588	0.92		0.70	0.94	0.91
34.L91B_MA_BasketballMath_403_Screen_2_16656	2.18		0.48	0.95	0.94
35.L91B_MA_BasketballMath_403_Screen_3_16657	1.04		0.70	0.96	0.94
36.L91B_MA_BasketballMath_403_Screen_4_16658	2.48		0.39	0.95	0.94
37.L91C_LA_Poetry_P100_Screen_2_12893	0.07	Yes	0.94	0.97	0.98
38.L91C_LA_Poetry_P100_Screen_3_12894	2.82	Yes	0.56	0.96	0.94
39.L91C_LA_Poetry_P100_Screen_4_12895	2.83	Yes	0.59	1.06	1.09
40.L91C_MA_WaterSale_P100_Screen_2_12720	1.60	Yes	0.81	0.95	0.88
41.L91C_MA_WaterSale_P100_Screen_3_12903	1.40	Yes	0.83	0.91	0.80
42.L91C_MA_WaterSale_P100_Screen_4_12907	2.18	Yes	0.73	0.93	0.88
43.L91C_SS_GlobalCoffeeProduction_P100_Screen_2_13036	1.07	Yes	0.88	0.94	0.86
44.L91C_SS_GlobalCoffeeProduction_P100_Screen_3_13037	2.66	Yes	0.68	0.95	0.93
45.L91C_SS_GlobalCoffeeProduction_P100_Screen_4_13038	3.60	Yes	0.48	1.02	1.02
46.L91B_SC_PlantLifeCycle_P100_Screen_2_11245	1.21	Yes	0.88	0.94	0.85
47.L91B_SC_PlantLifeCycle_P100_Screen_3_11343	2.65	Yes	0.68	0.92	0.88
48.L91B_SC_PlantLifeCycle_P100_Screen_4_11344	3.91	Yes	0.46	0.99	0.98
49.L91C_LA_LibraryVisit_403_Screen_2_16590	1.72		0.78	0.94	0.90
50.L91C_LA_LibraryVisit_403_Screen_3_16591	1.08		0.86	0.94	0.86
51.L91C_LA_LibraryVisit_403_Screen_4_16592	2.72		0.49	1.00	0.99
52.L91C_MA_angles_joga_P100_A202_Screen_2_13865	2.68	Yes	0.72	0.97	0.97
53.L91C_MA_angles_joga_P100_A202_Screen_3_13866	3.49	Yes	0.53	0.97	0.96
54.L91C_MA_angles_joga_P100_A202_Screen_4_13867	3.88	Yes	0.49	0.95	0.94

2.1.2 Reading

2.1.2.1 Grade 1

Table 2.1.2.1

Complete Item Analysis and Summary: Read 1 S403 Online

Item Type	Average Item Difficulty (in logits)	No. of Items	Average P-value	Average Infit Mean Square	Average Outfit Mean Square
Selected Response	-1.00	72	0.42	0.99	0.98

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	Outfit Mnsq
1.R01A_SI_GettingReady_P100_A203_Screen_2_13193	-2.11	Yes	0.65	0.94	0.84
2.R01A_SI_GettingReady_P100_A203_Screen_3_13194	-2.06	Yes	0.63	0.99	0.90
3.R01A_SI_GettingReady_P100_A203_Screen_4_13195	-1.52	Yes	0.56	1.12	1.20
4.R01B_SI_MorningMessage_P100_A203_Screen_2_13196	-1.11	Yes	0.48	1.09	1.18
5.R01B_SI_MorningMessage_P100_A203_Screen_3_13197	-1.54	Yes	0.54	1.09	1.11
6.R01B_SI_MorningMessage_P100_A203_Screen_4_13198	-0.64	Yes	0.42	1.31	1.49
7.R01A_LA_CatAdventure_203_P100_A301_Screen_2_13211	-2.30	Yes	0.37	0.98	0.99
8.R01A_LA_CatAdventure_203_P100_A301_Screen_3_13212	-1.67	Yes	0.29	0.98	0.97
9.R01A_LA_CatAdventure_203_P100_A301_Screen_4_13213	-2.28	Yes	0.38	1.00	1.00
10.R01A_MA_CountBalloons_301_P100_A301FT_Screen_2_13274	-4.07	Yes	0.76	1.04	1.08
11.R01A_MA_CountBalloons_301_P100_A301FT_Screen_3_13275	-1.10	Yes	0.14	0.98	0.98
12.R01A_MA_CountBalloons_301_P100_A301FT_Screen_4_13276	-1.13	Yes	0.15	0.96	0.93
13.R01C_SS_PetStore_P100_A202_Screen_2_13313	-4.46	Yes	0.83	0.99	0.98
14.R01C_SS_PetStore_P100_A202_Screen_3_13314	-2.42	Yes	0.33	0.97	0.96
15.R01C_SS_PetStore_P100_A202_Screen_4_13315	-2.58	Yes	0.36	0.98	0.97
16.R01B_SC_MoonCharacteristics_402_V1_Screen_2_15680	-4.32		0.76	0.97	0.95
17.R01B_SC_MoonCharacteristics_402_V1_Screen_3_15681	-2.77		0.43	0.97	0.97
18.R01B_SC_MoonCharacteristics_402_V1_Screen_4_15682	-2.77		0.42	0.96	0.95
19.R01A_LA_AVisitToThePond_P100_A203_Screen_2_13199	-1.94	Yes	0.26	0.96	0.95
20.R01A_LA_AVisitToThePond_P100_A203_Screen_3_13200	-2.17	Yes	0.35	0.92	0.90
21.R01A_LA_AVisitToThePond_P100_A203_Screen_4_13201	-1.09	Yes	0.16	0.95	0.90
22.R01A_MA_FamilyGatheringatBeach_jc_P100_A203_Screen_2_13202	-1.81	Yes	0.25	0.93	0.90
23.R01A_MA_FamilyGatheringatBeach_jc_P100_A203_Screen_3_13203	-1.75	Yes	0.27	1.00	1.02
24.R01A_MA_FamilyGatheringatBeach_jc_P100_A203_Screen_4_13204	-2.19	Yes	0.32	0.97	0.97
25.R01B_LA_ThankYou_402_V1_Screen_2_16039	-1.61	Yes	0.49	0.95	0.93

[redacted]

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	Fit Statistics Outfit Mnsq
61.R01C_LA_GrowingTaller_203_P100_A301_Screen_2_13241	0.50	Yes	0.43	0.83	0.79
62.R01C_LA_GrowingTaller_203_P100_A301_Screen_3_13242	0.57	Yes	0.40	0.83	0.79
63.R01C_LA_GrowingTaller_203_P100_A301_Screen_4_13243	0.47	Yes	0.42	1.06	1.06
64.R01C_MA_PuppetShow_dode_P100_A203_Screen_2_13256	0.57	Yes	0.43	1.04	1.04
65.R01C_MA_PuppetShow_dode_P100_A203_Screen_3_13257	0.61	Yes	0.43	0.98	0.96
66.R01C_MA_PuppetShow_dode_P100_A203_Screen_4_13258	0.14	Yes	0.53	0.97	0.93
67.R01B_SS_TwoSchools_403_Screen_2_17029	0.24		0.33	1.08	1.08
68.R01B_SS_TwoSchools_403_Screen_3_17030	-0.30		0.51	0.95	0.94
69.R01B_SS_TwoSchools_403_Screen_4_17031	0.11		0.40	1.01	1.00
70.R01B_SC_Fish_403_Screen_2_17131	-0.78		0.59	0.94	0.92
71.R01B_SC_Fish_403_Screen_3_17132	-0.43		0.55	0.91	0.88
72.R01B_SC_Fish_403_Screen_4_17133	0.35		0.34	1.04	1.03

2.1.2.2 *Grades 2–3*

Table 2.1.2.2

Complete Item Analysis and Summary: Read 2-3 S403 Online

[redacted]

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	 Outfit Mnsq
26.R 23B_LA_TreasureHunt_402_v1_Screen_3_15715	-0.16		0.40	1.03	1.03
27.R 23B_LA_TreasureHunt_402_v1_Screen_4_15716	0.36	Yes	0.33	1.00	1.00
28.R 23B_MA_RecyclingProject_402_v1_Screen_2_15626	-1.69	Yes	0.71	0.92	0.88
29.R 23B_MA_RecyclingProject_402_v1_Screen_3_15628	-0.79	Yes	0.51	0.96	0.95
30.R 23B_MA_RecyclingProject_402_v1_Screen_4_15629	0.42		0.27	1.06	1.07
31.R 23C_SS_UrbanNeighborhood_P100_A301_Screen_2_13371	-0.21	Yes	0.35	1.00	0.99
32.R 23C_SS_UrbanNeighborhood_P100_A301_Screen_3_13372	-1.27	Yes	0.63	0.90	0.87
33.R 23C_SS_UrbanNeighborhood_P100_A301_Screen_4_13373	-0.63	Yes	0.50	0.86	0.84
34.R 23B_SC_TempClimate_401_V1_Screen_2_13932	-0.08	Yes	0.39	1.00	0.99
35.R 23B_SC_TempClimate_401_V1_Screen_3_13936	0.45	Yes	0.24	1.02	1.02
36.R 23B_SC_TempClimate_401_V1_Screen_4_13937	0.42	Yes	0.26	1.05	1.07
37.R 23B_LA_ZebraStripes_203_P100_A301_altl_Screen_2_13410	-0.34	Yes	0.43	0.95	0.94
38.R 23B_LA_ZebraStripes_203_P100_A301_altl_Screen_3_13411	0.21	Yes	0.31	0.99	0.98
39.R 23B_LA_ZebraStripes_203_P100_A301_altl_Screen_4_13412	-0.11	Yes	0.39	0.93	0.91
40.R 23B_MA_Library_P100_A201_altl_Screen_2_13413	-0.47	Yes	0.49	0.90	0.89
41.R 23B_MA_Library_P100_A201_altl_Screen_3_13414	-0.76	Yes	0.47	0.93	0.93
42.R 23B_MA_Library_P100_A201_altl_Screen_4_13415	-0.14	Yes	0.43	0.97	0.97
43.R 23B_SS_ArtProducts_403_Screen_2_17041	-0.20		0.47	0.94	0.93
44.R 23B_SS_ArtProducts_403_Screen_3_17042	0.70		0.36	1.01	1.00
45.R 23B_SS_ArtProducts_403_Screen_4_17043	0.08		0.42	0.95	0.95
46.R 23C_SC_Pond_403_Screen_2_17153	-0.51		0.51	0.95	0.95
47.R 23C_SC_Pond_403_Screen_3_17154	-0.04		0.43	0.97	0.97
48.R 23C_SC_Pond_403_Screen_4_17155	0.99		0.35	1.03	1.03
49.R 23A_LA_KittenFable_203_P100_A203FT_Screen_2_13338	0.80	Yes	0.50	1.14	1.20
50.R 23A_LA_KittenFable_203_P100_A203FT_Screen_3_13339	0.38	Yes	0.57	0.86	0.82
51.R 23A_LA_KittenFable_203_P100_A203FT_Screen_4_13340	-0.25	Yes	0.71	0.83	0.73
52.R 23C_MA_ReadingProject_402_v2_Screen_2_16092	1.27		0.43	1.05	1.06
53.R 23C_MA_ReadingProject_402_v2_Screen_3_16094	0.90		0.50	1.03	1.02
54.R 23C_MA_ReadingProject_402_v2_Screen_4_16095	0.70		0.54	1.04	1.04
55.R 23B_SS_BoatsLongAgo_401_V2_Screen_2_14583	0.05	Yes	0.68	0.90	0.85
56.R 23B_SS_BoatsLongAgo_401_V2_Screen_3_14584	0.92	Yes	0.53	1.08	1.10
57.R 23B_SS_BoatsLongAgo_401_V2_Screen_4_14585	0.64	Yes	0.53	0.97	0.95
58.R 23B_SC_AnimalMovement_dode_P100_A301_altl_Screen_2_13419	-1.67	Yes	0.93	0.93	0.73
59.R 23B_SC_AnimalMovement_dode_P100_A301_altl_Screen_3_13420	-0.90	Yes	0.84	0.91	0.83
60.R 23B_SC_AnimalMovement_dode_P100_A301_altl_Screen_4_13421	0.70	Yes	0.62	0.94	0.92

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	Outfit Mnsq
61.R23B_LA_Zebra Stripes_203_P100_A301_Screen_2_13344	0.33	Yes	0.69	0.93	0.90
62.R23B_LA_Zebra Stripes_203_P100_A301_Screen_3_13345	1.24	Yes	0.48	0.91	0.89
63.R23B_LA_Zebra Stripes_203_P100_A301_Screen_4_13346	0.60	Yes	0.65	0.89	0.84
64.R23B_MA_Library_P100_A201_Screen_2_13356	0.07	Yes	0.69	0.89	0.84
65.R23B_MA_Library_P100_A201_Screen_3_13357	-0.13	Yes	0.72	0.89	0.82
66.R23B_MA_Library_P100_A201_Screen_4_13358	0.54	Yes	0.63	0.96	0.94
67.R23C_SS_ShopsInTown_403_Screen_2_17049	0.73		0.48	1.05	1.05
68.R23C_SS_ShopsInTown_403_Screen_3_17050	0.71		0.47	1.03	1.03
69.R23C_SS_ShopsInTown_403_Screen_4_17051	-0.01		0.67	0.92	0.88
70.R23C_SC_Insects_dode_P100_A301_Screen_2_13374	-0.30	Yes	0.78	0.92	0.87
71.R23C_SC_Insects_dode_P100_A301_Screen_3_13375	1.34	Yes	0.50	0.96	0.94
72.R23C_SC_Insects_dode_P100_A301_Screen_4_13376	0.66	Yes	0.62	1.04	1.04

2.1.2.3 *Grades 4–5*

Table 2.1.2.3

Complete Item Analysis and Summary: Read 4-5 S403 Online

Item Type	Average Item Difficulty (in logits)	No. of Items	Average P-value	Average Infit Mean Square	Average Outfit Mean Square
Selected Response	0.78	69	0.49	0.98	0.99

				Fit Statistics	
Name	**Item Difficulty (in logits)**	**Anchored?**	**P-value**	**Infit Mnsq**	**Outfit Mnsq**
1.R 45A_SI_WatchingTV_P 100_A301FT_Screen_2_13509	-3.19	Yes	0.98	0.88	0.50
2.R 45A_SI_WatchingTV_P 100_A301FT_Screen_3_13510	-1.06	Yes	0.87	0.95	0.99
3.R 45A_SI_WatchingTV_P 100_A301FT_Screen_4_13511	1.08	Yes	0.57	1.06	1.11
4.R 45B_SI_PhotoContest_P 100_A301FT_Screen_2_13449	0.25	Yes	0.69	0.96	0.98
5.R 45B_SI_PhotoContest_P 100_A301FT_Screen_3_13450	1.77	Yes	0.42	1.20	1.51
6.R 45B_SI_PhotoContest_P 100_A301FT_Screen_4_13451	-0.11	Yes	0.75	1.09	1.31
7.R 45A_LA_RecycleOrCompost_402_V2_Screen_2_15991	-1.81	Yes	0.68	0.92	0.86
8.R 45A_LA_RecycleOrCompost_402_V2_Screen_3_15992	-0.43		0.39	1.08	1.10
9.R 45A_LA_RecycleOrCompost_402_V2_Screen_4_15993	0.01	Yes	0.36	0.93	0.92
10.R 45A_MA_Supermarket_pasz_P 100_A203_Screen_2_13464	-1.61	Yes	0.63	0.94	0.91
11.R 45A_MA_Supermarket_pasz_P 100_A203_Screen_3_13465	-0.93	Yes	0.46	0.95	0.93
12.R 45A_MA_Supermarket_pasz_P 100_A203_Screen_4_13466	-0.43	Yes	0.39	0.97	0.96
13.R 45A_SS_ShoppingCart_401_V2_Screen_2_14714	-0.37	Yes	0.33	0.90	0.87
14.R 45A_SS_ShoppingCart_401_V2_Screen_3_14715	-0.75	Yes	0.41	1.05	1.06
15.R 45A_SS_ShoppingCart_401_V2_Screen_4_14716	0.20	Yes	0.21	0.98	0.97
16.R 45B_SC_Geysers_P 100_A301_45_Screen_2_13407	-0.72	Yes	0.45	0.98	0.97
17.R 45B_SC_Geysers_P 100_A301_45_Screen_3_13408	0.06	Yes	0.22	1.02	1.12
18.R 45B_SC_Geysers_P 100_A301_45_Screen_4_13409	-0.17	Yes	0.28	0.98	0.97
19.R 45B_LA_OntheTrain_P 100_A202_Screen_2_13467	0.47	Yes	0.44	1.07	1.14
20.R 45B_LA_OntheTrain_P 100_A202_Screen_3_13468	1.31	Yes	0.25	1.09	1.34
21.R 45B_LA_OntheTrain_P 100_A202_Screen_4_13469	1.15	Yes	0.35	1.08	1.22
22.R 45A_MA_BuyingSchoolSupplies_P 100_A201_Screen_2_13518	-1.94	Yes	0.66	0.93	0.89
23.R 45A_MA_BuyingSchoolSupplies_P 100_A201_Screen_3_13519	-1.12	Yes	0.49	0.93	0.91
24.R 45A_MA_BuyingSchoolSupplies_P 100_A201_Screen_4_13520	-0.31	Yes	0.32	0.93	0.92
25.R 45C_LA_CookingWithGrandpa_402_V1_Screen_2_16009	-0.10	Yes	0.63	0.87	0.81

[redacted]

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	Outfit Mnsq
61.R45B_MA_ClassSchedule_jc_P100_A201_Alt1_Screen_2_13533	1.43	Yes	0.67	0.94	0.91
62.R45B_MA_ClassSchedule_jc_P100_A201_Alt1_Screen_3_13534	0.99	Yes	0.72	0.95	0.94
63.R45B_MA_ClassSchedule_jc_P100_A201_Alt1_Screen_4_13535	2.38	Yes	0.44	0.93	0.91
64.R45C_SS_EleanorRoosevelt_203_P100_A301_Screen_2_13503	2.15	Yes	0.48	0.96	0.95
65.R45C_SS_EleanorRoosevelt_203_P100_A301_Screen_3_13504	2.57	Yes	0.45	0.97	0.96
66.R45C_SS_EleanorRoosevelt_203_P100_A301_Screen_4_13505	1.96	Yes	0.57	0.92	0.91
67.R45C_SC_QuakingAspen_403_Screen_2_17107	1.54		0.58	1.01	1.03
68.R45C_SC_QuakingAspen_403_Screen_3_17112	2.14		0.44	0.95	0.94
69.R45C_SC_QuakingAspen_403_Screen_4_17113	2.30		0.40	1.04	1.04

2.1.2.4 *Grades 6–8*

Table 2.1.2.4

Complete Item Analysis and Summary: Read 6-8 S403 Online

[redacted]

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	 Outfit Mnsq
26.R 68C_LA_Team Leader_402_V2_Screen_3_16023	0.24	Yes	0.67	0.93	0.95
27.R 68C_LA_Team Leader_402_V2_Screen_4_16024	-0.03	Yes	0.69	0.88	0.84
28.R 68C_MA_Ratios_401_V1_Screen_2_13962	1.39	Yes	0.48	0.97	0.96
29.R 68C_MA_Ratios_401_V1_Screen_3_13963	1.94	Yes	0.36	1.00	1.01
30.R 68C_MA_Ratios_401_V1_Screen_4_13964	2.02	Yes	0.31	1.02	1.06
31.R 68B_SS_RockAndRoll_203_P100_A301_Screen_2_13584	1.26	Yes	0.44	1.01	1.01
32.R 68B_SS_RockAndRoll_203_P100_A301_Screen_3_13585	0.57	Yes	0.66	0.91	0.87
33.R 68B_SS_RockAndRoll_203_P100_A301_Screen_4_13586	0.61	Yes	0.60	0.95	0.93
34.R 68B_SC_CellDivision_diab_P100_A301_Screen_2_13587	0.31	Yes	0.72	0.95	0.92
35.R 68B_SC_CellDivision_diab_P100_A301_Screen_3_13588	0.78	Yes	0.62	0.94	0.93
36.R 68B_SC_CellDivision_diab_P100_A301_Screen_4_13589	1.93	Yes	0.35	1.00	1.00
37.R 68C_LA_SolarOven_203_P100_A203_Screen_2_13602	1.09	Yes	0.51	0.95	0.93
38.R 68C_LA_SolarOven_203_P100_A203_Screen_3_13603	1.01	Yes	0.53	0.92	0.91
39.R 68C_LA_SolarOven_203_P100_A203_Screen_4_13604	1.87	Yes	0.39	0.92	0.91
40.R 68B_MA_newroom_diab_P100_A203_alt1_Screen_2_13650	1.09	Yes	0.56	0.94	0.94
41.R 68B_MA_newroom_diab_P100_A203_alt1_Screen_3_13651	2.53	Yes	0.24	0.97	0.96
42.R 68B_MA_newroom_diab_P100_A203_alt1_Screen_4_13652	2.41	Yes	0.25	1.03	1.05
43.R 68B_SS_Competition_403_Screen_2_17012	1.73		0.48	0.97	0.97
44.R 68B_SS_Competition_403_Screen_3_17014	2.56		0.31	1.00	1.01
45.R 68B_SS_Competition_403_Screen_4_17015	2.41		0.30	0.95	0.94
46.R 68B_SC_SunLayers_403_Screen_2_17127	2.11		0.36	1.00	1.00
47.R 68B_SC_SunLayers_403_Screen_3_17128	2.61		0.31	1.03	1.04
48.R 68B_SC_SunLayers_403_Screen_4_17129	3.21		0.21	1.02	1.03
49.R 68C_LA_MyPenPal_P100_A202_Screen_2_13659	1.26	Yes	0.66	0.95	0.97
50.R 68C_LA_MyPenPal_P100_A202_Screen_3_13660	1.89	Yes	0.54	0.99	1.00
51.R 68C_LA_MyPenPal_P100_A202_Screen_4_13661	1.22	Yes	0.66	1.06	1.16
52.R 68C_MA_Ratios_401_V2_Screen_2_14616	2.52	Yes	0.46	0.94	0.91
53.R 68C_MA_Ratios_401_V2_Screen_3_14617	3.66		0.26	1.01	1.03
54.R 68C_MA_Ratios_401_V2_Screen_4_14618	2.75	Yes	0.42	0.97	0.97
55.R 68C_SS_Nile River_401_V1_Screen_2_14491	1.72	Yes	0.70	0.92	0.89
56.R 68C_SS_Nile River_401_V1_Screen_3_14492	1.78	Yes	0.64	1.03	1.07
57.R 68C_SS_Nile River_401_V1_Screen_4_14493	2.61	Yes	0.47	0.98	0.97
58.R 68B_SC_RockCycle_DeRoSaJa_P100_A301_alt1_Screen_2_13656	2.35	Yes	0.59	0.97	0.96
59.R 68B_SC_RockCycle_DeRoSaJa_P100_A301_alt1_Screen_3_13657	1.62	Yes	0.77	1.00	1.04
60.R 68B_SC_RockCycle_DeRoSaJa_P100_A301_alt1_Screen_4_13658	3.59	Yes	0.34	1.05	1.06

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	Fit Statistics Outfit Mnsq
61.R68C_LA_Prodigy_203_P100_A301_Screen_2_13614	1.95	Yes	0.67	0.85	0.79
62.R68C_LA_Prodigy_203_P100_A301_Screen_3_13615	2.21	Yes	0.62	0.87	0.83
63.R68C_LA_Prodigy_203_P100_A301_Screen_4_13616	2.67	Yes	0.52	0.93	0.92
64.R68C_MA_FoodCollection_RaObAmSc_P100_A203_Screen_2_13617	3.33	Yes	0.39	1.03	1.03
65.R68C_MA_FoodCollection_RaObAmSc_P100_A203_Screen_3_13618	3.13	Yes	0.45	0.98	0.97
66.R68C_MA_FoodCollection_RaObAmSc_P100_A203_Screen_4_13619	3.03	Yes	0.43	0.99	0.98
67.R68C_SS_Spotlight_403_Screen_2_17025	3.22		0.30	1.03	1.03
68.R68C_SS_Spotlight_403_Screen_3_17026	2.04		0.64	0.91	0.88
69.R68C_SS_Spotlight_403_Screen_4_17027	2.63		0.43	0.98	0.98
70.R68C_SC_PlanetComposition_403_Screen_2_17123	3.19		0.29	0.98	0.98
71.R68C_SC_PlanetComposition_403_Screen_3_17124	2.55		0.48	0.95	0.94
72.R68C_SC_PlanetComposition_403_Screen_4_17125	3.20		0.27	1.04	1.06

2.1.2.5 Grades 9–12

Table 2.1.2.5

Complete Item Analysis and Summary: Read 9-12 S403 Online

[redacted]

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	Outfit Mnsq
26.R91B_LA_Fiction_402_V2_Screen_3_16059	0.67	Yes	0.74	0.86	0.78
27.R91B_LA_Fiction_402_V2_Screen_4_16060	2.00	Yes	0.42	1.05	1.10
28.R91B_MA_CakeMaking_JaGeKaKo_P100_A203_alt1_Screen_2_13767	-0.89	Yes	0.90	0.92	0.80
29.R91B_MA_CakeMaking_JaGeKaKo_P100_A203_alt1_Screen_3_13768	2.02	Yes	0.43	1.00	0.98
30.R91B_MA_CakeMaking_JaGeKaKo_P100_A203_alt1_Screen_4_13769	2.14	Yes	0.44	1.00	0.99
31.R91B_SS_Reading_203_P100_A301_Screen_2_13698	1.20	Yes	0.64	1.01	1.01
32.R91B_SS_Reading_203_P100_A301_Screen_3_13699	1.48	Yes	0.60	0.93	0.91
33.R91B_SS_Reading_203_P100_A301_Screen_4_13700	2.04	Yes	0.44	0.97	0.96
34.R91C_SC_PlantsInSpace_402_V2_Screen_2_16070	2.65	Yes	0.35	0.97	0.95
35.R91C_SC_PlantsInSpace_402_V2_Screen_3_16071	2.36	Yes	0.41	0.94	0.93
36.R91C_SC_PlantsInSpace_402_V2_Screen_4_16072	3.26	Yes	0.26	1.00	1.00
37.R91B_LA_Mattie_203_P100_A301_Screen_2_13704	0.80	Yes	0.69	0.89	0.85
38.R91B_LA_Mattie_203_P100_A301_Screen_3_13705	1.29	Yes	0.61	0.91	0.89
39.R91B_LA_Mattie_203_P100_A301_Screen_4_13706	1.94	Yes	0.45	0.95	0.94
40.R91B_MA_CircumferenceandDiameter_ryro_P100_A203_Screen_2_13707	1.60	Yes	0.54	0.93	0.92
41.R91B_MA_CircumferenceandDiameter_ryro_P100_A203_Screen_3_13708	2.16	Yes	0.47	0.94	0.93
42.R91B_MA_CircumferenceandDiameter_ryro_P100_A203_Screen_4_13709	2.72	Yes	0.33	0.99	0.98
43.R91C_SS_BillOfRights_403_Screen_2_17071	2.49		0.43	0.93	0.92
44.R91C_SS_BillOfRights_403_Screen_3_17072	3.50		0.31	1.00	0.99
45.R91C_SS_BillOfRights_403_Screen_4_17073	2.87		0.43	0.97	0.97
46.R91B_SC_Electromagnetism_403_Screen_2_17079	1.91		0.56	0.96	0.95
47.R91B_SC_Electromagnetism_403_Screen_3_17080	3.60		0.29	0.98	0.97
48.R91B_SC_Electromagnetism_403_Screen_4_17081	3.78		0.25	0.98	0.97
49.R91C_LA_Drama_402_V1_Screen_2_14975	1.87	Yes	0.80	0.94	0.92
50.R91C_LA_Drama_402_V1_Screen_3_14976	2.99		0.62	1.01	1.00
51.R91C_LA_Drama_402_V1_Screen_4_14977	4.01		0.39	1.20	1.31
52.R91C_MA_Functions_401_V1_Screen_2_13968	2.61	Yes	0.66	1.03	1.05
53.R91C_MA_Functions_401_V1_Screen_3_13969	3.47	Yes	0.52	1.02	1.01
54.R91C_MA_Functions_401_V1_Screen_4_13970	2.85	Yes	0.60	1.06	1.08
55.R91C_SS_Mesopotamia_401_V2_Screen_2_14634	2.09	Yes	0.75	0.94	0.92
56.R91C_SS_Mesopotamia_401_V2_Screen_3_14635	3.35	Yes	0.54	0.91	0.89
57.R91C_SS_Mesopotamia_401_V2_Screen_4_14636	3.11	Yes	0.59	0.99	0.98
58.R91C_SC_pHScale_P100_A201_Screen_2_13785	2.54	Yes	0.68	0.93	0.90
59.R91C_SC_pHScale_P100_A201_Screen_3_13786	4.30	Yes	0.37	0.98	0.98
60.R91C_SC_pHScale_P100_A201_Screen_4_13787	3.69	Yes	0.47	1.00	0.99

Name	Item Difficulty (in logits)	Anchored?	P-value	Fit Statistics Infit Mnsq	 Outfit Mnsq
61.R91C_LA_Malta_203_P100_A301_L_Screen_2_13731	3.95	Yes	0.45	1.02	1.02
62.R91C_LA_Malta_203_P100_A301_L_Screen_3_13732	2.92	Yes	0.62	0.90	0.87
63.R91C_LA_Malta_203_P100_A301_L_Screen_4_13733	2.87	Yes	0.71	0.94	0.91
64.R91C_MA_CircumferenceandDiameter_ryro_P100_A203_Screen_2_13722	2.50	Yes	0.76	0.92	0.86
65.R91C_MA_CircumferenceandDiameter_ryro_P100_A203_Screen_3_13723	2.82	Yes	0.70	0.91	0.86
66.R91C_MA_CircumferenceandDiameter_ryro_P100_A203_Screen_4_13724	3.96	Yes	0.44	1.00	0.99
67.R91C_SS_DeclarationOfIndependence_403_Screen_2_17075	3.24		0.45	1.03	1.03
68.R91C_SS_DeclarationOfIndependence_403_Screen_3_17076	2.40		0.74	0.90	0.84
69.R91C_SS_DeclarationOfIndependence_403_Screen_4_17077	4.02		0.33	1.00	0.98
70.R91C_SC_ElasticEnergy_403_Screen_2_17091	2.74		0.64	0.95	0.94
71.R91C_SC_ElasticEnergy_403_Screen_3_17092	3.33		0.41	1.06	1.06
72.R91C_SC_ElasticEnergy_403_Screen_4_17093	3.59		0.44	0.98	0.97

2.1.3 Writing

2.1.3.1 *Grade 1*

Table 2.1.3.1.1

Complete Task Analysis and Summary: Writ 1 A S403 Online

Task Type	**Average Task Difficulty (in logits)**	**No. of Tasks**	**Average Infit Mean Square**	**Average Outfit Mean Square**
Constructed Response	-2.17	4	0.68	0.94

Name	**Task Difficulty (in logits)**	**Anchored?**	**Fit Statistics**	
			Infit Mnsq	**Outfit Mnsq**
1.W01A_SI_PartAH8_202_P100_A302_14245	-7.22	Yes	1.35	2.43
2.W01A_SI_GoingToSchoolPartB_402_14775	-1.11	Yes	0.72	0.71
3.W01A_SI_GoingToSchoolPartC_402_15919	-0.07	Yes	0.58	0.57
4.W01A_SI_PartDSentencesAboutMe_P100_A302_14248	-0.27	Yes	0.08	0.08

Raw Score Distribution by Task	**Raw Score**	**Task 1**	**Task 2**	**Task 3**	**Task 4**
	0	2.54%	5.61%	8.60%	5.81%
	1	97.46%	91.82%	7.44%	6.84%
	2	N/A	1.30%	12.56%	13.91%
	3	N/A	1.27%	24.30%	23.78%
	4	N/A	N/A	45.86%	43.96%
	5	N/A	N/A	1.09%	4.93%
	6	N/A	N/A	0.14%	0.77%
	7	N/A	N/A	0.00%	0.00%
	8	N/A	N/A	0.00%	0.00%
	9	N/A	N/A	0.00%	0.00%

Table 2.1.3.1.2

Complete Task Analysis and Summary: Writ 1 B/C S403 Online

Task Type		**Average Task Difficulty (in logits)**	**No. of Tasks**	**Average Infit Mean Square**	**Average Outfit Mean Square**
Constructed Response		-0.10	3	0.54	0.52
				Fit Statistics	
Name		**Task Difficulty (in logits)**	**Anchored?**	**Infit Mnsq**	**Outfit Mnsq**
1.W01B_SI_SpecialMemory_403_16402		0.05		0.80	0.78
2.W01C_MS_FlowerFruit_402_14960		0.25	Yes	0.53	0.49
3.W01C_IT_GoingToLibrary_403_16276		-0.60		0.29	0.27
Raw Score Distribution by Task	**Raw Score**	**Task 1**	**Task 2**	**Task 3**	
	0	0.27%	1.05%	0.55%	
	1	0.40%	1.47%	0.40%	
	2	3.58%	3.44%	1.13%	
	3	3.73%	7.27%	2.71%	
	4	44.74%	32.06%	19.89%	
	5	35.86%	42.28%	50.47%	
	6	10.44%	11.42%	21.46%	
	7	0.97%	0.96%	3.10%	
	8	0.03%	0.03%	0.28%	
	9	0.00%	0.00%	0.01%	

2.1.3.2 Grades 2–3

Table 2.1.3.2.1

Complete Task Analysis and Summary: Writ 2-3 A S403 Online

Task Type	**Average Task Difficulty (in logits)**	**No. of Tasks**	**Average Infit Mean Square**	**Average Outfit Mean Square**
Constructed Response	0.14	3	0.42	0.41

Name	**Task Difficulty (in logits)**	**Anchored?**	**Fit Statistics**	
			Infit Mnsq	**Outfit Mnsq**
1.W23A_SI_AtTheLibrary_403_16405	0.01		0.54	0.53
2.W23A_LA_ComputerArt_403_16448	0.13		0.34	0.34
3.W23A_MS_AnimalHomes_402_14792	0.27	Yes	0.38	0.38

Raw Score Distribution by Task	**Raw Score**	**Task 1**	**Task 2**	**Task 3**
	0	6.95%	7.72%	9.75%
	1	9.83%	12.18%	10.01%
	2	10.23%	8.68%	9.07%
	3	10.47%	21.05%	32.13%
	4	28.60%	39.87%	34.39%
	5	30.83%	9.66%	4.33%
	6	3.09%	0.79%	0.29%
	7	0.00%	0.05%	0.03%
	8	0.00%	0.00%	0.00%
	9	0.00%	0.00%	0.00%

Table 2.1.3.2.2

Complete Task Analysis and Summary: Writ 2-3 B/C S403 Online

Task Type	Average Task Difficulty (in logits)	No. of Tasks	Average Infit Mean Square	Average Outfit Mean Square
Constructed Response	0.86	3	0.49	0.48

			Fit Statistics	
Name	Task Difficulty (in logits)	Anchored?	Infit Mnsq	Outfit Mnsq
1.W23B_SI_WorkChoices_403_16407	0.84		0.69	0.69
2.W23C_MS_HeartRateChanges_402_15549	0.94	Yes	0.46	0.46
3.W23C_IT_ClassEvent_403_16450	0.80		0.32	0.30

Raw Score Distribution by Task	Raw Score	Task 1	Task 2	Task 3
	0	0.28%	0.42%	0.73%
	1	0.43%	0.84%	0.56%
	2	2.20%	2.88%	1.62%
	3	8.46%	9.18%	6.48%
	4	38.55%	41.00%	26.79%
	5	37.00%	36.23%	34.72%
	6	11.79%	8.49%	22.50%
	7	1.23%	0.91%	5.59%
	8	0.05%	0.05%	0.89%
	9	0.00%	0.00%	0.12%

2.1.3.3 *Grades 4–5*

Table 2.1.3.3.1

Complete Task Analysis and Summary: Writ 4-5 A S403 Online

[redacted]

Table 2.1.3.3.2

Complete Task Analysis and Summary: Writ 4-5 B/C S403 Online

		Average		Average	Average
Task Type					
Constructed Response		1.85	3	0.55	0.53
				Fit Statistics	
Name		**Task Difficulty (in logits)**	**Anchored?**	**Infit Mnsq**	**Outfit Mnsq**
1.W45B_SI_LibraryRules_403_Screen_2_16344_17244		1.39		0.85	0.82
2.W45C_MS_TulipGrowth_402_Screen_4_15416_17207		2.00	Yes	0.51	0.49
3.W45C_IT_AncientChina_401_Screen_5_14477_17232		2.17		0.29	0.28
Raw Score Distribution by Task	**Raw Score**	**Task 1**	**Task 2**	**Task 3**	
	0	0.77%	0.99%	1.31%	
	1	0.51%	0.79%	0.64%	
	2	2.46%	2.97%	2.91%	
	3	4.56%	6.73%	11.17%	
	4	18.52%	29.15%	30.48%	
	5	39.41%	43.15%	37.30%	
	6	30.64%	14.52%	13.83%	
	7	2.97%	1.58%	2.19%	
	8	0.16%	0.11%	0.16%	
	9	0.01%	0.01%	0.01%	

2.1.3.4 Grades 6–8

Table 2.1.3.4.1

Complete Task Analysis and Summary: Writ 6-8 A S403 Online

Task Type	Average Task Difficulty (in logits)	No. of Tasks	Average Infit Mean Square	Average Outfit Mean Square
Constructed Response	0.93	3	0.40	0.43

Name	Task Difficulty (in logits)	Anchored?	Fit Statistics	
			Infit Mnsq	Outfit Mnsq
1.W68A_SI_StudyPrefs_403_Screen_4_16426_17254	0.89		0.38	0.41
2.W68A_LA_AccidentalSwap_403_Screen_4_16400_17249	0.48		0.40	0.43
3.W68A_MS_HeightComparison_402_Screen_4_14786_17188	1.41	Yes	0.41	0.46

Raw Score Distribution by Task	Raw Score	Task 1	Task 2	Task 3
	0	4.73%	4.42%	8.03%
	1	8.00%	6.40%	11.97%
	2	10.19%	7.61%	14.90%
	3	22.84%	17.74%	30.23%
	4	31.81%	29.03%	24.75%
	5	17.44%	25.66%	8.38%
	6	4.29%	7.60%	1.53%
	7	0.67%	1.36%	0.18%
	8	0.02%	0.15%	0.01%
	9	0.00%	0.02%	0.00%

Table 2.1.3.4.2

Complete Task Analysis and Summary: Writ 6-8 B/C S403 Online

Task Type	Average Task Difficulty (in logits)	No. of Tasks	Average Infit Mean Square	Average Outfit Mean Square
Constructed Response	1.40	3	0.52	0.49

			Fit Statistics	
Name	Task Difficulty (in logits)	Anchored?	Infit Mnsq	Outfit Mnsq
1.W68B_SI_SchoolComputers_403_Screen_2_16369_17257	1.13		0.68	0.65
2.W68C_MS_FishpH_402_Screen_5_15418_17176	1.49	Yes	0.57	0.55
3.W68C_IT_BookMovie_403_Screen_4_16388_17262	1.56		0.30	0.28

Raw Score Distribution by Task	Raw Score	Task 1	Task 2	Task 3
	0	0.14%	0.46%	0.41%
	1	0.16%	0.70%	0.27%
	2	1.13%	1.71%	2.03%
	3	2.21%	3.28%	4.99%
	4	10.56%	14.77%	17.28%
	5	45.81%	45.97%	40.66%
	6	34.69%	29.25%	29.74%
	7	4.61%	3.40%	3.85%
	8	0.61%	0.38%	0.65%
	9	0.09%	0.08%	0.12%

2.1.3.5 *Grades 9–12*

Table 2.1.3.5.1

Complete Task Analysis and Summary: Writ 9-12 A S403 Online

Task Type	**Average Task Difficulty (in logits)**	**No. of Tasks**	**Average Infit Mean Square**	**Average Outfit Mean Square**
Constructed Response	2.00	3	0.53	0.56

Name	**Task Difficulty (in logits)**	**Anchored?**	**Fit Statistics**	
			Infit Mnsq	**Outfit Mnsq**
1.W91A_SI_HealthEvents_403_Screen_3_16374_17272	2.19		0.52	0.56
2.W91A_LA_Sprinkler_403_Screen_4_11985_17268	1.64		0.49	0.51
3.W91A_MS_WaterTemp_402_Screen_5_15555_17166	2.17	Yes	0.58	0.62

Raw Score Distribution by Task	**Raw Score**	**Task 1**	**Task 2**	**Task 3**
	0	6.64%	4.67%	5.52%
	1	8.57%	4.96%	7.62%
	2	11.20%	6.81%	10.81%
	3	18.05%	13.48%	18.90%
	4	22.69%	20.69%	22.22%
	5	21.36%	29.91%	21.81%
	6	9.17%	15.75%	10.44%
	7	2.17%	3.49%	2.45%
	8	0.16%	0.23%	0.23%
	9	0.01%	0.02%	0.01%

Table 2.1.3.5.2

Complete Task Analysis and Summary: Writ 9-12 B/C S403 Online

Task Type	Average Task Difficulty (in logits)	No. of Tasks	Average Infit Mean Square	Average Outfit Mean Square
Constructed Response	1.85	3	0.72	0.69

Name	Task Difficulty (in logits)	Anchored?	Fit Statistics	
			Infit Mnsq	Outfit Mnsq
1.W91B_SI_Scholarship_403_Screen_2_16384_17275	1.42		1.00	0.96
2.W91B_MS_Viscosity_P100_A203_Screen_2_14281_14652	1.83	Yes	0.71	0.69
3.W91C_IT_Architecture_402_Screen_4_16506_17280	2.32		0.45	0.42

Raw Score Distribution by Task	Raw Score	Task 1	Task 2	Task 3
	0	0.64%	1.02%	1.77%
	1	0.18%	0.21%	0.37%
	2	0.76%	1.10%	2.31%
	3	2.03%	2.37%	6.63%
	4	7.66%	9.40%	16.50%
	5	27.78%	29.87%	31.40%
	6	40.06%	38.11%	28.30%
	7	17.54%	15.33%	10.40%
	8	3.13%	2.45%	2.07%
	9	0.22%	0.14%	0.25%

2.1.4 Speaking

2.1.4.1 Grade 1

Table 2.1.4.1

Complete Task Analysis and Summary: Spek 1 S403 Online

Task Type	Average Task Difficulty (in logits)	No. of Tasks	Average Infit Mean Square	Average Outfit Mean Square
Constructed Response	-1.55	9	0.84	0.71

Name	Tier	Task Difficulty (in logits)	Anchored?	Fit Statistics	
				Infit Mnsq	Outfit Mnsq
1.S01AP_SI_ClassroomJobs_403_Screen_3_P1_16137	Pre-A / A	-4.20	Yes	1.03	0.67
2.S01AC_SI_ClassroomJobs_403_Screen_7_P3_16149	A / BC	-0.52	Yes	0.78	0.79
3.S01C_SI_ClassroomJobs_403_Screen_11_P5_16158	BC	0.27	Yes	0.82	0.84
4.S01AP_LS_ArtTeacher_403_Screen_3_P1_16140	Pre-A / A	-4.42	Yes	0.86	0.38
5.S01AC_LS_ArtTeacher_403_Screen_9_P3_16262	A / BC	-0.30	Yes	0.76	0.78
6.S01C_LS_ArtTeacher_403_Screen_15_P5_16268	BC	-0.20	Yes	0.95	0.97
7.S01AP_MS_Elephants_402_Screen_4_P1_15004	Pre-A / A	-4.62	Yes	0.87	0.42
8.S01AC_MS_Elephants_402_Screen_11_P3_15011	A / BC	-0.07	Yes	0.76	0.78
9.S01C_MS_Elephants_402_Screen_19_P5_15019	BC	0.13	Yes	0.76	0.77

Raw Score Distribution by Task	Task	Raw Score				
		0	1	2	3	4
	Task 1	9.74%	5.03%	85.23%	NA	NA
	Task 2	5.82%	14.77%	49.60%	26.36%	3.44%
	Task 3	4.44%	26.69%	43.88%	22.09%	2.91%
	Task 4	8.97%	5.69%	85.34%	NA	NA
	Task 5	6.84%	17.01%	47.02%	25.53%	3.61%
	Task 6	6.00%	18.02%	38.90%	32.02%	5.05%
	Task 7	9.24%	5.26%	85.50%	NA	NA
	Task 8	6.49%	15.03%	53.45%	21.30%	3.73%
	Task 9	4.78%	17.26%	57.77%	17.99%	2.20%

2.1.4.2 *Grades 2–3*

Table 2.1.4.2

Complete Task Analysis and Summary: Spek 2-3 S403 Online

Task Type	**Average Task Difficulty (in logits)**	**No. of Tasks**	**Average Infit Mean Square**	**Average Outfit Mean Square**
Constructed Response	-1.66	9	0.81	0.63

				Fit Statistics	
Name	**Tier**	**Task Difficulty (in logits)**	**Anchored?**	**Infit Mnsq**	**Outfit Mnsq**
1.S23AP_SI_SpecialChoice_403_Screen_3_P1_16207	Pre-A / A	-4.71	Yes	0.97	0.39
2.S23AC_SI_SpecialChoice_403_Screen_7_P3_16211	A / BC	-0.49	Yes	0.82	0.82
3.S23C_SI_SpecialChoice_403_Screen_11_P5_16215	BC	0.63	Yes	0.74	0.79
4.S23AP_LS_StoreManager_402_Screen_3_P1_15106	Pre-A / A	-4.75	Yes	0.87	0.33
5.S23AC_LS_StoreManager_402_Screen_8_P3_15112	A / BC	-0.24	Yes	0.71	0.69
6.S23C_LS_StoreManager_402_Screen_14_P5_15119	BC	-1.16	Yes	0.77	0.80
7.S23AP_MS_RampHeight_402_Screen_4_P1_15085	Pre-A / A	-4.60	Yes	0.88	0.28
8.S23AC_MS_RampHeight_402_Screen_11_P3_15092	A / BC	-0.31	Yes	0.71	0.72
9.S23C_MS_RampHeight_402_Screen_18_P5_15102	BC	0.66	Yes	0.82	0.84

Raw Score Distribution by Task	**Task**			**Raw Score**		
		0	**1**	**2**	**3**	**4**
	Task 1	6.43%	4.06%	89.51%	NA	NA
	Task 2	4.04%	10.28%	49.57%	30.23%	5.86%
	Task 3	2.23%	35.68%	50.22%	9.92%	1.95%
	Task 4	6.26%	3.16%	90.58%	NA	NA
	Task 5	3.32%	3.76%	56.55%	27.40%	8.96%
	Task 6	2.67%	3.47%	26.10%	51.56%	16.20%
	Task 7	6.18%	3.60%	90.22%	NA	NA
	Task 8	3.30%	5.70%	47.71%	40.27%	3.02%
	Task 9	2.73%	19.99%	52.82%	21.98%	2.49%

2.1.4.3 *Grades 4–5*

Table 2.1.4.3

Complete Task Analysis and Summary: Spek 4-5 S403 Online

Task Type	**Average Task Difficulty (in logits)**	**No. of Tasks**	**Average Infit Mean Square**	**Average Outfit Mean Square**
Constructed Response	-0.22	9	0.73	0.54

Name	**Tier**	**Task Difficulty (in logits)**	**Anchored?**	**Fit Statistics**	
				Infit Mnsq	**Outfit Mnsq**
1.S45AP_SI_CafeteriaRules_403_Screen_3_P1_16174	Pre-A / A	-4.06	Yes	0.80	0.20
2.S45AC_SI_CafeteriaRules_403_Screen_8_P3_16179	A / BC	0.82	Yes	0.72	0.74
3.S45C_SI_CafeteriaRules_403_Screen_12_P5_16183	BC	2.22	Yes	0.81	0.85
4.S45AP_LS_Kwolek_403_Screen_3_P1_16301	Pre-A / A	-3.47	Yes	0.80	0.15
5.S45AC_LS_Kwolek_403_Screen_9_P3_16315	A / BC	1.89	Yes	0.61	0.62
6.S45C_LS_Kwolek_403_Screen_15_P5_16321	BC	2.15	Yes	0.64	0.66
7.S45AP_MS_Circuits_402_Screen_4_P1_15077	Pre-A / A	-3.76	Yes	0.76	0.15
8.S45AC_MS_Circuits_402_Screen_11_P3_15174	A / BC	0.74	Yes	0.71	0.72
9.S45C_MS_Circuits_402_Screen_18_P5_15254	BC	1.52	Yes	0.73	0.74

Raw Score Distribution by Task	**Task**	**Raw Score**				
		0	**1**	**2**	**3**	**4**
	Task 1	8.03%	7.57%	84.39%	NA	NA
	Task 2	3.25%	5.87%	52.55%	33.72%	4.61%
	Task 3	3.51%	29.13%	43.98%	20.60%	2.78%
	Task 4	9.97%	4.95%	85.08%	NA	NA
	Task 5	3.75%	14.74%	62.11%	16.91%	2.48%
	Task 6	2.42%	27.51%	51.24%	16.37%	2.45%
	Task 7	9.25%	7.57%	83.18%	NA	NA
	Task 8	2.77%	5.98%	46.03%	37.95%	7.27%
	Task 9	3.23%	17.91%	49.42%	26.92%	2.51%

2.1.4.4 *Grades 6–8*

Table 2.1.4.4

Complete Task Analysis and Summary: Spek 6-8 S403 Online

Task Type	**Average Task Difficulty (in logits)**	**No. of Tasks**	**Average Infit Mean Square**	**Average Outfit Mean Square**
Constructed Response	0.24	9	0.76	0.59

				Fit Statistics	
Name	**Tier**	**Task Difficulty (in logits)**	**Anchored?**	**Infit Mnsq**	**Outfit Mnsq**
1.S68AP_SI_ClassroomRules_403_Screen_3_P1_16218	Pre-A / A	-3.90	Yes	0.91	0.44
2.S68AC_SI_ClassroomRules_403_Screen_8_P3_16223	A / BC	1.51	Yes	0.70	0.72
3.S68C_SI_ClassroomRules_403_Screen_13_P5_16228	BC	2.72	Yes	0.66	0.68
4.S68AP_LS_Blodgett_402_Screen_3_P1_15163	Pre-A / A	-3.20	Yes	0.79	0.25
5.S68AC_LS_Blodgett_402_Screen_8_P3_15175	A / BC	1.51	Yes	0.70	0.71
6.S68C_LS_Blodgett_402_Screen_13_P5_15181	BC	2.37	Yes	0.73	0.74
7.S68AP_MS_Radar_402_Screen_4_P1_15319	Pre-A / A	-3.73	Yes	0.86	0.24
8.S68AC_MS_Radar_402_Screen_11_P3_15326	A / BC	1.86	Yes	0.70	0.72
9.S68C_MS_Radar_402_Screen_19_P5_15340	BC	3.02	Yes	0.76	0.78

Raw Score Distribution by Task	**Task**			**Raw Score**		
		0	**1**	**2**	**3**	**4**
	Task 1	6.19%	4.75%	89.06%	NA	NA
	Task 2	5.43%	10.04%	58.00%	24.19%	2.35%
	Task 3	4.62%	32.81%	48.12%	12.93%	1.52%
	Task 4	6.78%	4.91%	88.31%	NA	NA
	Task 5	5.17%	8.03%	45.33%	36.77%	4.71%
	Task 6	3.63%	25.33%	51.01%	17.58%	2.47%
	Task 7	7.81%	4.55%	87.63%	NA	NA
	Task 8	5.38%	12.02%	52.22%	27.51%	2.87%
	Task 9	5.87%	29.49%	48.18%	14.34%	2.13%

2.1.4.5 *Grades 9–12*

Table 2.1.4.5

Complete Task Analysis and Summary: Spek 9-12 S403 Online

[redacted]

2.2 DIF Analysis and Summary

Differential item analysis (DIF) attempts to investigate whether performances on items were influenced by factors extraneous to English language proficiency (i.e., the construct being measured on the test). In other words, DIF attempts to find items that may be functioning differently for different groups based on criteria irrelevant to what is being tested. The performance of students on ACCESS for ELLs Online items and tasks is compared by dividing students into two different groupings: first, males versus females; second, students of Hispanic ethnic background versus students of all other backgrounds. Students for whom gender or ethnicity[1] was unknown were excluded from both analyses. Two commonly used procedures for detecting DIF were used: one for dichotomously scored items (Listening and Reading), conducted prior to operational testing, and one for polytomously scored items (Writing and Speaking), conducted on population data subsequent to the close of operational testing.

Dichotomous Items

Following procedures that were originally proposed by the Educational Testing Service (ETS), the Mantel-Haenszel (M-H) chi-square statistic (Mantel & Haenszel, 1959) was used for dichotomous items. This procedure compares item-level performances of students in the two groups (e.g., males versus females) who are divided into subgroups based on their performance on the total test. It is assumed that, if there is no DIF, a similar percentage of students in each group should get the item correct at any ability level (based on performance on the total test). The M-H chi-square statistic is used to check the probability that the two groups performed comparably on each item across the ability groupings. The statistic is transformed into the "M-H delta" scale. This scale is symmetrical around zero, with a delta zero interpreted as indicating that neither group is favored. A positive result indicates that one group is favored; a negative result indicates that the other group is favored.

The existing M-H procedure was designed for fixed forms, where all test takers take exactly the same set of items; therefore, the test takers can be matched on the number-correct score when computing the M-H statistic. In the multistage computerized adaptive test (CAT) condition, however, not all students take exactly the same set of items; thus, it is not possible to match students on the number-correct score. Instead, a CAT M-H DIF procedure (Zwick, Thayer, & Wingersky, 1993) was used to examine DIF for the Listening and Reading domains. First, the examinee's expected true score for the entire item pool is derived. To derive the expected true score, each examinee's Rasch ability estimate is transformed into the expected true score metric by calculating the sum of the item response functions in the operational item pool, which is evaluated at the estimated ability level of the test taker. The expected true score of the examinees is used as the matching variable for the M-H DIF procedure. Once examinees are matched on the expected true score, the ordinary M-H DIF procedure and the ETS evaluation criterion for severity of M-H DIF can be applied. In CAL's implementation of this method, examinees are matched for M-H DIF analysis on the basis of this expected true score using two-unit intervals, as recommended by Zwick and Bridgeman (2014). A two-step purification process was used in conducting the DIF analysis; that is, items with C-level DIF in the first pass were removed from the matching variable in the second stage, and the DIF was then recalculated for the remaining items.

---

[1] In the dataset, Hispanic ethnicity, as well as each of the race categories, is coded as a binary variable (Y/blank). Ethnicity information is counted as "Unknown" in cases where the student is recorded as blank for Hispanic ethnicity and also blank for every race category.

Because DIF is measured on a continuous scale, and because most items are likely to show some degree of DIF, it is useful to have guidelines to determine when the level of DIF requires further review of the item. We follow the guidance provided by ETS (Zieky, 1993) to classify items into DIF levels as follows:

- A (no DIF), when the absolute value of delta is <1.0
- B (weak DIF), when the absolute value of delta is 1.0 to 1.5
- C (strong DIF), when the absolute value of the delta is >1.5

Polytomous Items

For polytomous items (i.e., Writing and Speaking tasks), a similar approach is used. It is based on the M-H chi-square statistic and the standardized mean difference following procedures again developed by ETS (Zwick, Donoghue, & Grima, 1993; Allen, Carlson, & Zalanak, 1999). The DIF procedures developed by ETS for polytomous items were used to identify tasks that exhibit DIF. JMetrik (Meyer, 2018), an open source computer program for psychometric analysis, was used in conducting the analyses. The procedures implemented in JMetrik first calculate the Cochran-Mantel-Haenszel chi-square statistic for testing statistical 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 standardized mean difference between the performances of the two groups being compared is calculated. The standardized mean difference compares the means of the two groups, adjusting for differences in the distribution of the groups across the values of the total raw scores. To standardize the outcome, this difference is divided by the item score range and serves as an effect size measure for the Cochran-Mantel-Haenszel chi-square statistic. This effect size measure (reported as standardized P-DIF in JMetrik) ranges from -1 to 1, which may present some interpretation challenges. To mitigate this, the absolute value is taken in JMetrik (Meyer, 2018), thereby restricting the range of the rescaled effect size (standardized P-DIF*) to fall between 0 and 1. The effect size flagging criterion for polytomous items, proposed by ETS (Allen et al., 1999), is also rescaled to the standardized P-DIF* metric (Meyer, 2018).

Following guidance proposed by ETS for the National Assessment of Educational Progress (Allen et al., 1999), ACCESS for ELLs Writing and Speaking tasks are classified into three DIF levels as follows:

- AA (no DIF), when the Cochran-Mantel-Haenszel chi-square statistic is not significant or when it is significant and standardized P-DIF* is <0.05
- BB (weak DIF), when the Cochran-Mantel-Haenszel chi-square statistic is significant and standardized P-DIF* is ≥ 0.05 but <0.10
- CC (strong DIF), when the Cochran-Mantel-Haenszel chi-square statistic is significant and standardized P-DIF* is ≥ 0.10

The tables in this section provide a summary of the findings of the DIF analyses at the top, followed by information for any item or task which showed B, BB, C, or CC-level DIF. The first column gives the DIF level: A, B, or C for dichotomous items or AA, BB, or CC for polytomous tasks (i.e., Writing and Speaking tasks). The next columns show the contrasting groups in the DIF analyses: either male versus female or Hispanic versus non-Hispanic other ethnicities. The top part of the table summarizes the number of items which exhibit DIF falling into each of the three categories (A, B, or C for Listening and Reading, and AA, BB, or CC for Writing and Speaking). Any items that show B (or BB) or C (or CC)–level DIF are reported in the bottom part of the table.

Items and tasks that show C-level (or CC-level) DIF are investigated by a team of content experts to determine if any construct-irrelevant factors can be identified that may contribute to DIF. For dichotomous items, DIF analysis is conducted prior to operational testing, and the content of any items that show C-level DIF is reviewed prior to item selection. Items of concern are removed from the item selection process. For polytomous items, if content experts identify concerning construct-irrelevant factors, the task will be removed from the test for the next operational year.

2.2.1 Listening

2.2.1.1 Grade 1

Table 2.2.1.1

DIF Analysis and Summary: List 1 S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	**Favoring Male (M)**	**Favoring Female (F)**	**Favoring Hispanic (H)**	**Favoring Other (O)**
A	29	24	24	30
B	1	0	0	0
C	0	0	0	0

	Male/Female		Hispanic/Other	
Name	**DIF Level**	**Favored Group**	**DIF Level**	**Favored Group**
17.L01B_SC_PropertiesOfObjects_402_Screen_3_14952	B	M	A	H

2.2.1.2 Grades 2–3

Table 2.2.1.2

DIF Analysis and Summary: List 2-3 S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	**Favoring Male (M)**	**Favoring Female (F)**	**Favoring Hispanic (H)**	**Favoring Other (O)**
A	27	26	24	29
B	0	0	1	0
C	1	0	0	0

	Male/Female		Hispanic/Other	
Name	**DIF Level**	**Favored Group**	**DIF Level**	**Favored Group**
4.L23A_SI_ThankYouCards_P100_Screen_2_13788	C	M	A	O
19.L23C_LA_Emilys Ball_402_Screen_2_14879	A	M	B	H

2.2.1.3 Grades 4–5

Table 2.2.1.3

DIF Analysis and Summary: List 4-5 S403 Online

[redacted]

2.2.1.4 Grades 6–8

Table 2.2.1.4

DIF Analysis and Summary: List 6-8 S403 Online

[redacted]

2.2.1.5 *Grades 9–12*

Table 2.2.1.5

DIF Analysis and Summary: List 9-12 S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	**Favoring Male (M)**	**Favoring Female (F)**	**Favoring Hispanic (H)**	**Favoring Other (O)**
A	25	29	24	27
B	0	0	1	2
C	0	0	0	0
	Male/Female		**Hispanic/Other**	
Name	**DIF Level**	**Favored Group**	**DIF Level**	**Favored Group**
37.L91C_LA_Poetry_P100_Screen_2_12893	A	F	B	O
52.L91C_MA_angles_joga_P100_A202_Screen_2_13865	A	F	B	O
54.L91C_MA_angles_joga_P100_A202_Screen_4_13867	A	F	B	H

2.2.2 Reading

2.2.2.1 *Grade 1*

Table 2.2.2.1

DIF Analysis and Summary: Read 1 S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	**Favoring Male (M)**	**Favoring Female (F)**	**Favoring Hispanic (H)**	**Favoring Other (O)**
A	34	38	32	39
B	0	0	1	0
C	0	0	0	0
	Male/Female		**Hispanic/Other**	
Name	**DIF Level**	**Favored Group**	**DIF Level**	**Favored Group**
13.R01C_SS_PetStore_P100_A202_Screen_2_13313	A	M	B	H

2.2.2.2 *Grades 2–3*

Table 2.2.2.2

DIF Analysis and Summary: Read 2-3 S403 Online

[redacted]

2.2.2.3 *Grades 4–5*

Table 2.2.2.3

DIF Analysis and Summary: Read 4-5 S403 Online

[redacted]

2.2.2.4 *Grades 6–8*

Table 2.2.2.4

DIF Analysis and Summary: Read 6-8 S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	**Favoring Male (M)**	**Favoring Female (F)**	**Favoring Hispanic (H)**	**Favoring Other (O)**
A	43	28	39	33
B	1	0	0	0
C	0	0	0	0

	Male/Female		Hispanic/Other	
Name	**DIF Level**	**Favored Group**	**DIF Level**	**Favored Group**
17.R68A_SC_LeafInsects_402_V1_Screen_3_15713	B	M	A	H

2.2.2.5 *Grades 9–12*

Table 2.2.2.5

DIF Analysis and Summary: Read 9-12 S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	**Favoring Male (M)**	**Favoring Female (F)**	**Favoring Hispanic (H)**	**Favoring Other (O)**
A	34	35	39	33
B	2	1	0	0
C	0	0	0	0

	Male/Female		Hispanic/Other	
Name	**DIF Level**	**Favored Group**	**DIF Level**	**Favored Group**
25.R91B_LA_Fiction_402_V2_Screen_2_16058	B	F	A	O
46.R91B_SC_Electromagnetism_403_Screen_2_17079	B	M	A	H
52.R91C_MA_Functions_401_V1_Screen_2_13968	B	M	A	H

2.2.3 Writing

2.2.3.1 *Grade 1*

Table 2.2.3.1.1

DIF Analysis and Summary: Writ 1 A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	2	2	2	2
	0	0	0	0
CC	0	0	0	0

Table 2.2.3.1.2

DIF Analysis and Summary: Writ 1 B/C S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	1	2	1	2
BB	0	0	0	0
CC	0	0	0	0

2.2.3.2 *Grades 2–3*

Table 2.2.3.2.1

DIF Analysis and Summary: Writ 2-3 A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	2	1	1	2
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.3.2.2

DIF Analysis and Summary: Writ 2-3 B/C S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	1	2	1	2
BB	0	0	0	0
CC	0	0	0	0

2.2.3.3 Grades 4–5

Table 2.2.3.3.1

DIF Analysis and Summary: Writ 4-5 A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	2	1	1	2
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.3.3.2

DIF Analysis and Summary: Writ 4-5 B/C S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	2	1	1	2
BB	0	0	0	0
CC	0	0	0	0

2.2.3.4 Grades 6–8

Table 2.2.3.4.1

DIF Analysis and Summary: Writ 6-8 A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	1	2	1	2
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.3.4.2

DIF Analysis and Summary: Writ 6-8 B/C S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	1	2	1	2
BB	0	0	0	0
CC	0	0	0	0

2.2.3.5 *Grades 9–12*

Table 2.2.3.5.1

DIF Analysis and Summary: Writ 9-12 A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	1	2	2	1
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.3.5.2

DIF Analysis and Summary: Writ 9-12 B/C S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	2	1	2	1
BB	0	0	0	0
CC	0	0	0	0

2.2.4 Speaking

2.2.4.1 Grade 1

Table 2.2.4.1.1

DIF Analysis and Summary: Spek 1 Pre-A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	2	1	2	1
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.4.1.2

DIF Analysis and Summary: Spek 1 A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	2	4	3	3
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.4.1.3

DIF Analysis and Summary: Spek 1 B/C S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	3	3	5	1
BB	0	0	0	0
CC	0	0	0	0

2.2.4.2 *Grades 2–3*

Table 2.2.4.2.1

DIF Analysis and Summary: Spek 2-3 Pre-A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Male (M)	ing Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	1	2	1	2
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.4.2.2

DIF Analysis and Summary: Spek 2-3 A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	3	3	1	5
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.4.2.3

DIF Analysis and Summary: Spek 2-3 B/C S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	3	3	3	3
BB	0	0	0	0
CC	0	0	0	0

2.2.4.3 *Grades 4–5*

Table 2.2.4.3.1

DIF Analysis and Summary: Spek 4-5 Pre-A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	2	1	2	1
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.4.3.2

DIF Analysis and Summary: Spek 4-5 A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	3	3	3	3
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.4.3.3

DIF Analysis and Summary: Spek 4-5 B/C S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	3	3	3	3
BB	0	0	0	0
CC	0	0	0	0

2.2.4.4 *Grades 6–8*

Table 2.2.4.4.1

DIF Analysis and Summary: Spek 6-8 Pre-A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	1	2	2	1
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.4.4.2

DIF Analysis and Summary: Spek 6-8 A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	1	5	3	3
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.4.4.3

DIF Analysis and Summary: Spek 6-8 B/C S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
	2	4	3	3
	0	0	0	0
CC	0	0	0	0

2.2.4.5 Grades 9–12

Table 2.2.4.5.1

DIF Analysis and Summary: Spek 9-12 Pre-A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	1	2	2	1
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.4.5.2

DIF Analysis and Summary: Spek 9-12 A S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	2	4	3	3
BB	0	0	0	0
CC	0	0	0	0

Table 2.2.4.5.3

DIF Analysis and Summary: Spek 9-12 B/C S403 Online

DIF Summary	Male/Female		Hispanic/Other	
DIF Level	Favoring Male (M)	Favoring Female (F)	Favoring Hispanic (H)	Favoring Other (O)
AA	2	4	2	4
BB	0	0	0	0

2.3 Raw Score Distribution for Speaking and Writing

Figures and tables in this section provide raw score information for Speaking and Writing. For each grade-level cluster and tier combination, the figure 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 received each raw score.

Each table in this section summarizes results for a grade-level cluster and tier combination (e.g., Speaking 4–5 Tier A). For each table, results are broken down by grade and also presented for the grade-level cluster as a whole for that tier. The following information is included in each table:

- The number of students in the analyses (the number of students who were not absent, invalid, refused, exempt, or in the wrong grade-level 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

2.3.1 Listening

The ACCESS Online Listening test is a multistage adaptive assessment. As students do not all take the same set of items in the test, raw score distributions are not presented.

2.3.2 Reading

The ACCESS Online Reading test is a multistage adaptive assessment. As students do not all take the same set of items in the test, raw score distributions are not presented.

2.3.3 Writing

2.3.3.1 Grade 1

Table 2.3.3.1.1

Raw Score Descriptive Statistics: Writ 1 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	154,550	0	28	14.24	4.92
Total	154,550	0	28	14.24	4.92

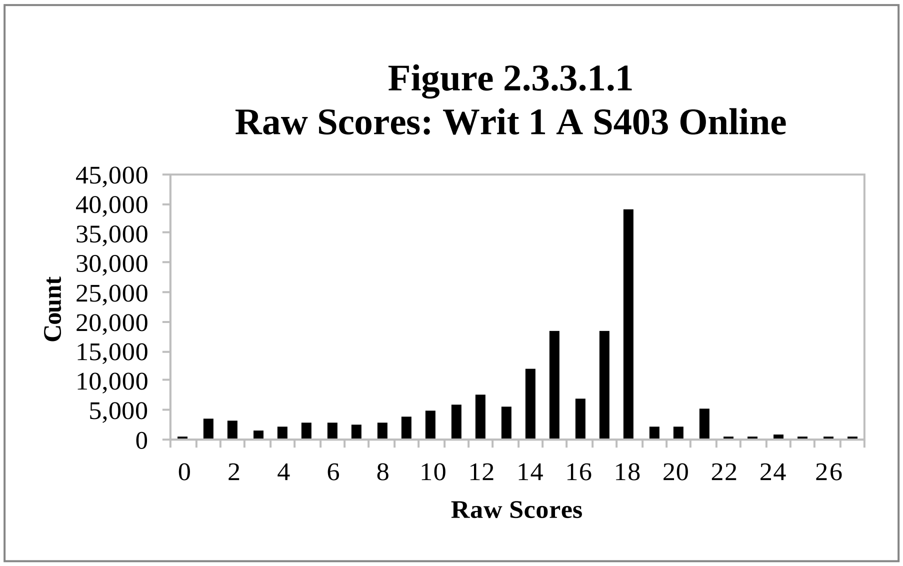


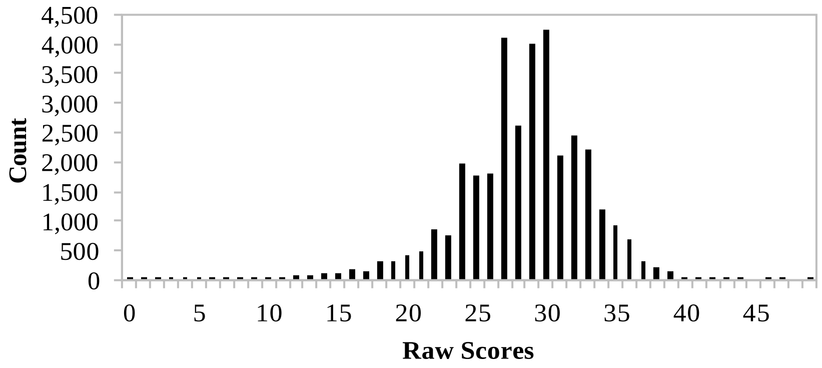
Figure 2.3.3.1.1
Raw Scores: Writ 1 A S403 Online

Table 2.3.3.1.2

Raw Score Descriptive Statistics: Writ 1 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	35,093	0	50	28.24	4.93
Total	35,093	0	50	28.24	4.93

Figure 2.3.3.1.2
Raw Scores: Writ 1 B/C S403 Online



2.3.3.2 *Grades 2–3*

Table 2.3.3.2.1

Raw Score Descriptive Statistics: Writ 2-3 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	56,167	0	18	9.18	4.02
3	37,447	0	19	9.73	4.03
Total	93,614	0	19	9.40	4.03

Figure 2.3.3.2.1
Raw Scores: Writ 2-3 AS403 Online

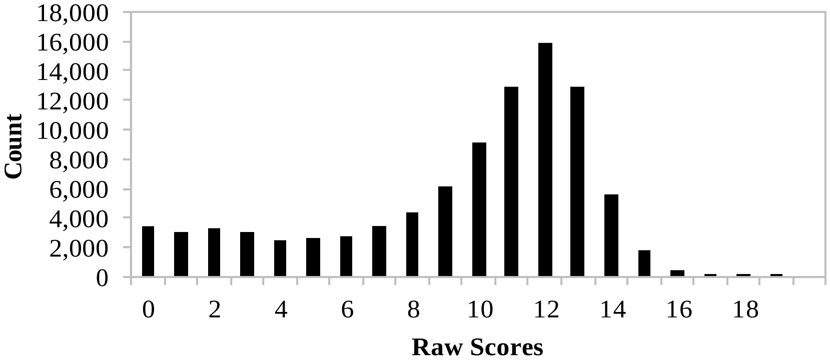
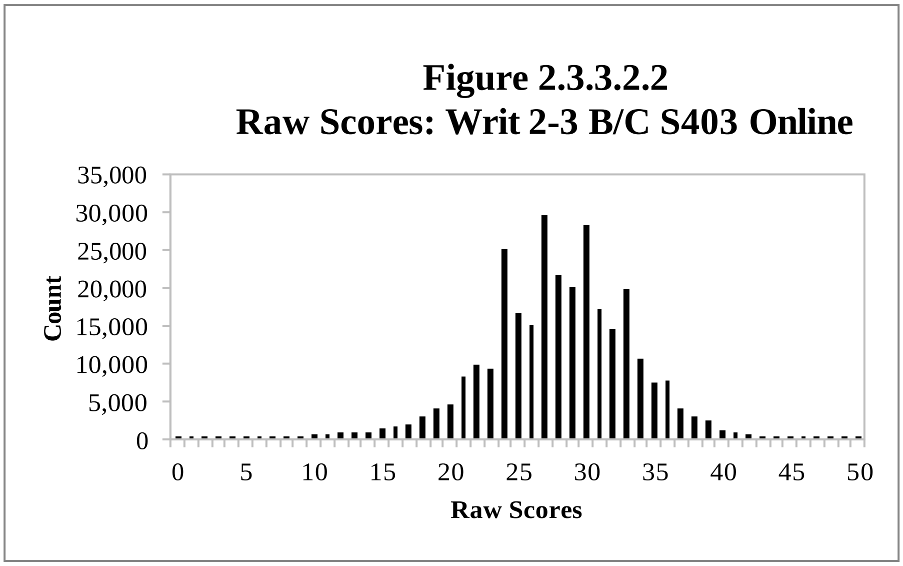


Table 2.3.3.2.2

Raw Score Descriptive Statistics: Writ 2-3 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	138,684	0	48	25.99	5.50
3	157,979	0	51	29.43	5.16
Total	296,663	0	51	27.82	5.59



2.3.3.3 *Grades 4–5*

Table 2.3.3.3.1

Raw Score Descriptive Statistics: Writ 4-5 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	22,213	0	19	7.65	4.30
5	21,996	0	20	8.70	4.30
Total	44,209	0	20	8.17	4.33

Figure 2.3.3.3.1
Raw Scores: Writ 4-5 AS403 Online

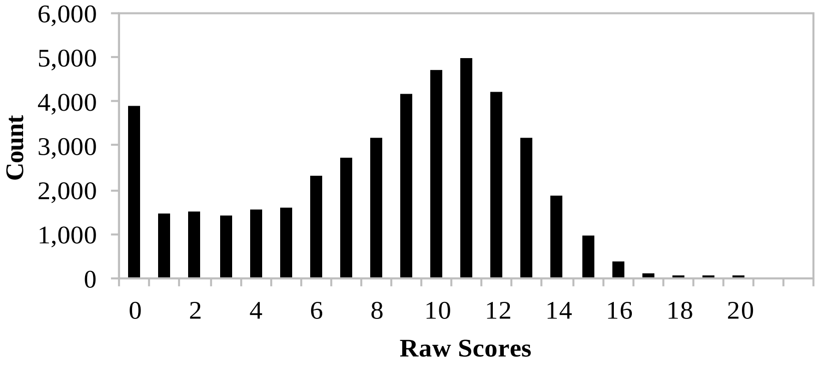
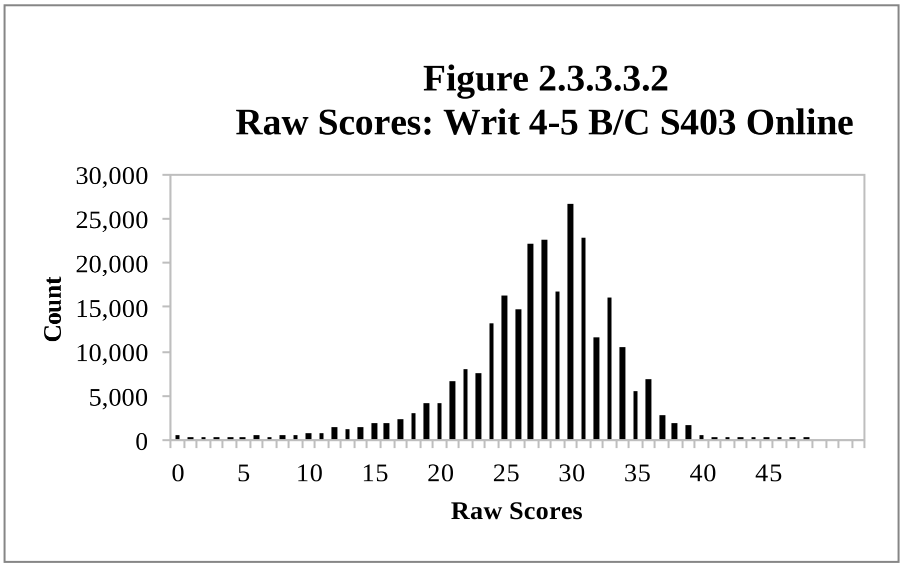


Table 2.3.3.3.2

Raw Score Descriptive Statistics: Writ 4-5 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	146,575	0	47	26.59	5.96
5	115,225	0	48	28.82	5.37
Total	261,800	0	48	27.57	5.81



2.3.3.4 *Grades 6–8*

Table 2.3.3.4.1

Raw Score Descriptive Statistics: Writ 6-8 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	31,021	0	20	9.29	3.93
7	34,794	0	22	10.12	3.99
8	34,272	0	25	10.57	4.06
Total	100,087	0	25	10.02	4.03

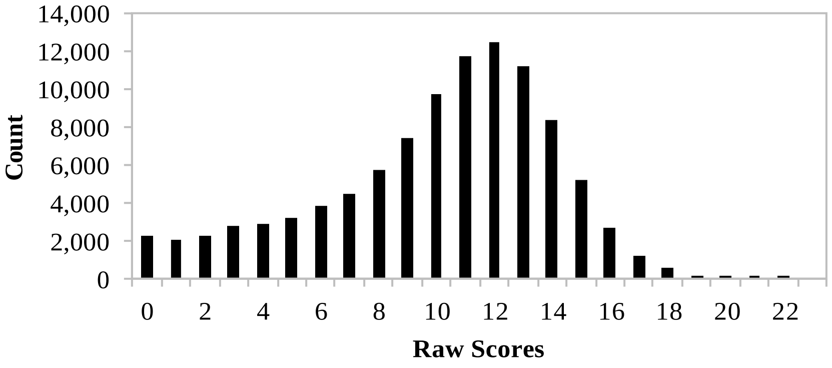


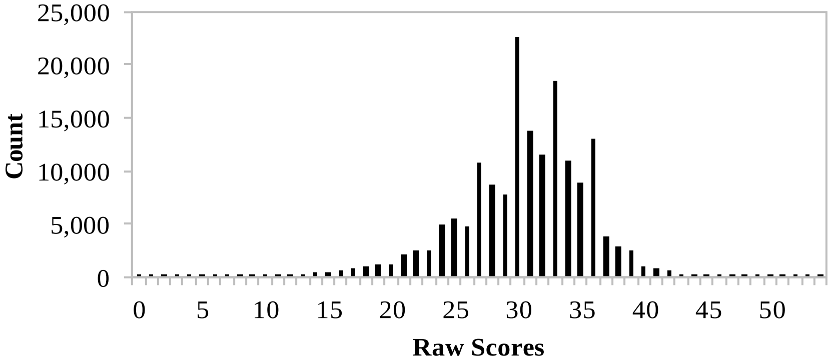
Figure 2.3.3.4.1
Raw Scores: Writ 6-8 AS403 Online

Table 2.3.3.4.2

Raw Score Descriptive Statistics: Writ 6-8 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	67,227	0	52	29.05	5.25
7	53,575	0	52	30.83	5.01
8	47,805	0	54	32.12	4.90
Total	168,607	0	54	30.48	5.23

Figure 2.3.3.4.2
Raw Scores: Writ 6-8 B/C S403 Online



2.3.3.5 Grades 9–12

Table 2.3.3.5.1

Raw Score Descriptive Statistics: Writ 9-12 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	41,531	0	24	10.23	5.05
10	28,776	0	24	11.54	4.38
11	23,202	0	23	12.28	4.15
12	15,363	0	24	12.40	4.24
Total	108,872	0	24	11.32	4.67

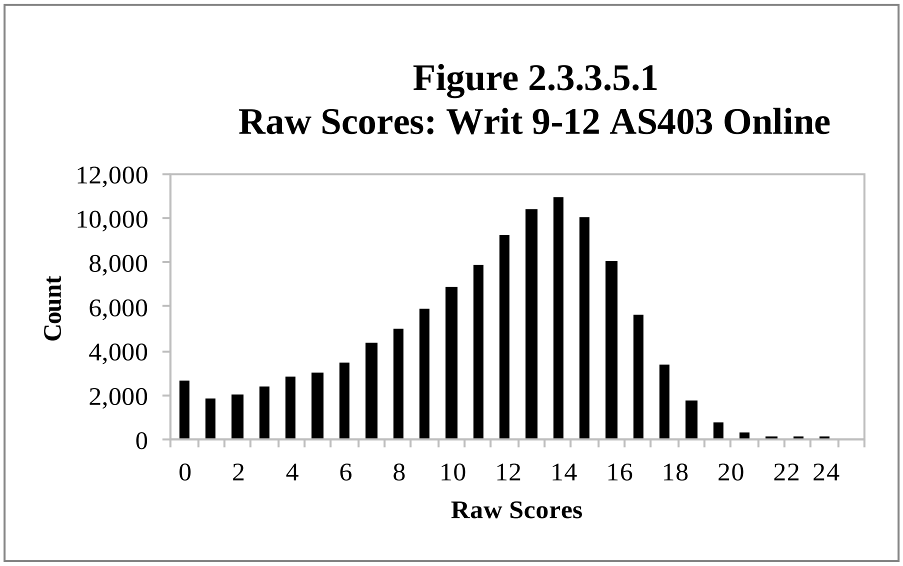
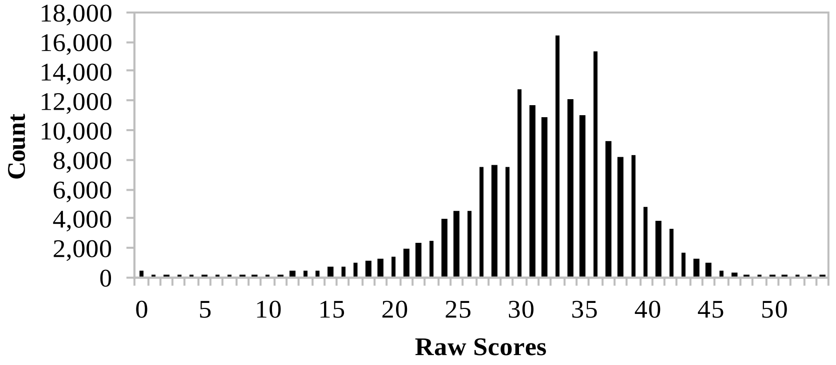


Table 2.3.3.5.2

Raw Score Descriptive Statistics: Writ 9-12 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	53,578	0	53	31.39	6.42
10	49,396	0	54	31.91	6.53
11	46,009	0	52	32.67	6.61
12	35,772	0	52	32.26	7.00
Total	184,755	0	54	32.02	6.63

Figure 2.3.3.5.2
Raw Scores: Writ 9-12 B/C S403 Online



2.3.4 Speaking

2.3.4.1 *Grade 1*

Table 2.3.4.1.1

Raw Score Descriptive Statistics: Spek 1 Pre-A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	5,158	0	6	4.40	2.15
Total	5,158	0	6	4.40	2.15

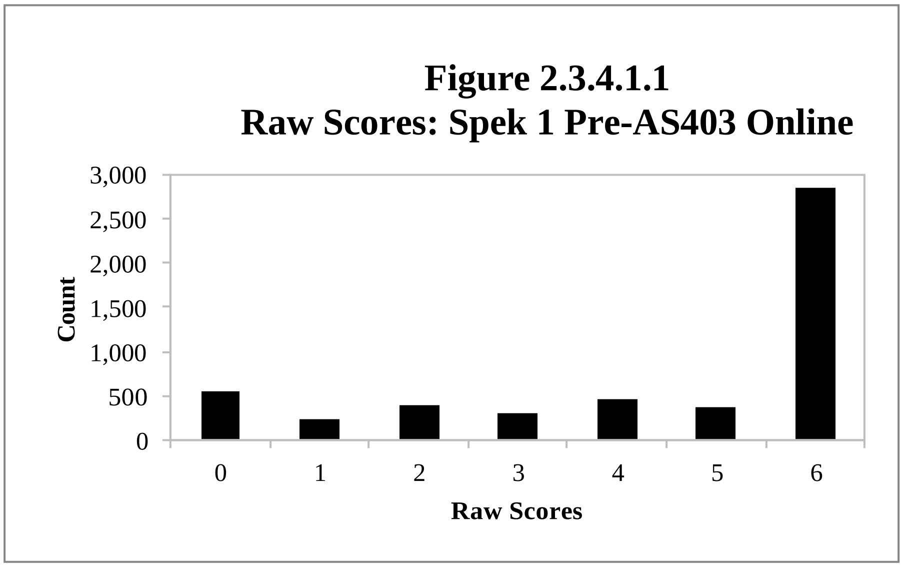


Table 2.3.4.1.2

Raw Score Descriptive Statistics: Spek 1 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	61,323	0	18	10.56	3.46
Total	61,323	0	18	10.56	3.46

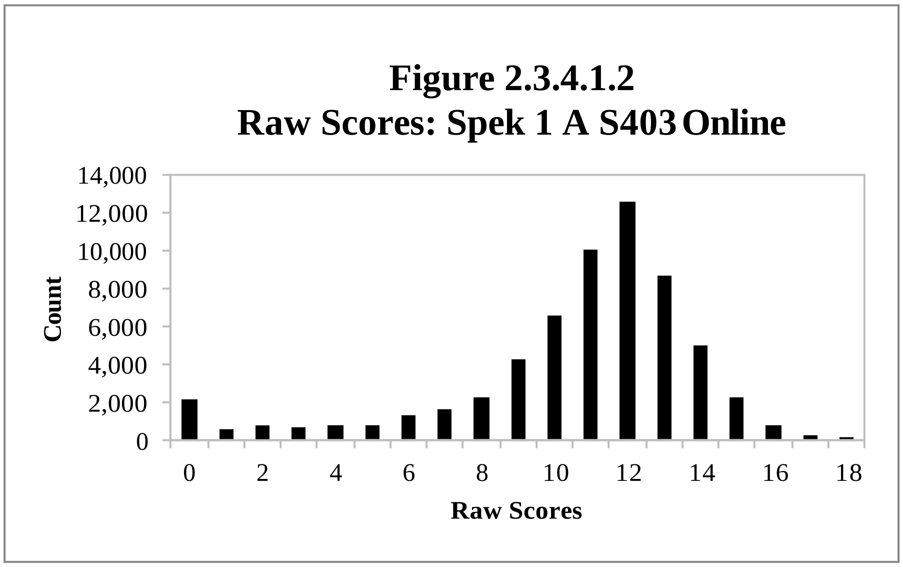
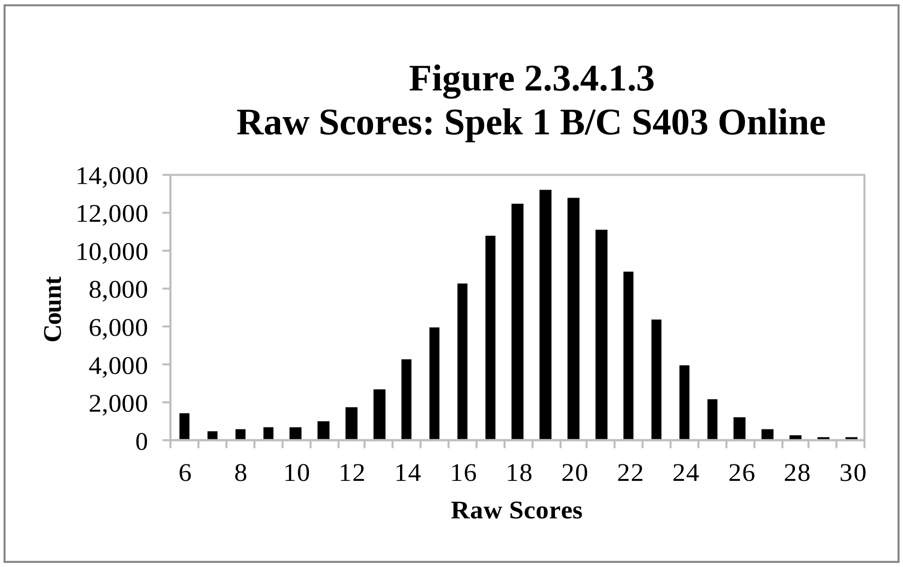


Table 2.3.4.1.3

Raw Score Descriptive Statistics: Spek 1 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	111,043	6	30	18.59	3.75
Total	111,043	6	30	18.59	3.75



2.3.4.2 *Grades 2–3*

Table 2.3.4.2.1

Raw Score Descriptive Statistics: Spek 2-3 Pre-A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	6,879	0	6	4.89	1.93
3	9,480	0	6	4.90	1.92
Total	16,359	0	6	4.90	1.92

Figure 2.3.4.2.1
Raw Scores: Spek 2-3 Pre-AS403 Online

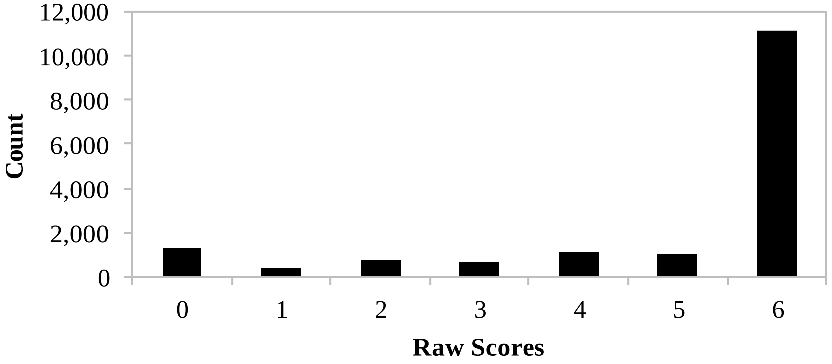


Table 2.3.4.2.2

Raw Score Descriptive Statistics: Spek 2-3 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	41,584	0	18	10.94	3.01
3	37,429	0	18	11.81	2.64
Total	79,013	0	18	11.35	2.87

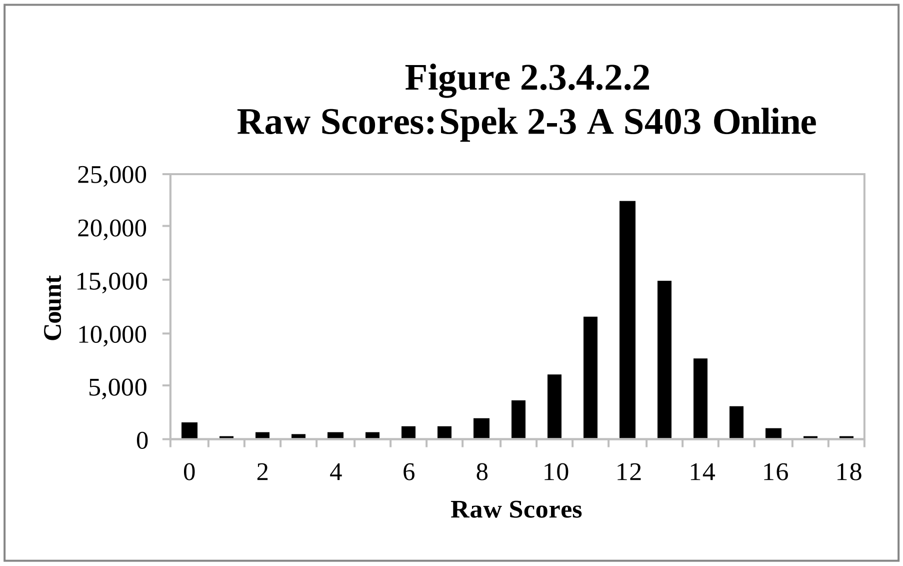
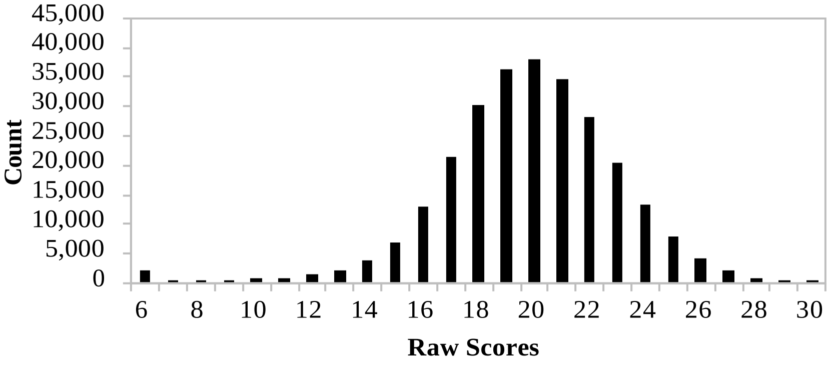


Table 2.3.4.2.3

Raw Score Descriptive Statistics: Spek 2-3 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	133,785	6	30	19.20	3.29
3	136,710	6	30	20.36	3.14
Total	270,495	6	30	19.78	3.27

Figure 2.3.4.2.3
Raw Scores: Spek 2-3 B/C S403 Online



2.3.4.3 *Grades 4–5*

Table 2.3.4.3.1

Raw Score Descriptive Statistics: Spek 4-5 Pre-A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	2,390	0	6	4.04	2.19
5	3,621	0	6	4.35	2.08
Total	6,011	0	6	4.23	2.13

Figure 2.3.4.3.1
Raw Scores: Spek 4-5 Pre-AS403 Online

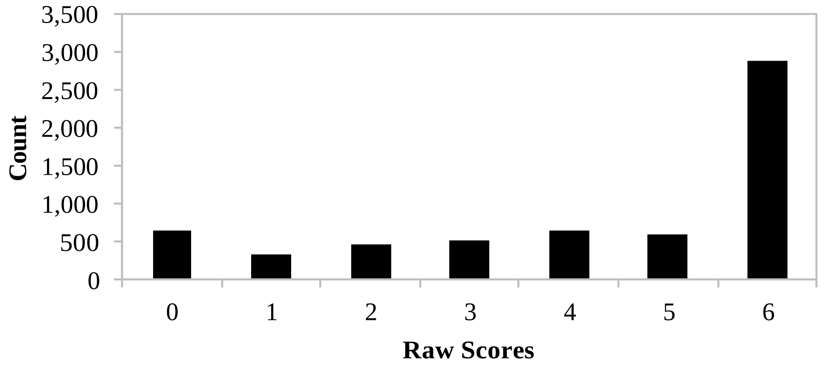


Table 2.3.4.3.2

Raw Score Descriptive Statistics: Spek 4-5 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	15,726	0	18	10.01	3.07
5	12,275	0	18	10.31	2.88
Total	28,001	0	18	10.14	2.99

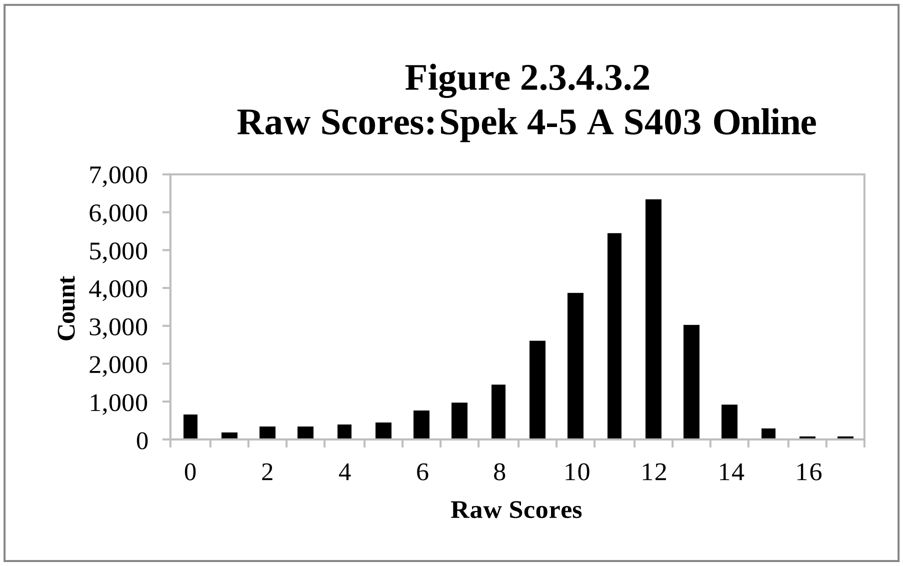


Table 2.3.4.3.3

Raw Score Descriptive Statistics: Spek 4-5 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	155,646	6	30	18.64	3.35
5	124,000	6	30	18.96	3.46
Total	279,646	6	30	18.78	3.40

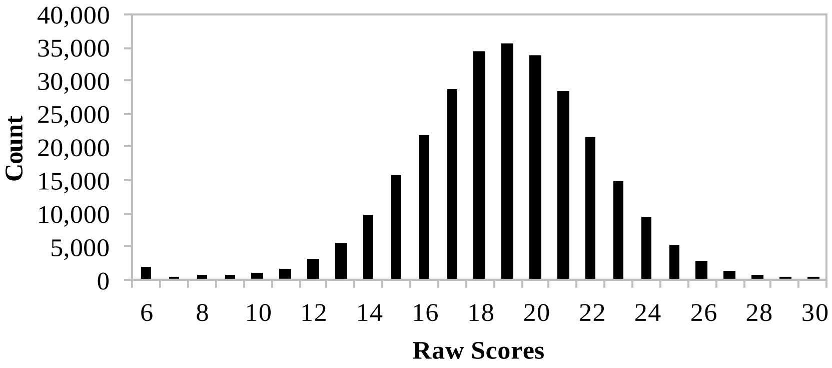


Figure 2.3.4.3.3
Raw Scores: Spek 4-5 B/C S403 Online

2.3.4.4 *Grades 6–8*

Table 2.3.4.4.1

Raw Score Descriptive Statistics: Spek 6-8 Pre-A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,811	0	6	4.78	1.89
7	3,854	0	6	4.83	1.88
8	4,829	0	6	4.86	1.83
Total	11,494	0	6	4.83	1.86

Figure 2.3.4.4.1
Raw Scores: Spek 6-8 Pre-AS403 Online

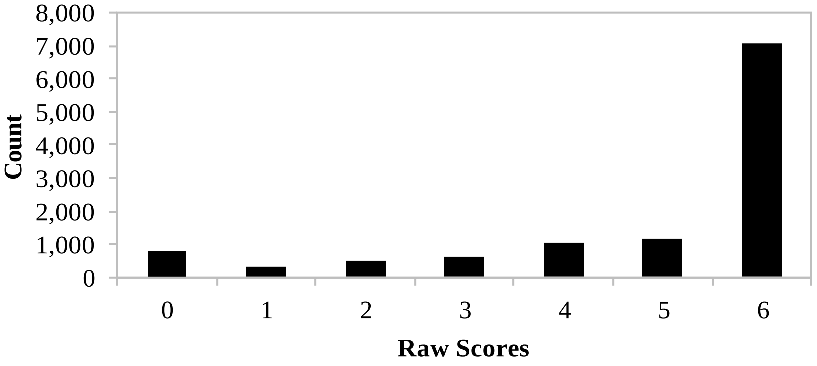


Table 2.3.4.4.2

Raw Score Descriptive Statistics: Spek 6-8 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	17,887	0	17	10.39	2.99
7	15,034	0	18	10.13	3.08
8	24,908	0	18	10.83	3.03
Total	57,829	0	18	10.51	3.05

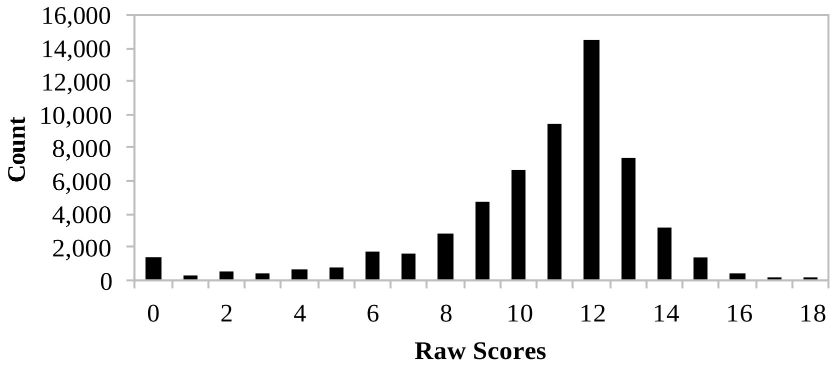


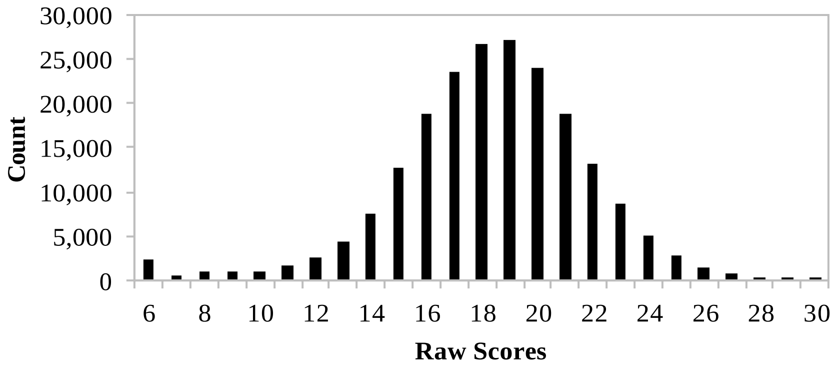
Figure 2.3.4.4.2
Raw Scores: Spek 6-8 A S403 Online

Table 2.3.4.4.3

Raw Score Descriptive Statistics: Spek 6-8 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	80,038	6	30	17.94	3.29
7	71,216	6	30	18.19	3.52
8	54,262	6	30	18.99	3.56
Total	205,516	6	30	18.30	3.47

Figure 2.3.4.4.3
Raw Scores: Spek 6-8 B/C S403 Online



2.3.4.5 *Grades 9–12*

Table 2.3.4.5.1

Raw Score Descriptive Statistics: Spek 9-12 Pre-A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	5,440	0	6	4.92	1.86
10	6,455	0	6	5.31	1.58
11	6,209	0	6	5.44	1.47
12	4,998	0	6	5.38	1.58
Total	23,102	0	6	5.27	1.63

Figure 2.3.4.5.1
Raw Scores: Spek 9-12 Pre-AS403 Online

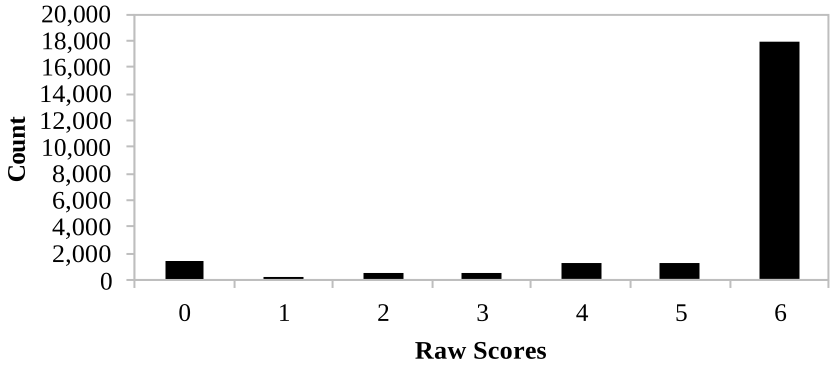


Table 2.3.4.5.2

Raw Score Descriptive Statistics: Spek 9-12 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	54,492	0	18	10.18	3.11
10	31,840	0	17	10.53	2.91
11	14,041	0	17	10.26	3.04
12	23,318	0	18	11.03	3.05
Total	123,691	0	18	10.44	3.06

Figure 2.3.4.5.2
Raw Scores: Spek 9-12 AS403 Online

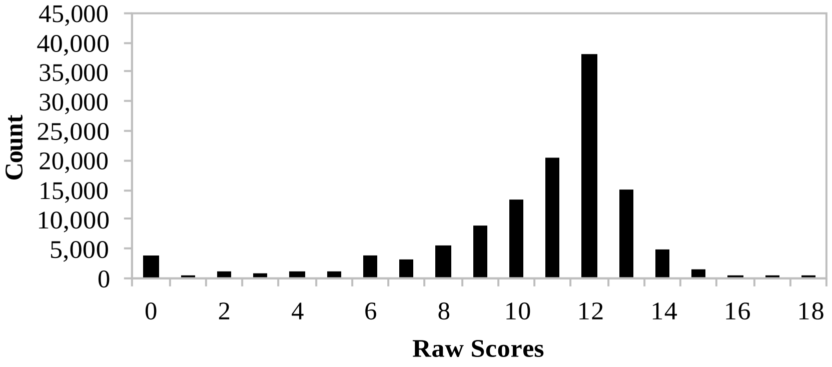
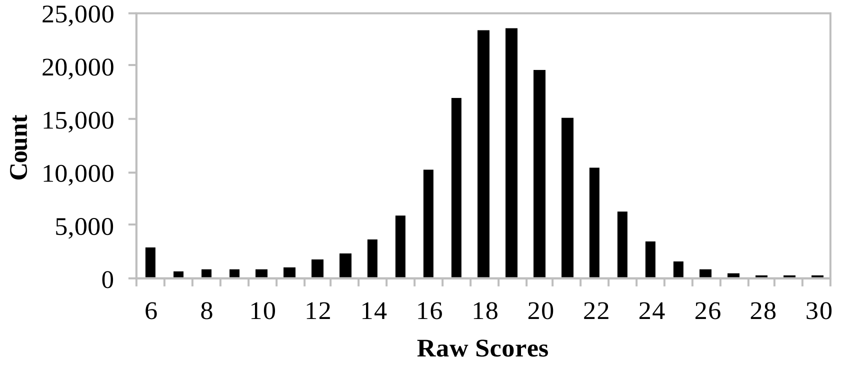


Table 2.3.4.5.3

Raw Score Descriptive Statistics: Spek 9-12 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	37,139	6	30	18.41	3.22
10	41,507	6	30	18.38	3.43
11	49,948	6	30	18.17	3.69
12	23,922	6	30	18.84	3.69
Total	152,516	6	30	18.39	3.52

Figure 2.3.4.5.3
Raw Scores: Spek 9-12 B/C S403 Online



2.4 Scale Score Distribution

Figures and tables in this section relate to the ACCESS for ELLs scale scores on each test form. For each test form, raw scores were converted to vertically equated scale scores. Scale score distribution is presented by grade-level cluster. For Writing and Speaking, it is also presented by grade-level cluster and tier.

For each test form, the figure shows the distribution of the scale scores. Scale scores are plotted on the horizontal axis.

For Listening and Reading, scale scores are grouped into units of five scale score points (e.g., 100-104, 105-109, 110-114, etc.).

For Speaking and Writing, each individual scale score point is plotted for each test form. For figures that summarize both test forms in a cluster, scale scores are grouped into units of five scale score points.

The number of students with scale scores falling into each range is plotted on the vertical axis.

The tables in this section show, by grade and by total for the grade-level cluster:

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

2.4.1 Listening

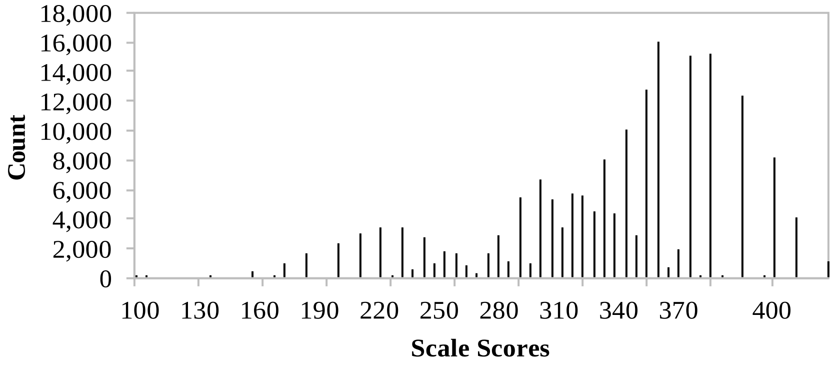
2.4.1.1 Grade 1

Table 2.4.1.1

Scale Score Descriptive Statistics: List 1 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	181,107	104	428	325.28	54.33
Total	181,107	104	428	325.28	54.33

Figure 2.4.1.1
Scale Scores: List 1 S403 Online

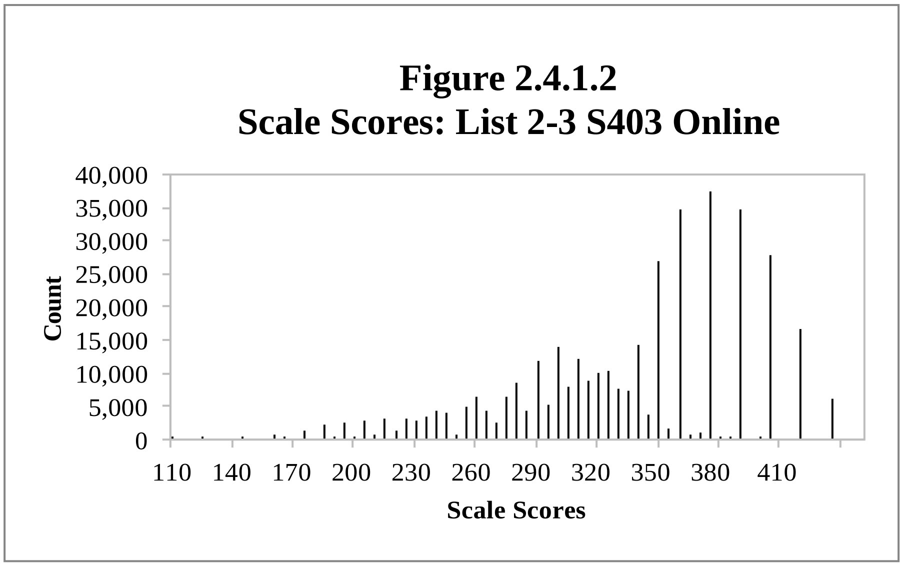


2.4.1.2 *Grades 2–3*

Table 2.4.1.2

Scale Score Descriptive Statistics: List 2-3 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	185,843	112	436	325.54	56.06
3	186,821	112	436	348.97	55.65
Total	372,664	112	436	337.28	57.07



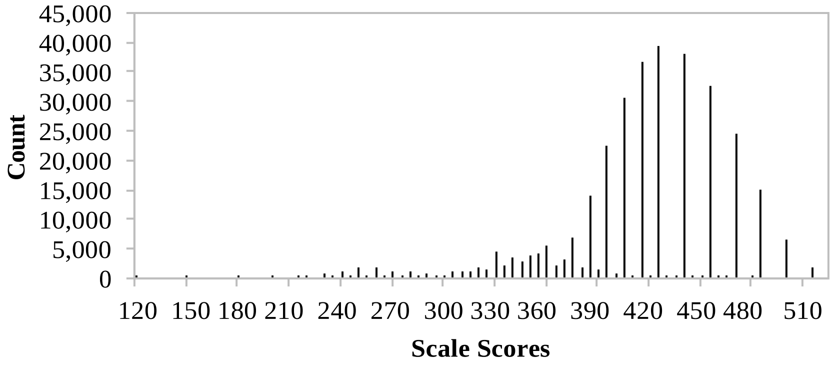
2.4.1.3 *Grades 4–5*

Table 2.4.1.3

Scale Score Descriptive Statistics: List 4-5 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	178,833	120	519	411.73	48.57
5	144,131	120	519	420.77	51.83
Total	322,964	120	519	415.77	50.25

Figure 2.4.1.3
Scale Scores: List 4-5 S403 Online



2.4.1.4 *Grades 6–8*

Table 2.4.1.4

Scale Score Descriptive Statistics: List 6-8 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	102,133	132	522	392.60	43.03
7	90,916	132	522	396.97	48.26
8	84,738	132	522	402.27	53.05
Total	277,787	132	522	396.98	48.14

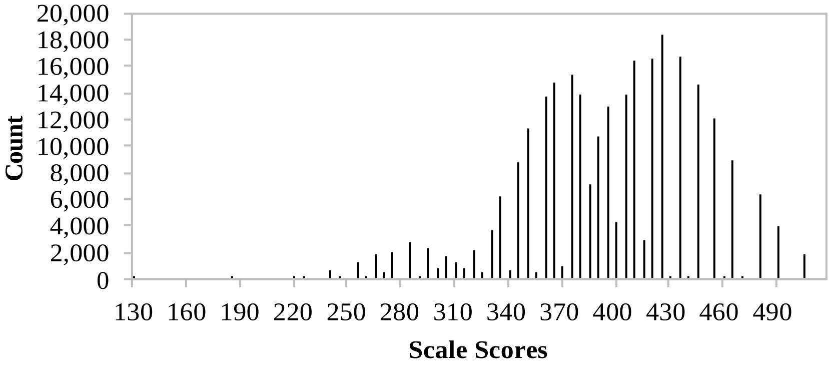


Figure 2.4.1.4
Scale Scores: List 6-8 S403 Online

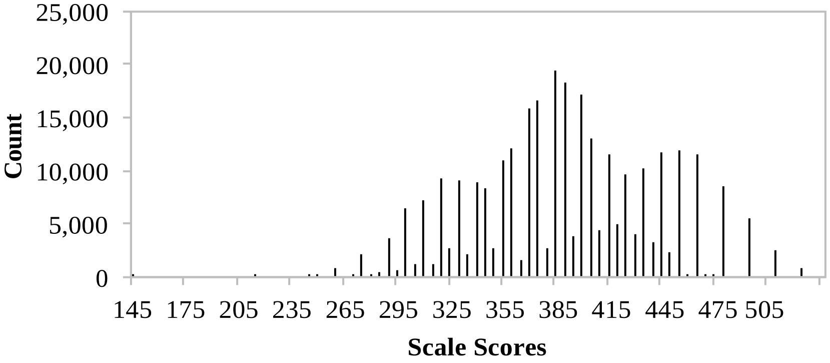
2.4.1.5 *Grades 9–12*

Table 2.4.1.5

Scale Score Descriptive Statistics: List 9-12 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	97,431	148	527	387.25	48.91
10	80,344	148	527	391.72	51.46
11	71,529	215	527	396.59	52.38
12	52,461	215	527	396.51	51.83
Total	301,765	148	527	392.26	51.09

Figure 2.4.1.5
Scale Scores: List 9-12 S403 Online



2.4.2 Reading

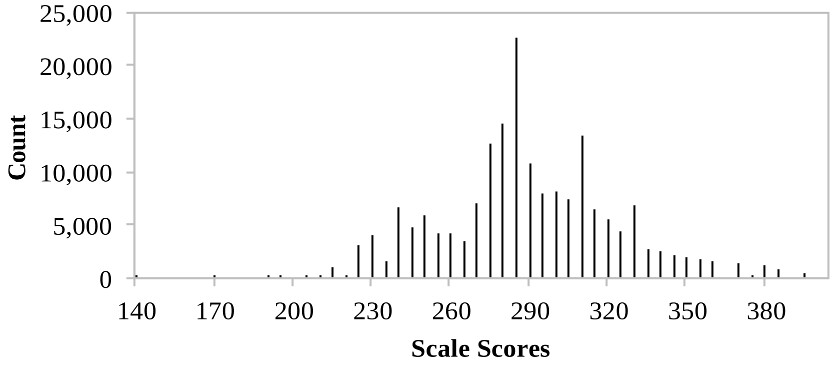
2.4.2.1 Grade 1

Table 2.4.2.1

Scale Score Descriptive Statistics: Read 1 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	183,421	141	407	291.83	33.39
Total	183,421	141	407	291.83	33.39

Figure 2.4.2.1
Scale Scores: Read 1 S403 Online



2.4.2.2 *Grades 2–3*

Table 2.4.2.2

Scale Score Descriptive Statistics: Read 2-3 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	186,415	198	418	318.59	30.36
3	186,956	158	418	333.70	34.71
Total	373,371	158	418	326.16	33.48

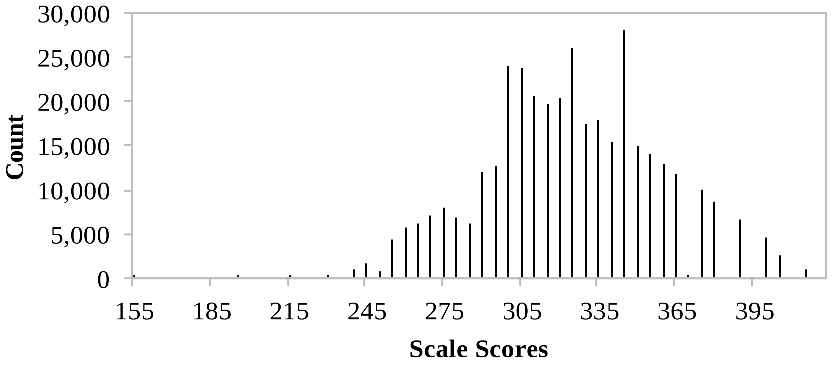


Figure 2.4.2.2
Scale Scores: Read 2-3 S403 Online

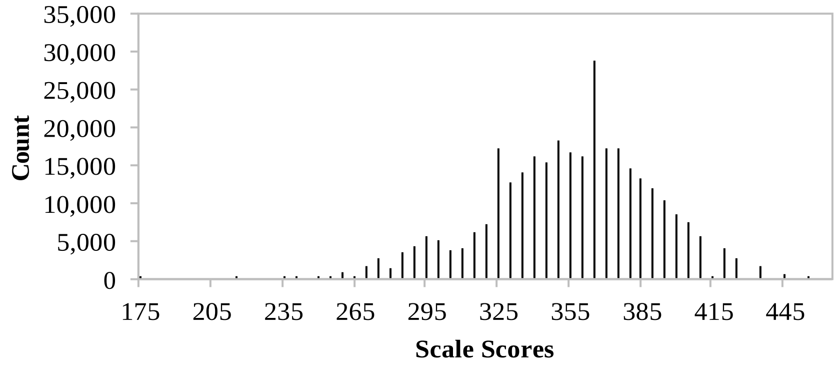
2.4.2.3 *Grades 4–5*

Table 2.4.2.3

Scale Score Descriptive Statistics: Read 4-5 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	176,492	175	456	354.98	32.24
5	142,149	217	456	359.86	35.18
Total	318,641	175	456	357.15	33.67

Figure 2.4.2.3
Scale Scores: Read 4-5 S403 Online



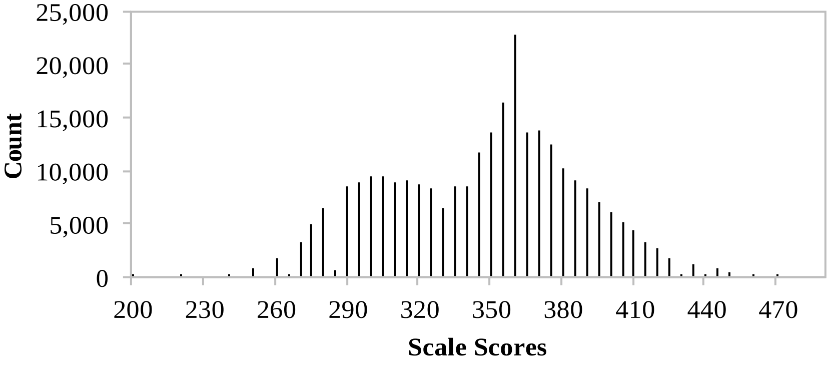
2.4.2.4 *Grades 6–8*

Table 2.4.2.4

Scale Score Descriptive Statistics: Read 6-8 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	102,022	200	472	343.10	36.19
7	91,277	200	472	348.97	39.29
8	84,658	200	472	355.45	41.24
Total	277,957	200	472	348.79	39.13

Figure 2.4.2.4
Scale Scores: Read 6-8 S403 Online



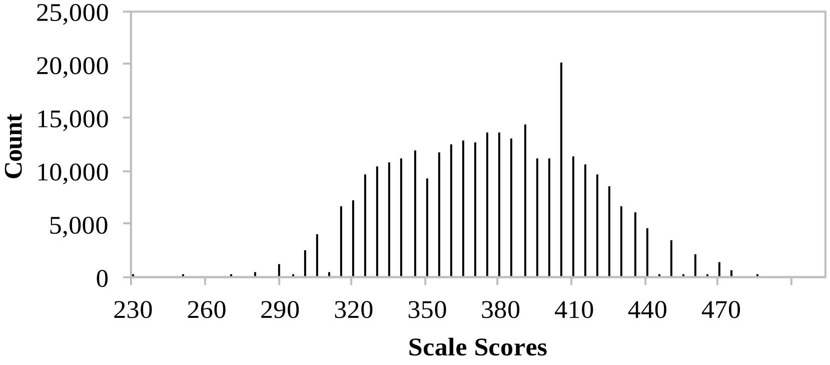
2.4.2.5 Grades 9–12

Table 2.4.2.5

Scale Score Descriptive Statistics: Read 9-12 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	96,429	233	488	370.19	37.45
10	79,330	250	488	378.29	37.25
11	70,173	250	488	384.67	37.58
12	51,887	233	488	385.38	36.87
Total	297,819	233	488	378.41	37.86

Figure 2.4.2.5
Scale Scores: Read 9-12 S403 Online



2.4.3 Writing

2.4.3.1 Grade 1

Table 2.4.3.1.1

Scale Score Descriptive Statistics: Writ 1 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	154,550	111	342	249.03	30.15
Total	154,550	111	342	249.03	30.15

Figure 2.4.3.1.1
Scale Scores:Writ 1 A S403 Online

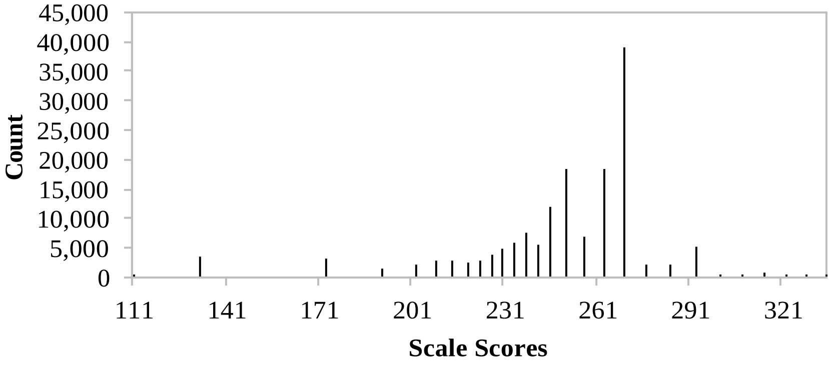


Table 2.4.3.1.2

Scale Score Descriptive Statistics: Writ 1 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	35,093	111	408	295.36	24.82
Total	35,093	111	408	295.36	24.82

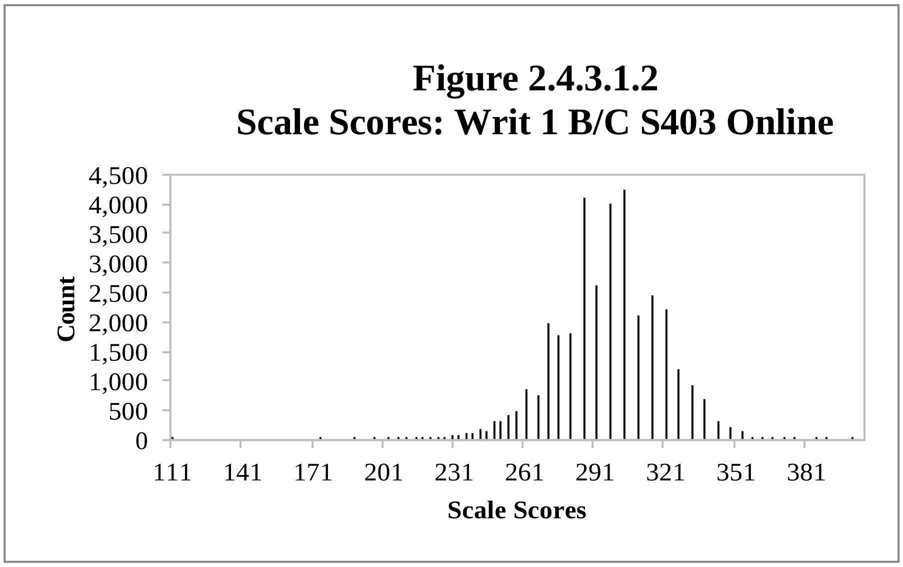
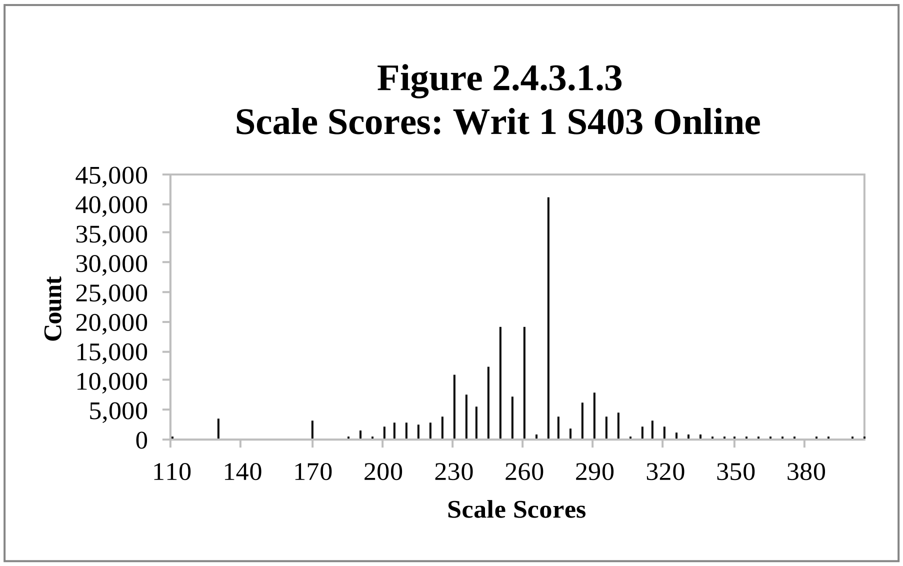


Table 2.4.3.1.3

Scale Score Descriptive Statistics: Writ 1 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	189,643	111	408	257.60	34.33
Total	189,643	111	408	257.60	34.33



2.4.3.2 *Grades 2–3*

Table 2.4.3.2.1

Scale Score Descriptive Statistics: Writ 2-3 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	56,167	133	348	258.22	35.71
3	37,447	133	358	262.96	36.10
Total	93,614	133	358	260.11	35.94

Figure 2.4.3.2.1
Scale Scores: Writ 2-3 AS403 Online

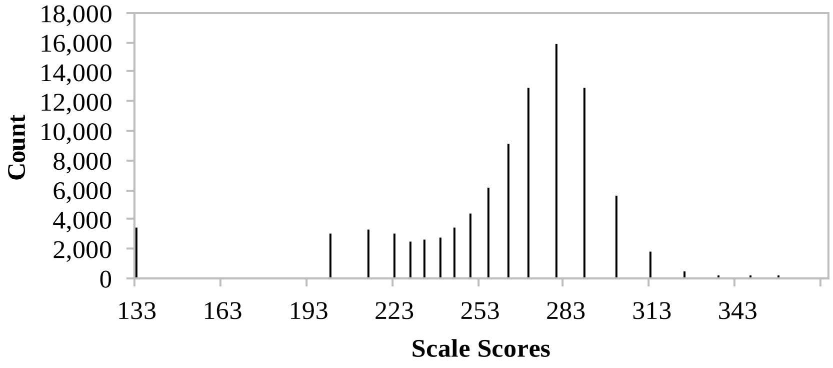


Table 2.4.3.2.2

Scale Score Descriptive Statistics: Writ 2-3 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	138,684	133	423	312.18	26.62
3	157,979	133	443	330.09	27.31
Total	296,663	133	443	321.72	28.43

Figure 2.4.3.2.2
Scale Scores: Writ 2-3 B/C S403 Online

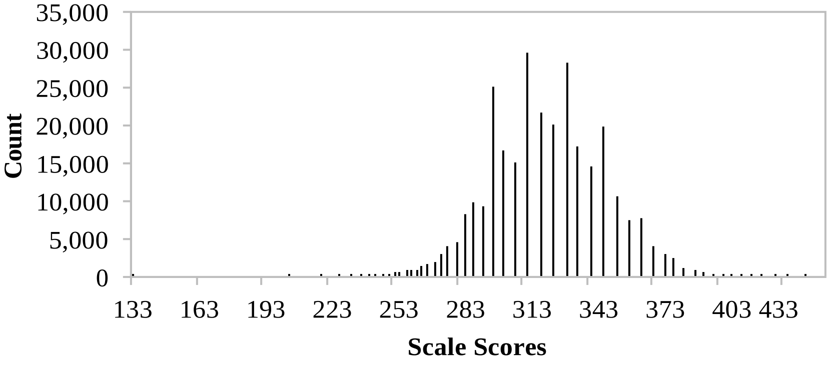
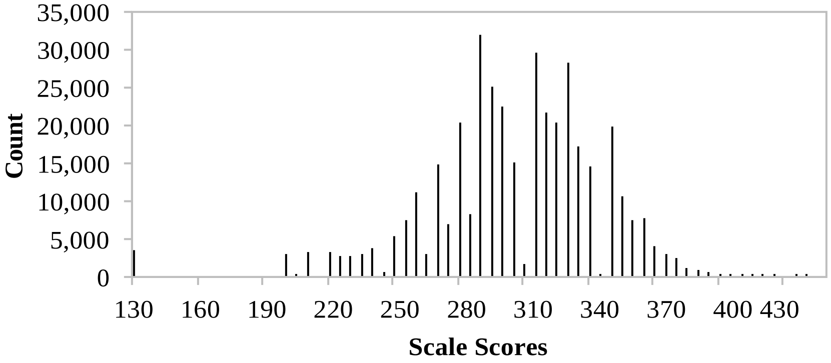


Table 2.4.3.2.3

Scale Score Descriptive Statistics: Writ 2-3 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	194,851	133	423	296.62	38.33
3	195,426	133	443	317.23	39.38
Total	390,277	133	443	306.94	40.20

Figure 2.4.3.2.3
Scale Scores: Writ 2-3 S403 Online



2.4.3.3 *Grades 4–5*

Table 2.4.3.3.1

Scale Score Descriptive Statistics: Writ 4-5 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	22,213	155	386	270.99	46.08
5	21,996	155	396	280.70	44.29
Total	44,209	155	396	275.82	45.46

Figure 2.4.3.3.1
Scale Scores: Writ 4-5 AS403 Online

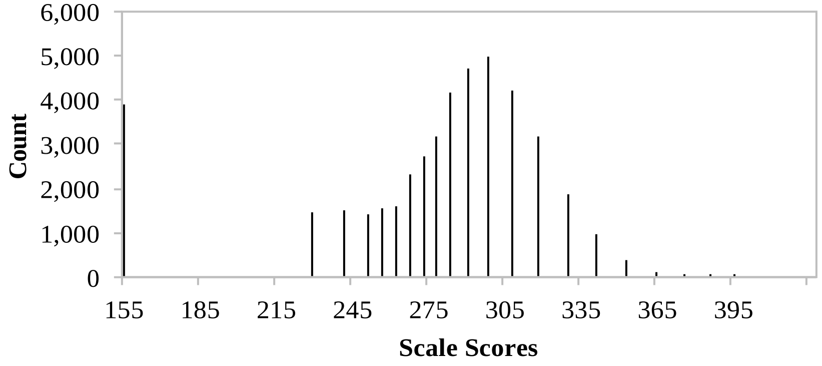


Table 2.4.3.3.2

Scale Score Descriptive Statistics: Writ 4-5 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	146,575	155	449	346.11	29.23
5	115,225	155	454	357.53	27.63
Total	261,800	155	454	351.14	29.10

Figure 2.4.3.3.2
Scale Scores: Writ 4-5 B/C S403 Online

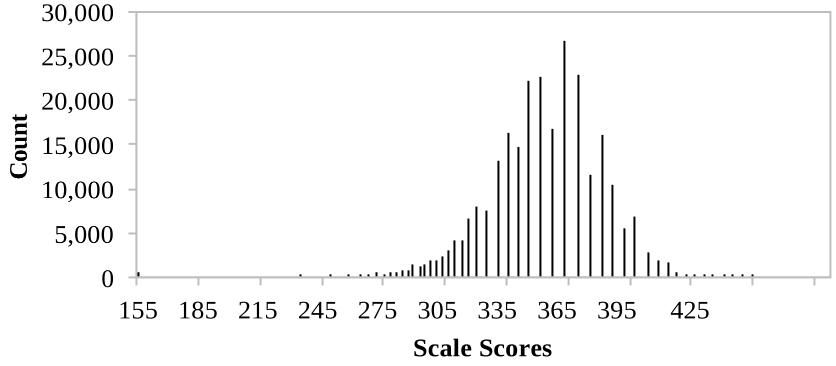
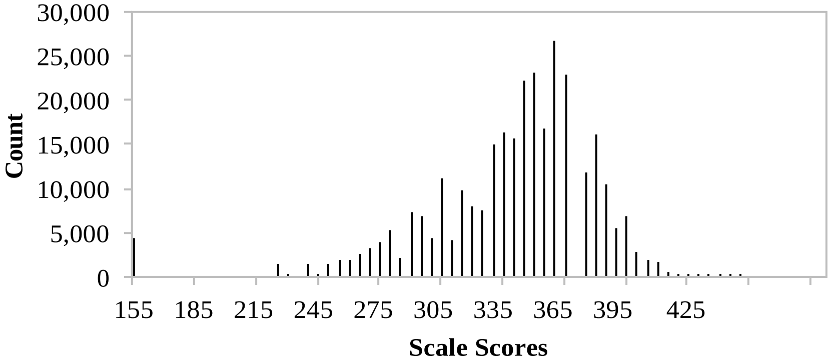


Table 2.4.3.3.3

Scale Score Descriptive Statistics: Writ 4-5 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	168,788	155	449	336.22	40.82
5	137,221	155	454	345.22	41.83
Total	306,009	155	454	340.26	41.52

Figure 2.4.3.3.3
Scale Scores: Writ 4-5 S403 Online



2.4.3.4 *Grades 6–8*

Table 2.4.3.4.1

Scale Score Descriptive Statistics: Writ 6-8 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	31,021	188	389	282.73	31.83
7	34,794	188	407	290.00	32.99
8	34,272	188	438	294.14	34.19
Total	100,087	188	438	289.16	33.38

Figure 2.4.3.4.1
Scale Scores: Writ 6-8 AS403 Online

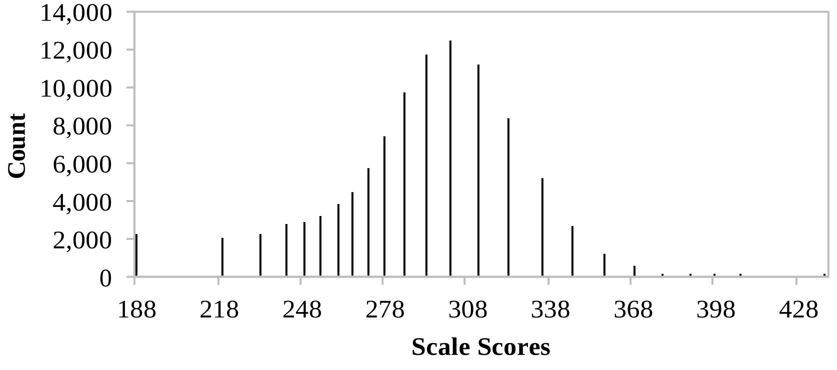


Table 2.4.3.4.2

Scale Score Descriptive Statistics: Writ 6-8 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	67,227	188	470	344.99	27.02
7	53,575	188	470	354.62	26.58
8	47,805	188	520	361.70	26.17
Total	168,607	188	520	352.79	27.52

Figure 2.4.3.4.2
Scale Scores: Writ 6-8 B/C S403 Online

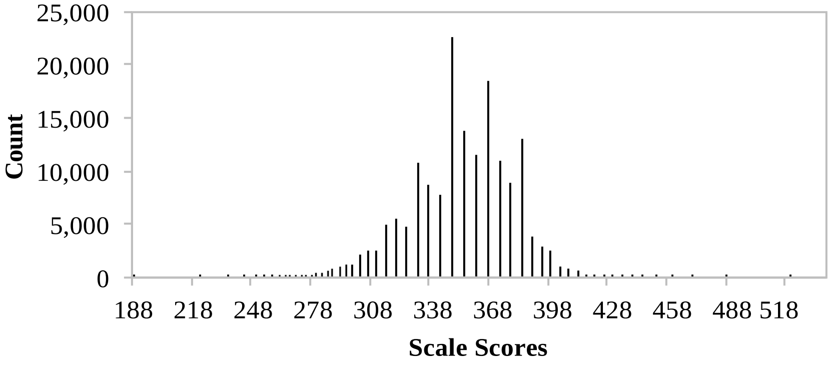
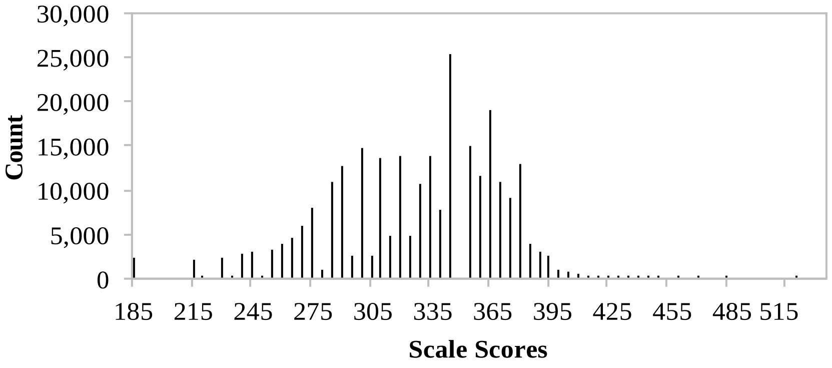


Table 2.4.3.4.3

Scale Score Descriptive Statistics: Writ 6-8 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	98,248	188	470	325.33	40.71
7	88,369	188	470	329.18	43.05
8	82,077	188	520	333.49	44.69
Total	268,694	188	520	329.09	42.86

Figure 2.4.3.4.3
Scale Scores: Writ 6-8 S403 Online



2.4.3.5 *Grades 9–12*

Table 2.4.3.5.1

Scale Score Descriptive Statistics: Writ 9-12 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	41,531	232	454	322.10	42.35
10	28,776	232	454	332.51	38.08
11	23,202	232	445	338.77	37.16
12	15,363	232	454	339.84	38.12
Total	108,872	232	454	330.91	40.27

Figure 2.4.3.5.1
Scale Scores: Writ 9-12 AS403 Online

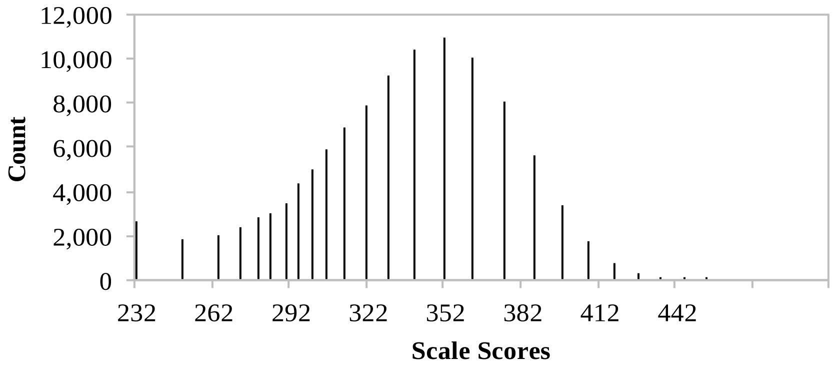


Table 2.4.3.5.2

Scale Score Descriptive Statistics: Writ 9-12 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	53,578	232	504	372.24	32.62
10	49,396	232	535	375.04	33.19
11	46,009	232	486	379.14	33.52
12	35,772	232	486	377.16	35.08
Total	184,755	232	535	375.66	33.59

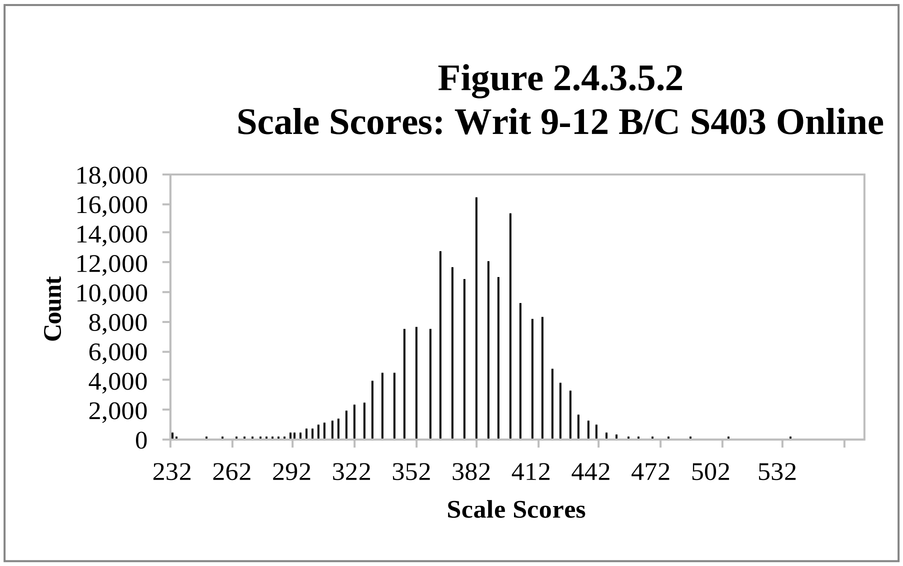
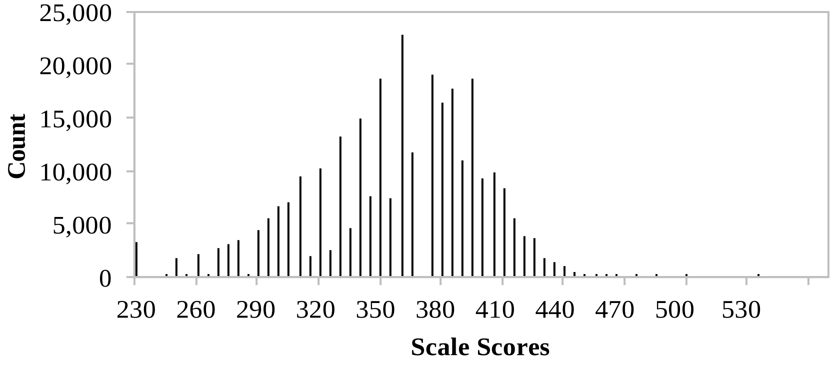


Table 2.4.3.5.3

Scale Score Descriptive Statistics: Writ 9-12 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	95,109	232	504	350.34	44.73
10	78,172	232	535	359.39	40.63
11	69,211	232	486	365.61	39.66
12	51,135	232	486	365.95	39.88
Total	293,627	232	535	359.07	42.17

Figure 2.4.3.5.3
Scale Scores: Writ 9-12 S403 Online



2.4.4 Speaking

2.4.4.1 Grade 1

Table 2.4.4.1.1

Scale Score Descriptive Statistics: Spek 1 Pre-A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	5,158	106	175	154.96	26.41
Total	5,158	106	175	154.96	26.41

Figure 2.4.4.1.1

Scale Scores: Spek 1 Pre-AS403 Online

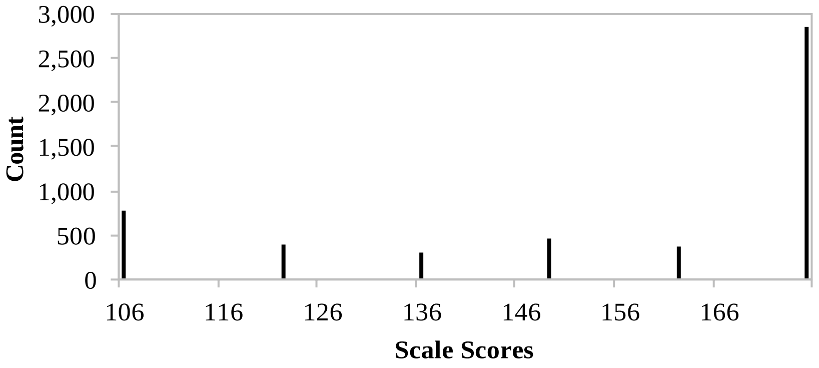


Table 2.4.4.1.2

Scale Score Descriptive Statistics: Spek 1 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	61,323	106	384	232.96	54.42
Total	61,323	106	384	232.96	54.42

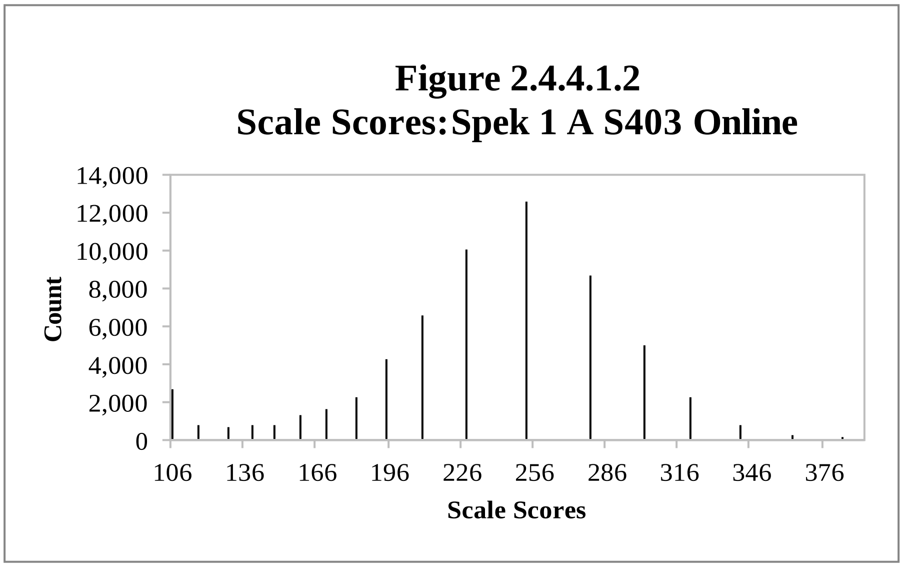


Table 2.4.4.1.3

Scale Score Descriptive Statistics: Spek 1 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	111,043	106	403	266.18	42.88
Total	111,043	106	403	266.18	42.88

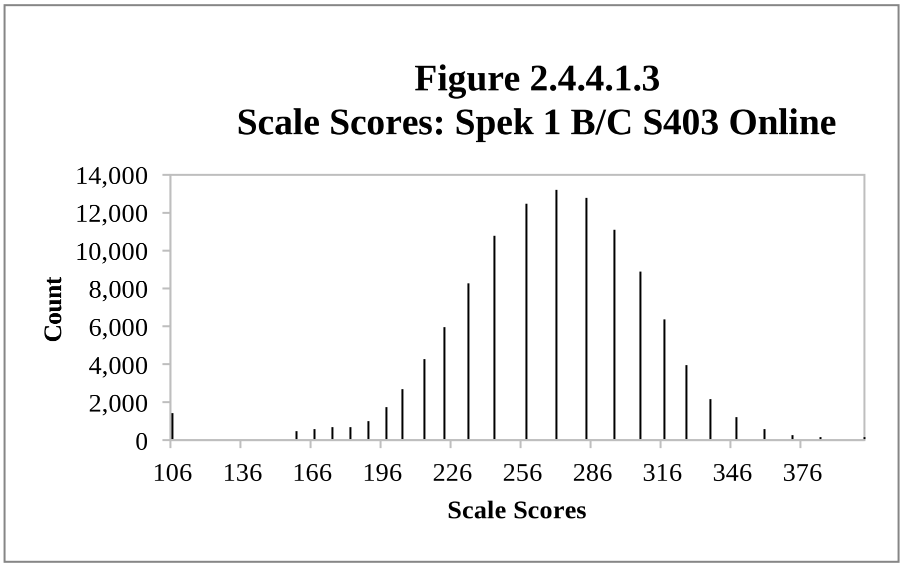
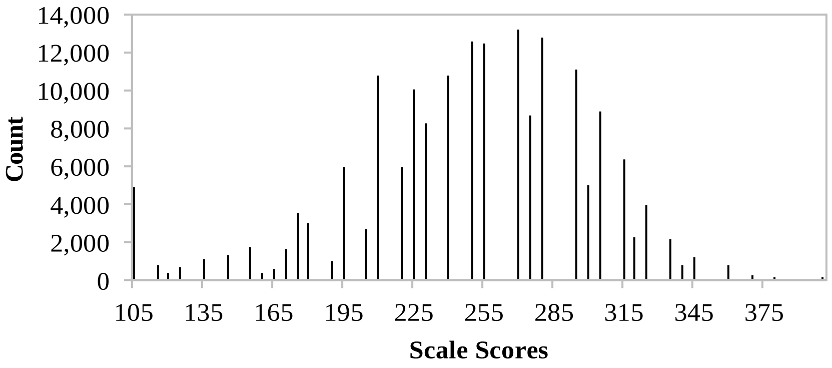


Table 2.4.4.1.4

Scale Score Descriptive Statistics: Spek 1 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	177,524	106	403	251.47	52.13
Total	177,524	106	403	251.47	52.13

Figure 2.4.4.1.4
Scale Scores: Spek 1 S403 Online



2.4.4.2 *Grades 2–3*

Table 2.4.4.2.1

Scale Score Descriptive Statistics: Spek 2-3 Pre-A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	6,879	118	169	156.83	19.61
3	9,480	118	169	156.92	19.43
Total	16,359	118	169	156.88	19.51

Figure 2.4.4.2.1

Scale Scores: Spek 2-3 Pre-AS403 Online

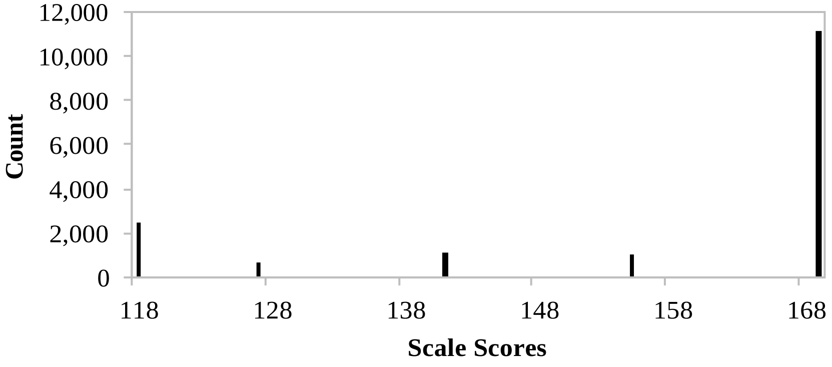


Table 2.4.4.2.2

Scale Score Descriptive Statistics: Spek 2-3 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	41,584	118	383	236.86	47.88
3	37,429	118	383	253.09	45.41
Total	79,013	118	383	244.55	47.42

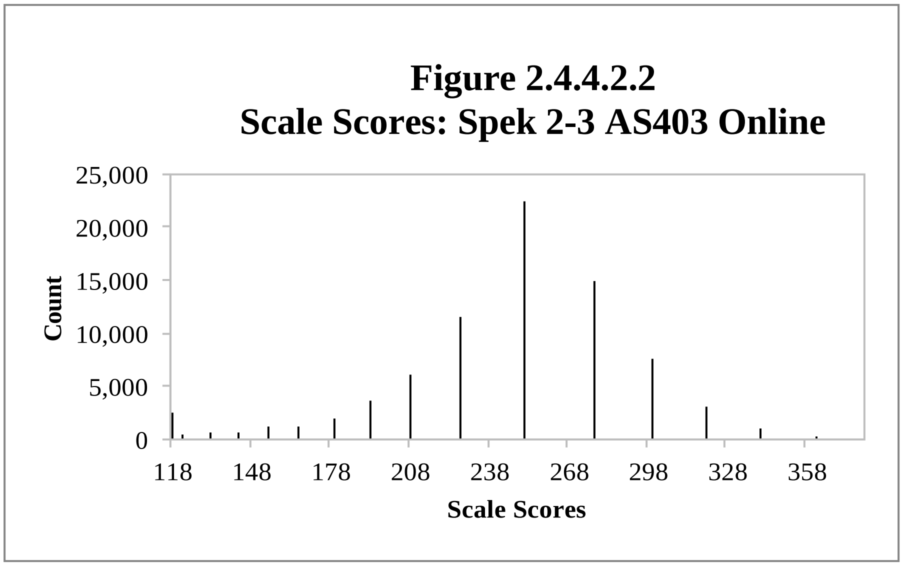


Table 2.4.4.2.3

Scale Score Descriptive Statistics: Spek 2-3 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	133,785	118	425	272.04	37.40
3	136,710	118	425	285.33	35.89
Total	270,495	118	425	278.76	37.24

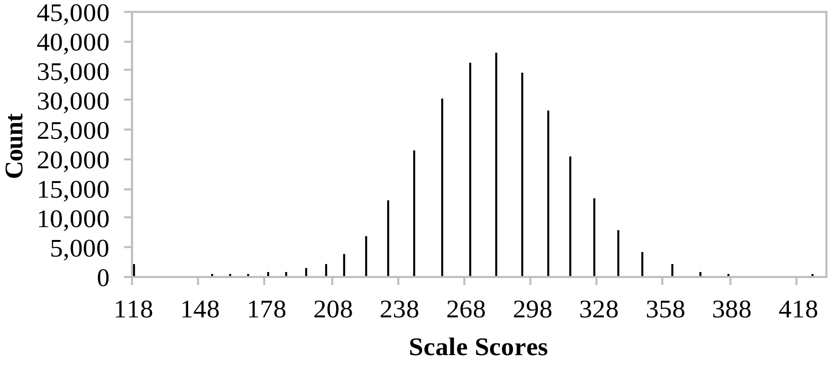
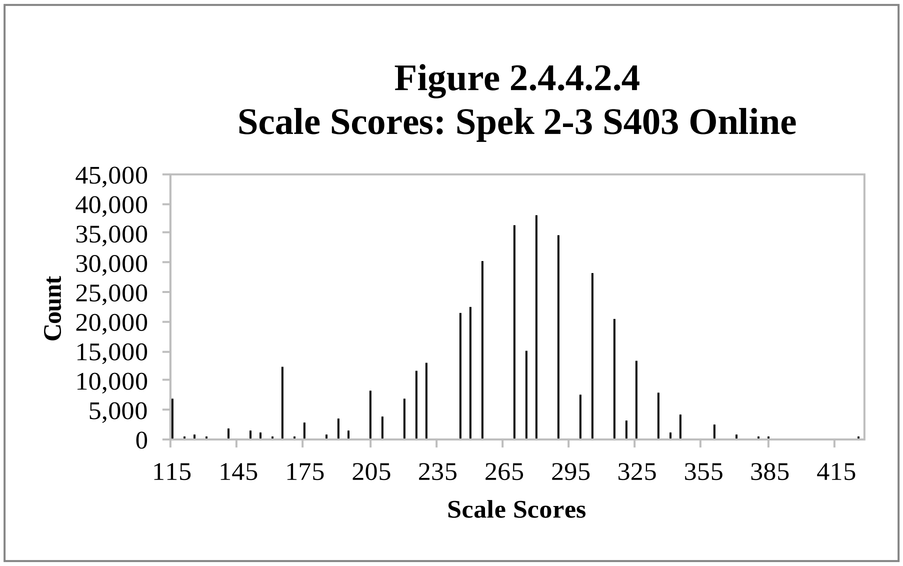


Figure 2.4.4.2.3
Scale Scores: Spek 2-3 B/C S403 Online

Table 2.4.4.2.4

Scale Score Descriptive Statistics: Spek 2-3 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	182,248	118	425	259.67	46.84
3	183,619	118	425	272.13	47.83
Total	365,867	118	425	265.92	47.75



2.4.4.3 Grades 4–5

Table 2.4.4.3.1

Scale Score Descriptive Statistics: Spek 4-5 Pre-A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	2,390	130	194	170.21	25.42
5	3,621	130	194	173.94	24.44
Total	6,011	130	194	172.46	24.90

Figure 2.4.4.3.1
Scale Scores: Spek 4-5 Pre-AS403 Online

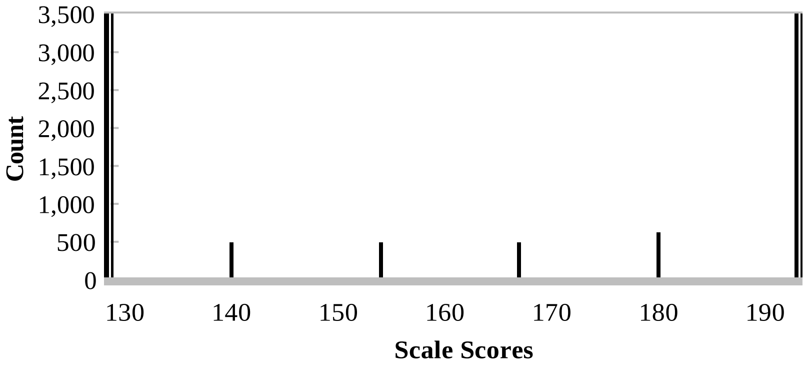


Table 2.4.4.3.2

Scale Score Descriptive Statistics: Spek 4-5 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	15,726	130	427	259.36	51.03
5	12,275	130	427	264.58	49.26
Total	28,001	130	427	261.65	50.33

Figure 2.4.4.3.2

Scale Scores: Spek 4-5 AS403 Online

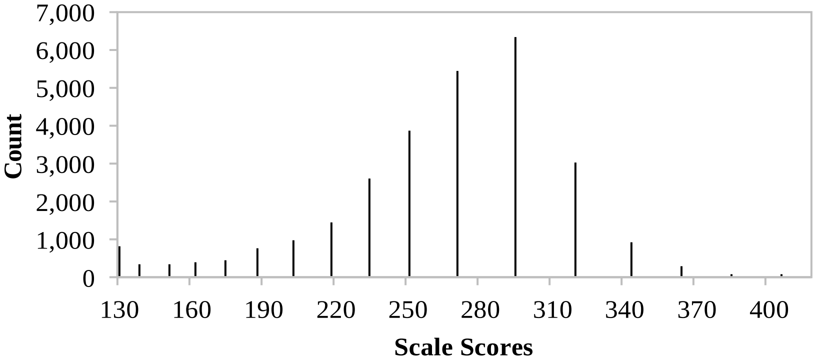


Table 2.4.4.3.3

Scale Score Descriptive Statistics: Spek 4-5 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	155,646	130	450	314.81	39.51
5	124,000	130	450	318.43	40.81
Total	279,646	130	450	316.42	40.13

Figure 2.4.4.3.3
Scale Scores: Spek 4-5 B/C S403 Online

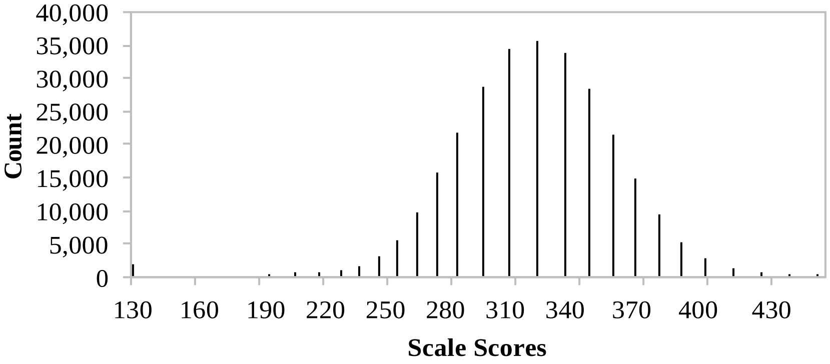
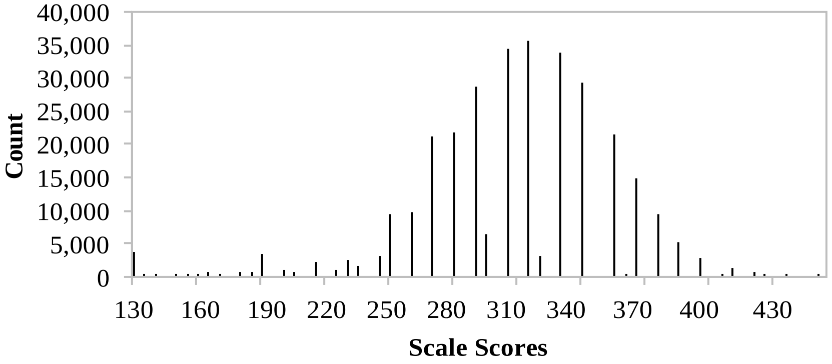


Table 2.4.4.3.4

Scale Score Descriptive Statistics: Spek 4-5 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	173,762	130	450	307.80	46.47
5	139,896	130	450	309.97	49.27
Total	313,658	130	450	308.77	47.75

Figure 2.4.4.3.4
Scale Scores: Spek 4-5 S403 Online



2.4.4.4 *Grades 6–8*

Table 2.4.4.4.1

Scale Score Descriptive Statistics: Spek 6-8 Pre-A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	2,811	148	201	186.79	20.25
7	3,854	148	201	187.49	19.86
8	4,829	148	201	187.65	19.53
Total	11,494	148	201	187.39	19.82

Figure 2.4.4.4.1
Scale Scores: Spek 6-8 Pre-AS403 Online

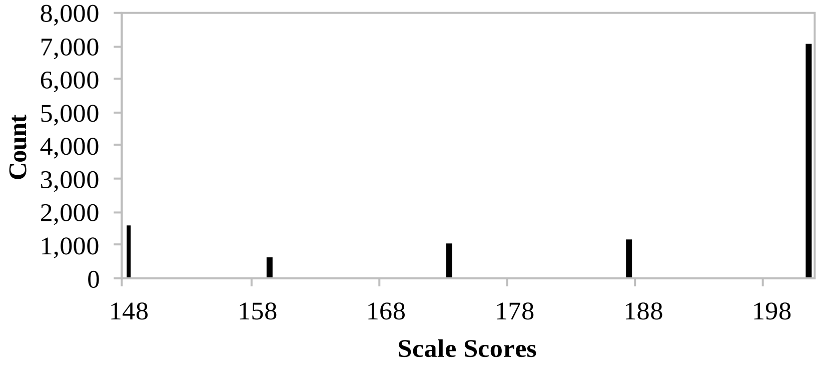


Table 2.4.4.4.2

Scale Score Descriptive Statistics: Spek 6-8 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	17,887	148	419	279.70	52.06
7	15,034	148	440	274.90	52.86
8	24,908	148	440	289.11	53.67
Total	57,829	148	440	282.50	53.31

Figure 2.4.4.4.2
Scale Scores: Spek 6-8 AS403 Online

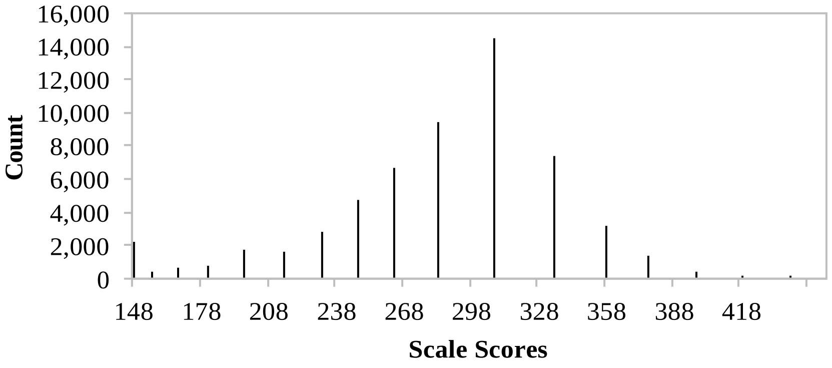


Table 2.4.4.4.3

Scale Score Descriptive Statistics: Spek 6-8 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	80,038	148	467	324.30	39.23
7	71,216	148	467	327.22	42.07
8	54,262	148	467	336.39	42.71
Total	205,516	148	467	328.51	41.45

Figure 2.4.4.4.3
Scale Scores: Spek 6-8 B/C S403 Online

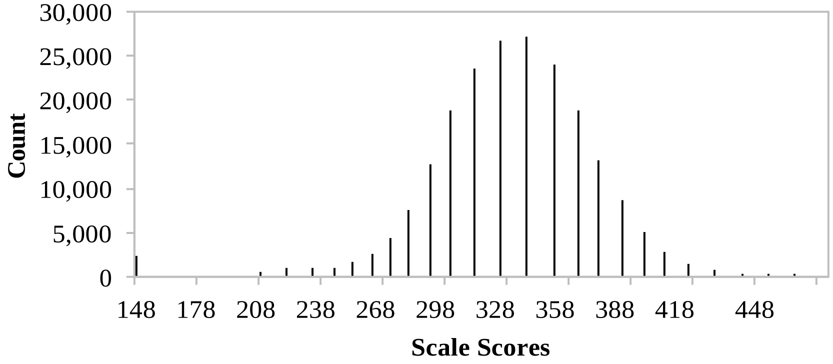
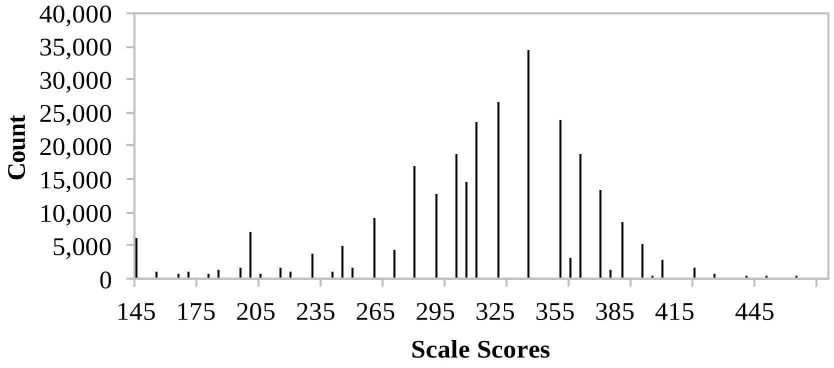


Table 2.4.4.4.4

Scale Score Descriptive Statistics: Spek 6-8 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	100,736	148	467	312.55	49.58
7	90,104	148	467	312.52	54.39
8	83,999	148	467	313.82	58.99
Total	274,839	148	467	312.92	54.17

Figure 2.4.4.4.4
Scale Scores: Spek 6-8 S403 Online



2.4.4.5 *Grades 9–12*

Table 2.4.4.5.1

Scale Score Descriptive Statistics: Spek 9-12 Pre-A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	5,440	172	225	212.74	19.46
10	6,455	172	225	217.21	16.46
11	6,209	172	225	218.79	15.17
12	4,998	172	225	218.29	15.90
Total	23,102	172	225	216.82	16.93

Figure 2.4.4.5.1
Scale Scores: Spek 9-12 Pre-AS403 Online

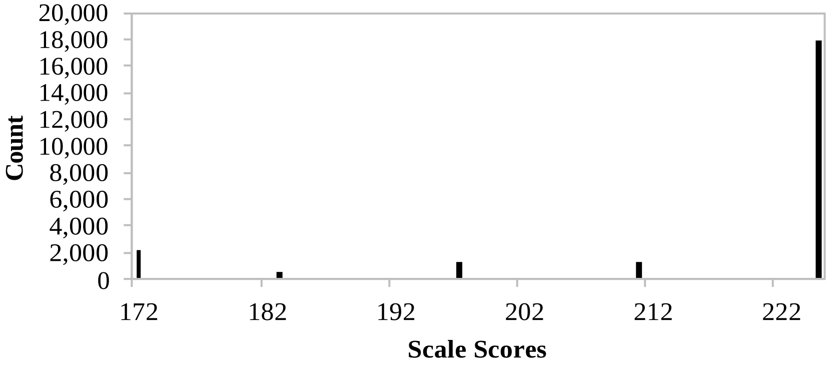


Table 2.4.4.5.2

Scale Score Descriptive Statistics: Spek 9-12 A S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	54,492	172	455	292.14	49.74
10	31,840	172	434	298.11	47.14
11	14,041	172	434	293.24	47.84
12	23,318	172	455	309.39	49.71
Total	123,691	172	455	297.05	49.28

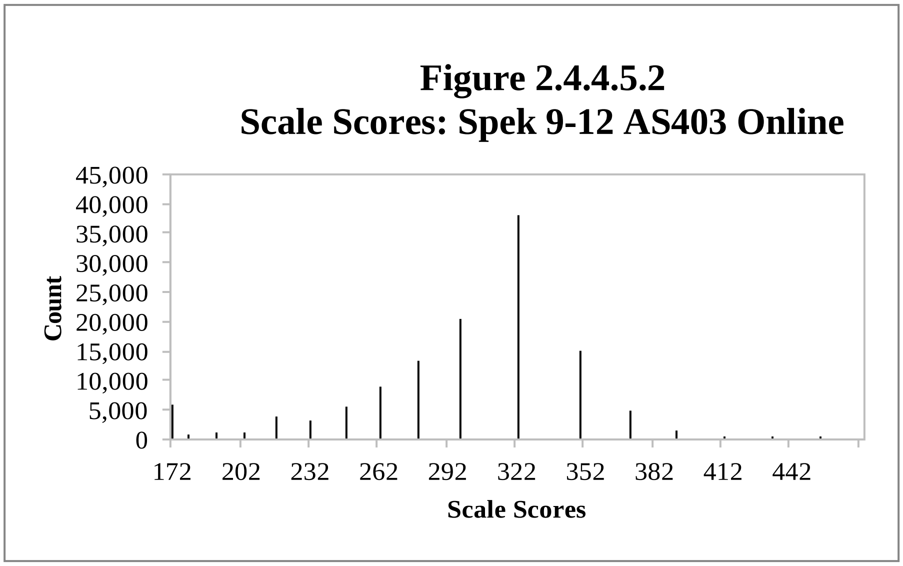


Table 2.4.4.5.3

Scale Score Descriptive Statistics: Spek 9-12 B/C S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	37,139	172	476	335.64	38.20
10	41,507	172	476	335.30	40.57
11	49,948	172	476	332.75	43.65
12	23,922	172	476	340.76	43.80
Total	152,516	172	476	335.40	41.65

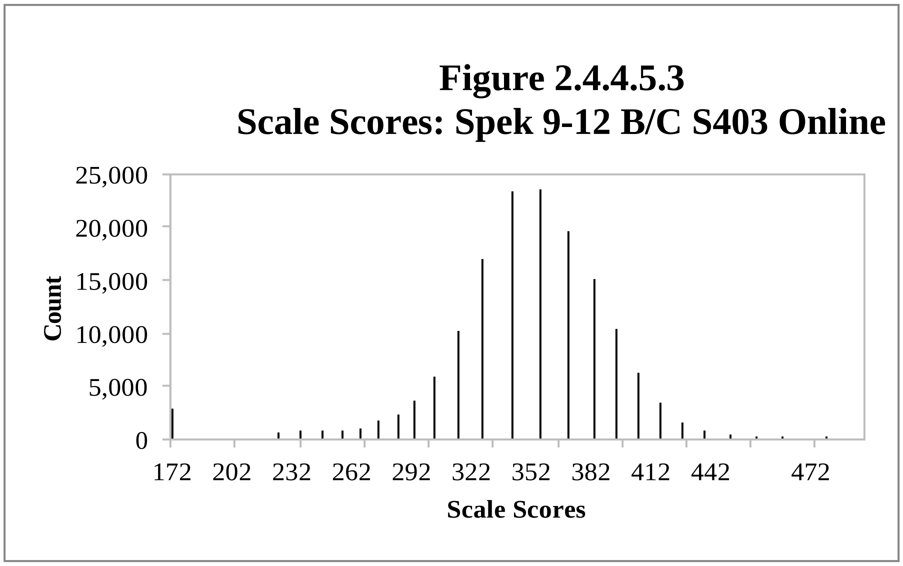
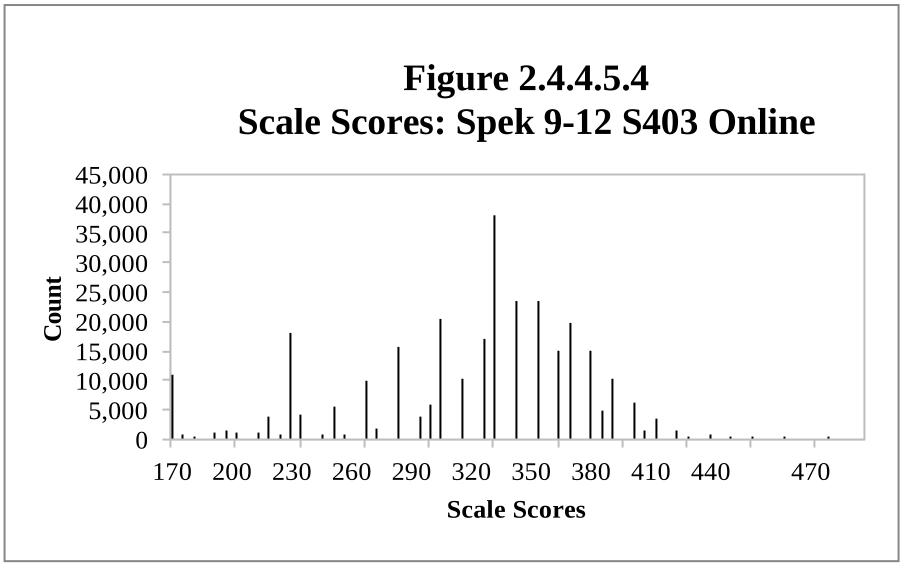


Table 2.4.4.5.4

Scale Score Descriptive Statistics: Spek 9-12 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	97,071	172	476	304.33	53.82
10	79,802	172	476	310.91	53.38
11	70,198	172	476	314.77	54.51
12	52,238	172	476	315.04	56.73
Total	299,309	172	476	310.40	54.57



2.5 Proficiency Level Distribution

The figures and tables in this section provide information on the proficiency level distribution of the students who took the test form based on their performance. Proficiency level distribution is presented by grade-level cluster. For Writing and Speaking, it is also presented by grade-level cluster and tier. Note that the Speaking Pre-A tier is intended for students at the beginning stages of English language proficiency, and students assigned to this tier are not expected to show proficiency above PL 1.

In the tables presented in this section, each row shows, by grade and by total for the grade-level cluster:

- The WIDA proficiency level designation (1–6)
- The number of students (count) whose performance on the test form placed them into that proficiency level in the domain being tested
- The percentage of students, out of the total number of students taking the form, who were placed into that proficiency level in the domain being tested

In the figure, the horizontal axis shows the six WIDA proficiency levels. The vertical axis shows the percentage of students. Each bar shows the percentage of students who were placed into each proficiency level in the domain being tested on this test form.

2.5.1 Listening

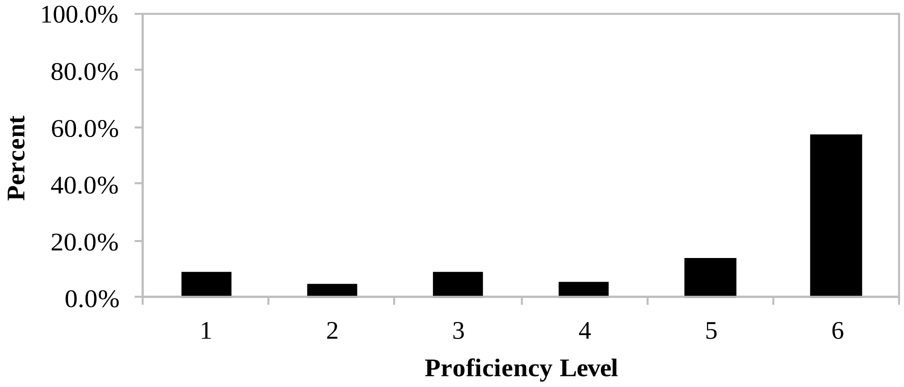
2.5.1.1 Grade 1

Table 2.5.1.1

Proficiency Level Distribution: List 1 S403 Online

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	16,004	8.84%	16,004	8.84%
2	8,117	4.48%	8,117	4.48%
3	16,930	9.35%	16,930	9.35%
4	10,511	5.80%	10,511	5.80%
5	25,558	14.11%	25,558	14.11%
6	103,987	57.42%	103,987	57.42%
Total	181,107	100.00%	181,107	100.00%

Figure 2.5.1.1
Proficiency Level: List 1 S403 Online



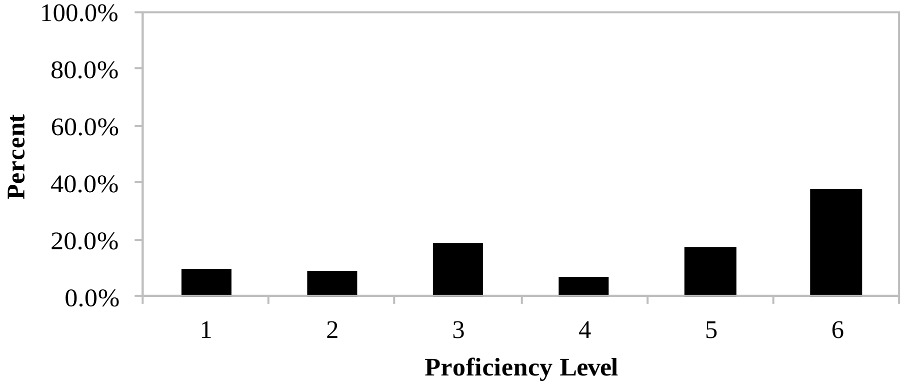
2.5.1.2 *Grades 2–3*

Table 2.5.1.2

Proficiency Level Distribution: List 2-3 S403 Online

	Grade 2		Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	19,280	10.37%	16,257	8.70%	35,537	9.54%
2	17,281	9.30%	16,541	8.85%	33,822	9.08%
3	39,292	21.14%	30,252	16.19%	69,544	18.66%
4	14,007	7.54%	11,877	6.36%	25,884	6.95%
5	34,505	18.57%	31,826	17.04%	66,331	17.80%
6	61,478	33.08%	80,068	42.86%	141,546	37.98%
Total	185,843	100.00%	186,821	100.00%	372,664	100.00%

Figure 2.5.1.2
Proficiency Level: List 2-3 S403 Online



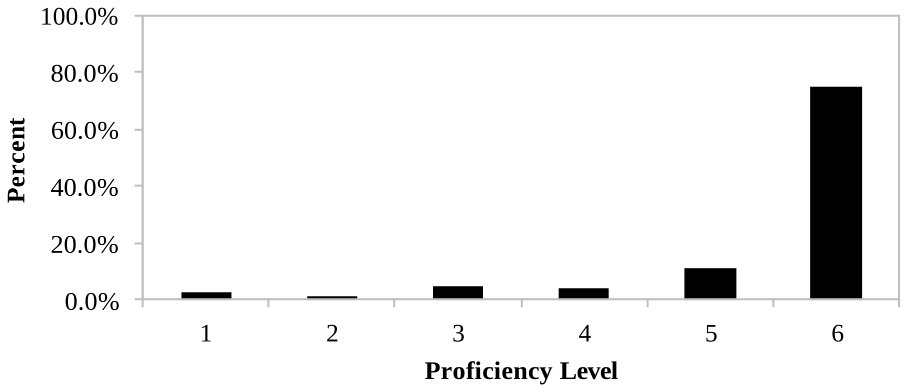
2.5.1.3 *Grades 4–5*

Table 2.5.1.3

Proficiency Level Distribution: List 4-5 S403 Online

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	4,149	2.32%	4,658	3.23%	8,807	2.73%
2	2,530	1.41%	2,776	1.93%	5,306	1.64%
3	9,527	5.33%	7,066	4.90%	16,593	5.14%
4	7,668	4.29%	6,242	4.33%	13,910	4.31%
5	20,003	11.19%	16,473	11.43%	36,476	11.29%
6	134,956	75.46%	106,916	74.18%	241,872	74.89%
Total	178,833	100.00%	144,131	100.00%	322,964	100.00%

Figure 2.5.1.3
Proficiency Level: List 4-5 S403 Online



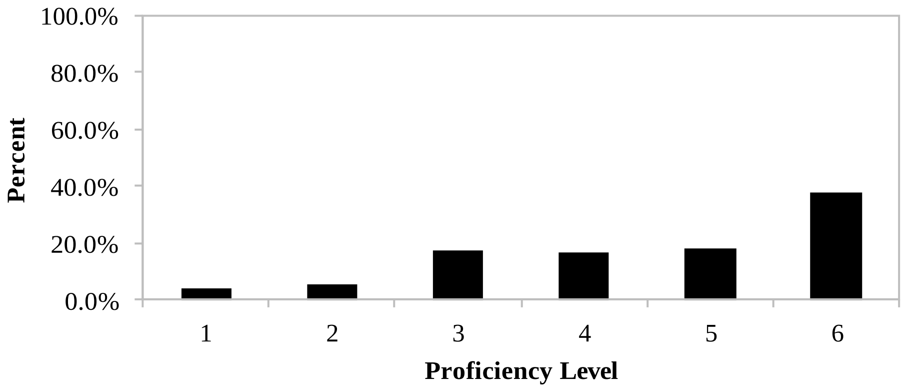
2.5.1.4 Grades 6–8

Table 2.5.1.4

Proficiency Level Distribution: List 6-8 S403 Online

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,982	2.92%	3,934	4.33%	4,869	5.75%	11,785	4.24%
2	4,475	4.38%	5,261	5.79%	6,449	7.61%	16,185	5.83%
3	16,732	16.38%	17,538	19.29%	15,109	17.83%	49,379	17.78%
4	16,370	16.03%	16,225	17.85%	13,574	16.02%	46,169	16.62%
5	22,048	21.59%	15,114	16.62%	12,484	14.73%	49,646	17.87%
6	39,526	38.70%	32,844	36.13%	32,253	38.06%	104,623	37.66%
Total	102,133	100.00%	90,916	100.00%	84,738	100.00%	277,787	100.00%

Figure 2.5.1.4
Proficiency Level: List 6-8 S403 Online



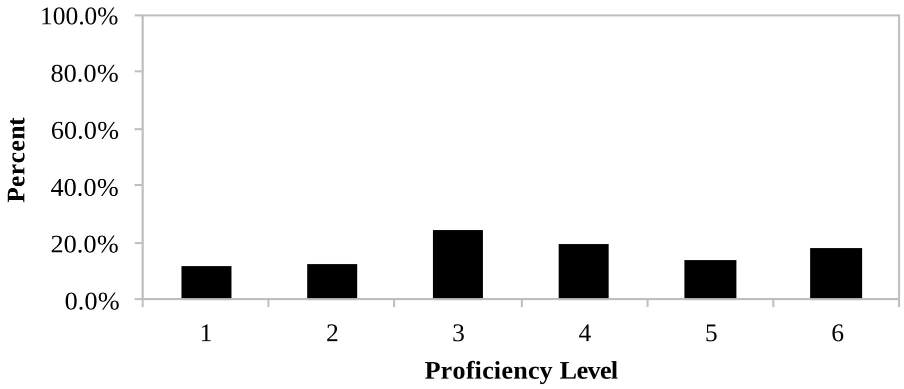
2.5.1.5 Grades 9–12

Table 2.5.1.5

Proficiency Level Distribution: List 9-12 S403 Online

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	7,598	7.80%	9,665	12.03%	10,424	14.57%	7,971	15.19%	35,658	11.82%
2	14,155	14.53%	10,525	13.10%	6,931	9.69%	6,730	12.83%	38,341	12.71%
3	24,186	24.82%	18,780	23.37%	18,215	25.47%	11,974	22.82%	73,155	24.24%
4	19,509	20.02%	15,406	19.18%	12,864	17.98%	10,197	19.44%	57,976	19.21%
5	13,233	13.58%	10,491	13.06%	10,669	14.92%	6,843	13.04%	41,236	13.66%
6	18,750	19.24%	15,477	19.26%	12,426	17.37%	8,746	16.67%	55,399	18.36%
Total	97,431	100.00%	80,344	100.00%	71,529	100.00%	52,461	100.00%	301,765	100.00%

Figure 2.5.1.5
Proficiency Level: List 9-12 S403 Online



2.5.2 Reading

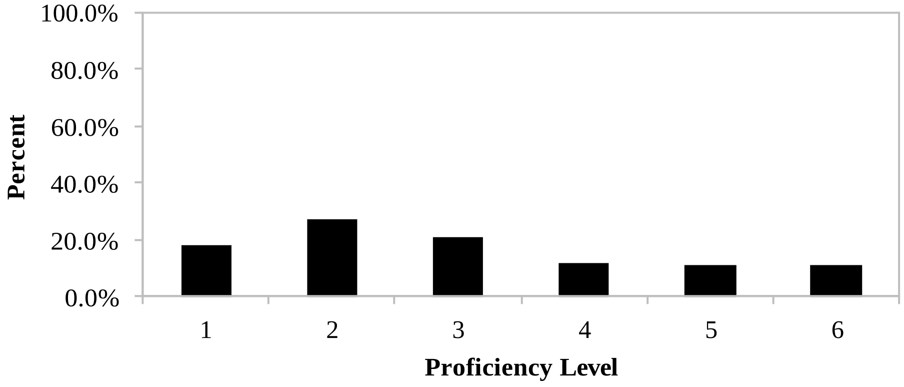
2.5.2.1 *Grade 1*

Table 2.5.2.1

Proficiency Level Distribution: Read 1 S403 Online

Level	Grade 1		Total	
	Count	**Percent**	**Count**	**Percent**
1	33,816	18.44%	33,816	18.44%
2	49,853	27.18%	49,853	27.18%
3	38,074	20.76%	38,074	20.76%
4	21,768	11.87%	21,768	11.87%
5	19,998	10.90%	19,998	10.90%
6	19,912	10.86%	19,912	10.86%
Total	183,421	100.00%	183,421	100.00%

Figure 2.5.2.1
Proficiency Level: Read 1 S403 Online



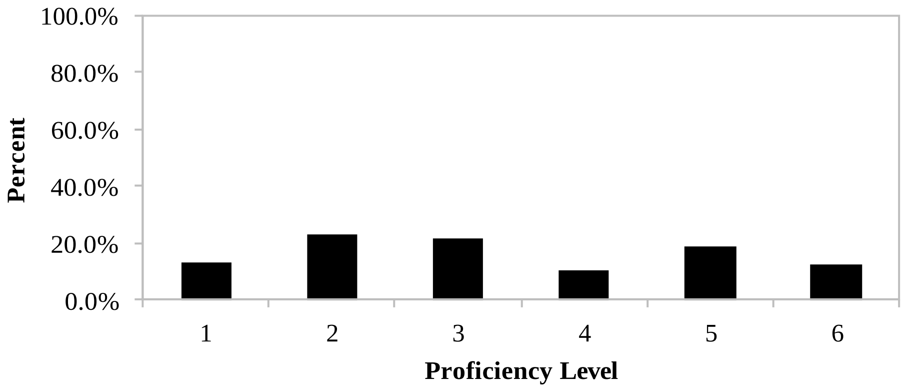
2.5.2.2 Grades 2–3

Table 2.5.2.2

Proficiency Level Distribution: Read 2-3 S403 Online

Level	Grade 2		Grade 3		Total	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	22,892	12.28%	27,673	14.80%	50,565	13.54%
2	42,882	23.00%	44,247	23.67%	87,129	23.34%
3	45,239	24.27%	36,203	19.36%	81,442	21.81%
4	21,205	11.38%	17,754	9.50%	38,959	10.43%
5	34,136	18.31%	35,013	18.73%	69,149	18.52%
6	20,061	10.76%	26,066	13.94%	46,127	12.35%
Total	186,415	100.00%	186,956	100.00%	373,371	100.00%

Figure 2.5.2.2
Proficiency Level: Read 2-3 S403 Online



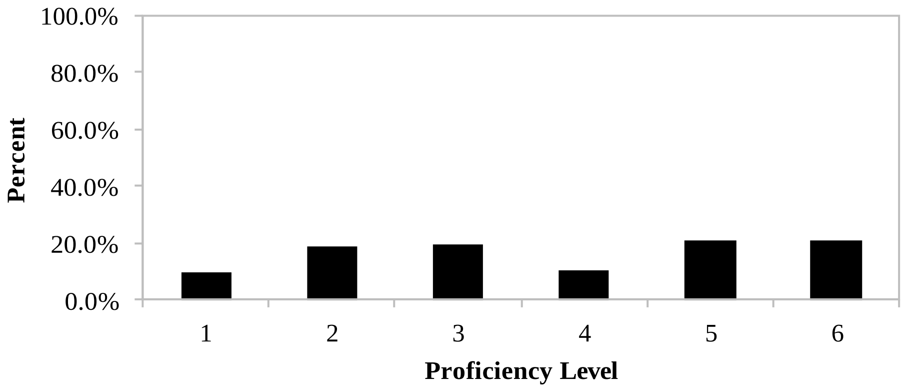
2.5.2.3 *Grades 4–5*

Table 2.5.2.3

Proficiency Level Distribution: Read 4-5 S403 Online

Level	**Grade 4**		**Grade 5**		**Total**	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	14,447	8.19%	16,536	11.63%	30,983	9.72%
2	30,894	17.50%	28,253	19.88%	59,147	18.56%
3	33,552	19.01%	28,215	19.85%	61,767	19.38%
4	21,406	12.13%	11,917	8.38%	33,323	10.46%
5	38,129	21.60%	27,748	19.52%	65,877	20.67%
6	38,064	21.57%	29,480	20.74%	67,544	21.20%
Total	176,492	100.00%	142,149	100.00%	318,641	100.00%

Figure 2.5.2.3
Proficiency Level: Read 4-5 S403 Online



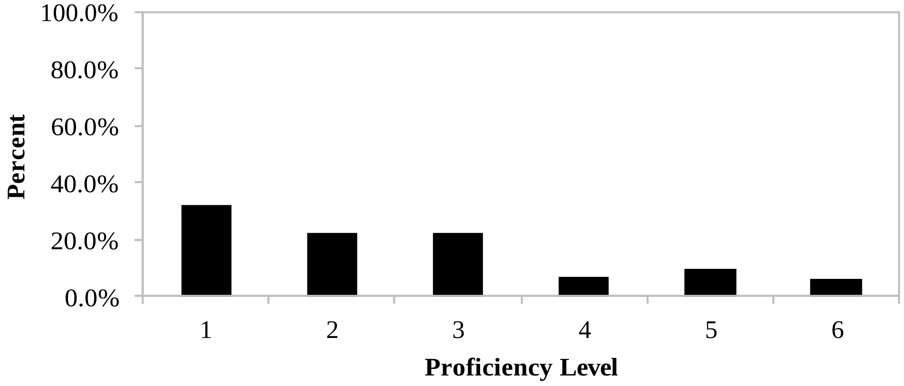
2.5.2.4 *Grades 6–8*

Table 2.5.2.4

Proficiency Level Distribution: Read 6-8 S403 Online

Level	**Grade 6**		**Grade 7**		**Grade 8**		**Total**	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	31,825	31.19%	29,924	32.78%	27,146	32.07%	88,895	31.98%
2	23,449	22.98%	19,833	21.73%	19,502	23.04%	62,784	22.59%
3	24,836	24.34%	21,103	23.12%	16,662	19.68%	62,601	22.52%
4	7,856	7.70%	6,360	6.97%	5,433	6.42%	19,649	7.07%
5	9,764	9.57%	7,596	8.32%	10,129	11.96%	27,489	9.89%
6	4,292	4.21%	6,461	7.08%	5,786	6.83%	16,539	5.95%
Total	102,022	100.00%	91,277	100.00%	84,658	100.00%	277,957	100.00%

Figure 2.5.2.4
Proficiency Level: Read 6-8 S403 Online



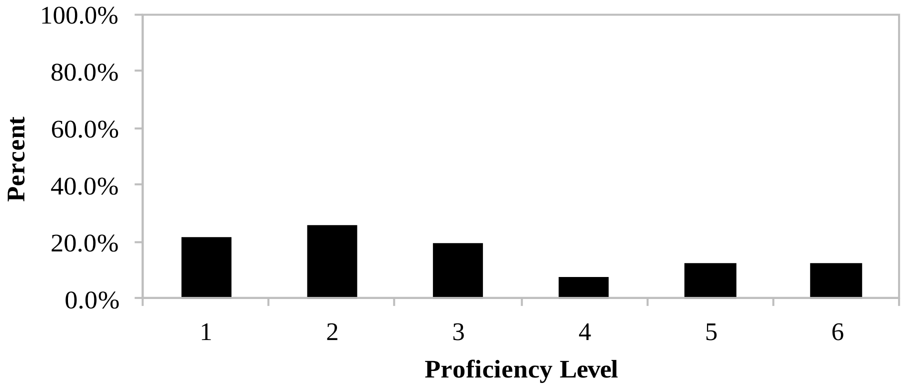
2.5.2.5 *Grades 9–12*

Table 2.5.2.5

Proficiency Level Distribution: Read 9-12 S403 Online

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	23,743	24.62%	16,769	21.14%	13,656	19.46%	10,977	21.16%	65,145	21.87%
2	24,163	25.06%	19,604	24.71%	18,573	26.47%	14,469	27.89%	76,809	25.79%
3	19,044	19.75%	16,490	20.79%	12,974	18.49%	10,204	19.67%	58,712	19.71%
4	6,243	6.47%	6,432	8.11%	5,751	8.20%	4,276	8.24%	22,702	7.62%
5	11,403	11.83%	11,284	14.22%	8,817	12.56%	6,224	12.00%	37,728	12.67%
6	11,833	12.27%	8,751	11.03%	10,402	14.82%	5,737	11.06%	36,723	12.33%
Total	96,429	100.00%	79,330	100.00%	70,173	100.00%	51,887	100.00%	297,819	100.00%

Figure 2.5.2.5
Proficiency Level: Read9-12 S403 Online



2.5.3 Writing

2.5.3.1 Grade 1

Table 2.5.3.1.1

Proficiency Level Distribution: Writ 1 A S403 Online

Level	Grade 1		Total	
	Count	Percent	Count	Percent
1	35,964	23.27%	35,964	23.27%
2	107,345	69.46%	107,345	69.46%
3	11,240	7.27%	11,240	7.27%
4	1	0.00%	1	0.00%
5	0	0.00%	0	0.00%
6	0	0.00%	0	0.00%
Total	154,550	100.00%	154,550	100.00%

Figure 2.5.3.1.1
Proficiency Level: Writ 1 A S403 Online

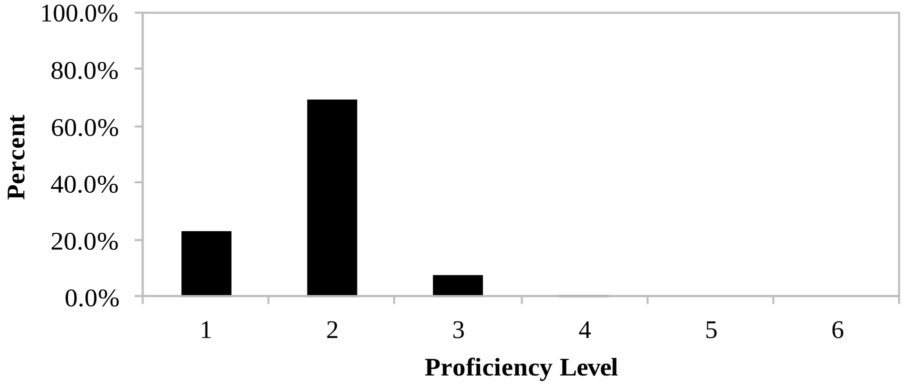


Table 2.5.3.1.2

Proficiency Level Distribution: Writ 1 B/C S403 Online

Level	Grade 1		Total	
	Count	Percent	Count	Percent
1	596	1.70%	596	1.70%
2	5,567	15.86%	5,567	15.86%
3	27,420	78.14%	27,420	78.14%
4	1,505	4.29%	1,505	4.29%
5	4	0.01%	4	0.01%
6	1	0.00%	1	0.00%
Total	35,093	100.00%	35,093	100.00%

Figure 2.5.3.1.2
Proficiency Level: Writ 1 B/C S403 Online

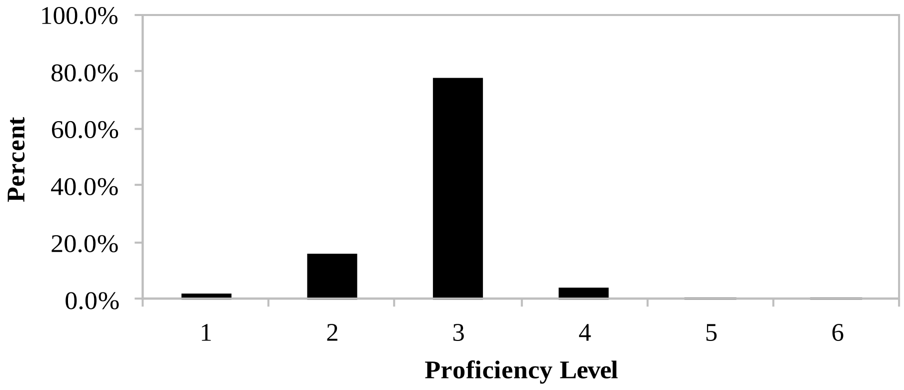
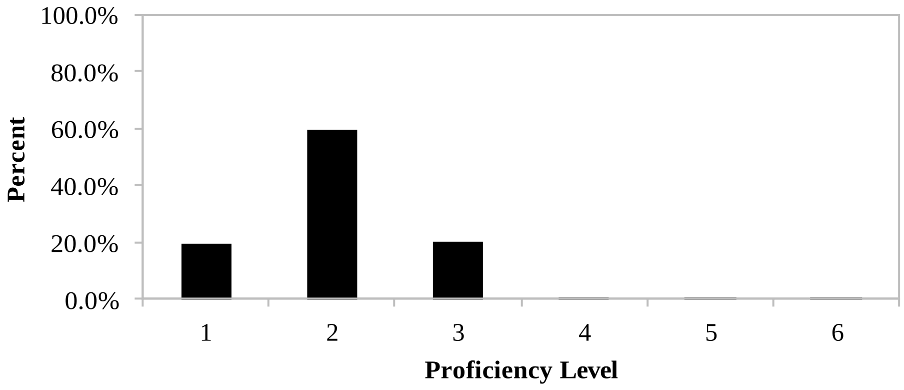


Table 2.5.3.1.3

Proficiency Level Distribution: Writ 1 S403 Online

Level	Grade 1		Total	
	Count	Percent	Count	Percent
1	36,560	19.28%	36,560	19.28%
2	112,912	59.54%	112,912	59.54%
3	38,660	20.39%	38,660	20.39%
4	1,506	0.79%	1,506	0.79%
5	4	0.00%	4	0.00%
6	1	0.00%	1	0.00%
Total	189,643	100.00%	189,643	100.00%

Figure 2.5.3.1.3
Proficiency Level: Writ 1 S403 Online



2.5.3.2 *Grades 2–3*

Table 2.5.3.2.1

Proficiency Level Distribution: Writ 2-3 A S403 Online

	Grade 2		**Grade 3**		**Total**	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	13,151	23.41%	8,905	23.78%	22,056	23.56%
2	22,624	40.28%	18,528	49.48%	41,152	43.96%
3	20,389	36.30%	9,993	26.69%	30,382	32.45%
4	3	0.01%	21	0.06%	24	0.03%
5	0	0.00%	0	0.00%	0	0.00%
6	0	0.00%	0	0.00%	0	0.00%
Total	56,167	100.00%	37,447	100.00%	93,614	100.00%

Figure 2.5.3.2.1
Proficiency Level: Writ 2-3 A S403 Online

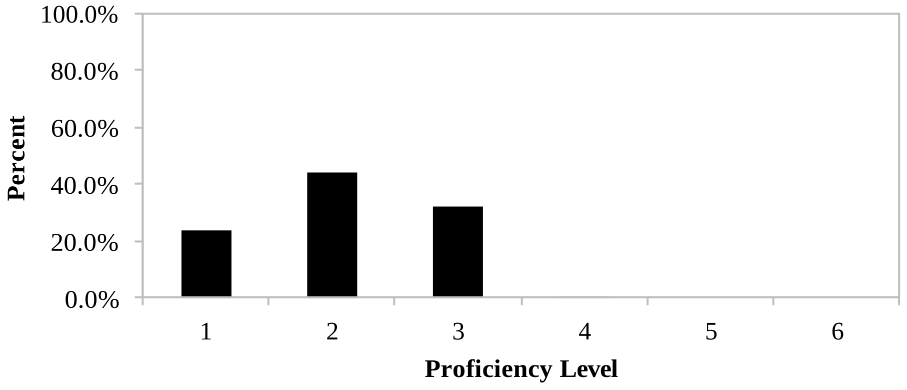


Table 2.5.3.2.2

Proficiency Level Distribution: Writ 2-3 B/C S403 Online

	Grade 2		Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	870	0.63%	446	0.28%	1,316	0.44%
2	12,607	9.09%	5,847	3.70%	18,454	6.22%
3	106,511	76.80%	107,082	67.78%	213,593	72.00%
4	18,500	13.34%	43,475	27.52%	61,975	20.89%
5	189	0.14%	1,066	0.67%	1,255	0.42%
6	7	0.01%	63	0.04%	70	0.02%
Total	138,684	100.00%	157,979	100.00%	296,663	100.00%

Figure 2.5.3.2.2
Proficiency Level: Writ 2-3 B/C S403 Online

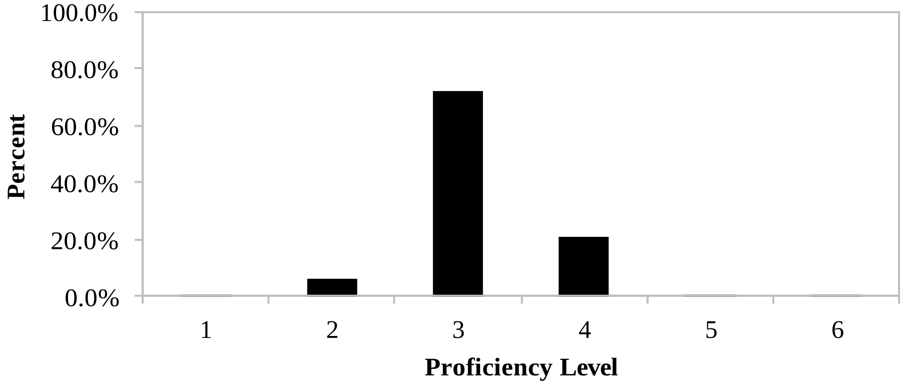
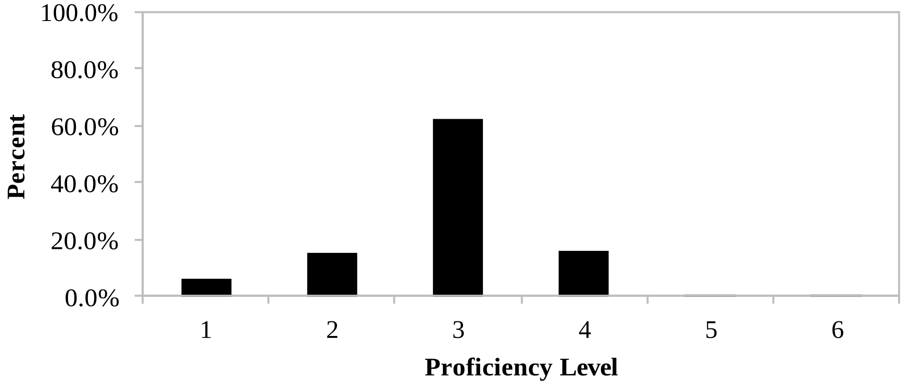


Table 2.5.3.2.3

Proficiency Level Distribution: Writ 2-3 S403 Online

Level	Grade 2		Grade 3		Total	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	14,021	7.20%	9,351	4.78%	23,372	5.99%
2	35,231	18.08%	24,375	12.47%	59,606	15.27%
3	126,900	65.13%	117,075	59.91%	243,975	62.51%
4	18,503	9.50%	43,496	22.26%	61,999	15.89%
5	189	0.10%	1,066	0.55%	1,255	0.32%
6	7	0.00%	63	0.03%	70	0.02%
Total	194,851	100.00%	195,426	100.00%	390,277	100.00%

Figure 2.5.3.2.3
Proficiency Level: Writ 2-3 S403 Online



2.5.3.3 Grades 4–5

Table 2.5.3.3.1

Proficiency Level Distribution: Writ 4-5 A S403 Online

Level	Grade 4		Grade 5		Total	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	6,526	29.38%	4,940	22.46%	11,466	25.94%
2	6,685	30.09%	8,000	36.37%	14,685	33.22%
3	8,834	39.77%	8,923	40.57%	17,757	40.17%
4	168	0.76%	133	0.60%	301	0.68%
5	0	0.00%	0	0.00%	0	0.00%
6	0	0.00%	0	0.00%	0	0.00%
Total	22,213	100.00%	21,996	100.00%	44,209	100.00%

Figure 2.5.3.3.1

Proficiency Level: Writ 4-5 A S403 Online

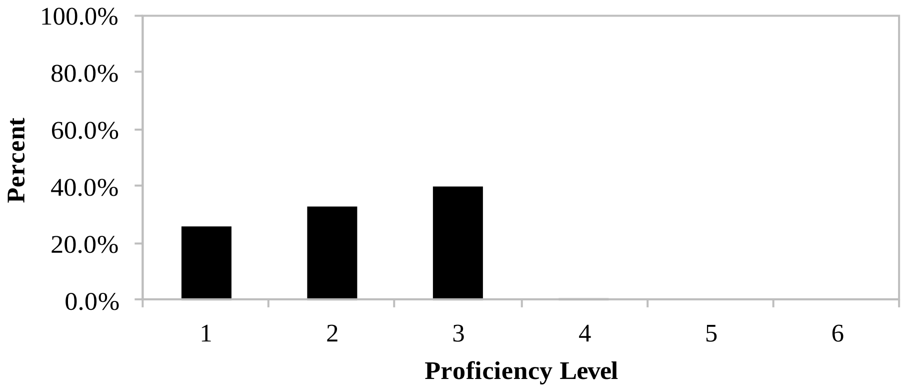


Table 2.5.3.3.2

Proficiency Level Distribution: Writ 4-5 B/C S403 Online

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	963	0.66%	304	0.26%	1,267	0.48%
2	2,486	1.70%	1,522	1.32%	4,008	1.53%
3	71,111	48.52%	47,825	41.51%	118,936	45.43%
4	69,508	47.42%	62,233	54.01%	131,741	50.32%
5	2,377	1.62%	3,248	2.82%	5,625	2.15%
6	130	0.09%	93	0.08%	223	0.09%
Total	146,575	100.00%	115,225	100.00%	261,800	100.00%

Figure 2.5.3.3.2
Proficiency Level: Writ 4-5 B/C S403 Online

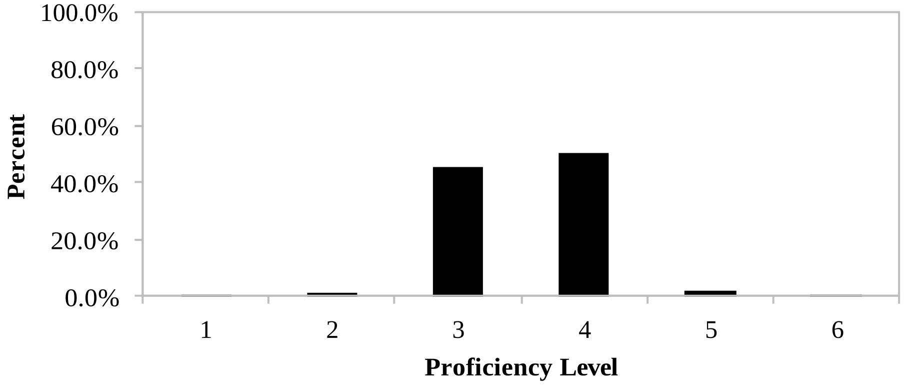
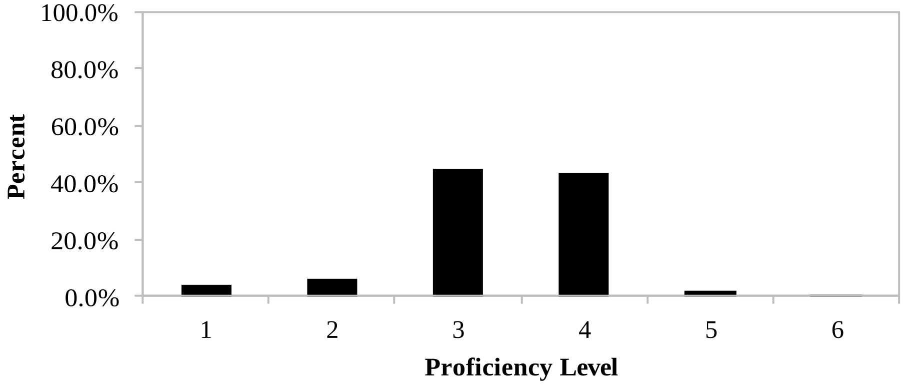


Table 2.5.3.3.3

Proficiency Level Distribution: Writ 4-5 S403 Online

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	7,489	4.44%	5,244	3.82%	12,733	4.16%
2	9,171	5.43%	9,522	6.94%	18,693	6.11%
3	79,945	47.36%	56,748	41.36%	136,693	44.67%
4	69,676	41.28%	62,366	45.45%	132,042	43.15%
5	2,377	1.41%	3,248	2.37%	5,625	1.84%
6	130	0.08%	93	0.07%	223	0.07%
Total	168,788	100.00%	137,221	100.00%	306,009	100.00%

Figure 2.5.3.3.3
Proficiency Level: Writ 4-5 S403 Online



2.5.3.4 *Grades 6–8*

Table 2.5.3.4.1

Proficiency Level Distribution: Writ 6-8 A S403 Online

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	8,641	27.86%	9,881	28.40%	10,974	32.02%	29,496	29.47%
2	12,170	39.23%	14,368	41.29%	10,842	31.64%	37,380	37.35%
3	10,140	32.69%	10,281	29.55%	12,292	35.87%	32,713	32.68%
4	70	0.23%	264	0.76%	163	0.48%	497	0.50%
5	0	0.00%	0	0.00%	1	0.00%	1	0.00%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	31,021	100.00%	34,794	100.00%	34,272	100.00%	100,087	100.00%

Figure 2.5.3.4.1
Proficiency Level: Writ 6-8 A S403 Online

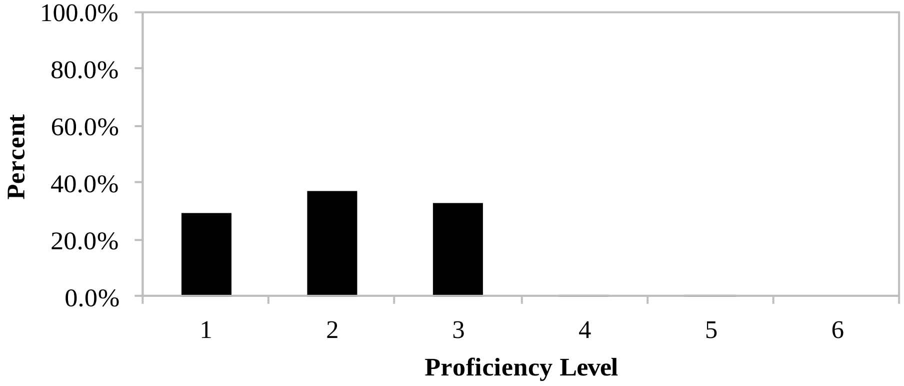


Table 2.5.3.4.2

Proficiency Level Distribution: Writ 6-8 B/C S403 Online

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	336	0.50%	182	0.34%	282	0.59%	800	0.47%
2	3,132	4.66%	2,212	4.13%	1,514	3.17%	6,858	4.07%
3	41,721	62.06%	29,621	55.29%	27,093	56.67%	98,435	58.38%
4	21,900	32.58%	21,387	39.92%	18,671	39.06%	61,958	36.75%
5	132	0.20%	162	0.30%	231	0.48%	525	0.31%
6	6	0.01%	11	0.02%	14	0.03%	31	0.02%
Total	67,227	100.00%	53,575	100.00%	47,805	100.00%	168,607	100.00%

Figure 2.5.3.4.2
Proficiency Level: Writ 6-8 B/C S403 Online

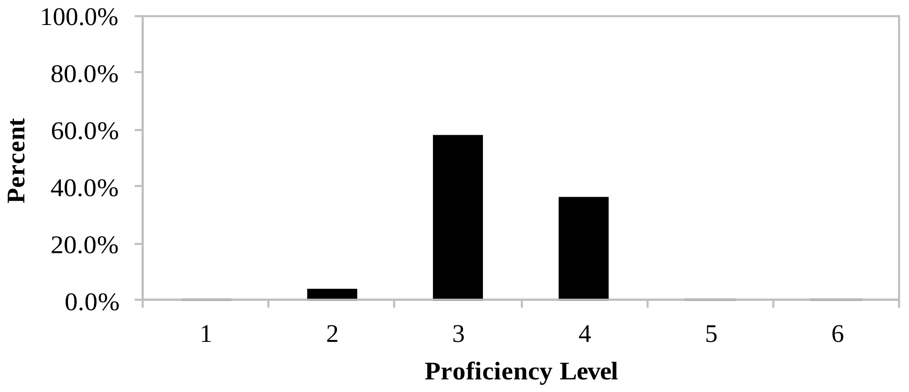
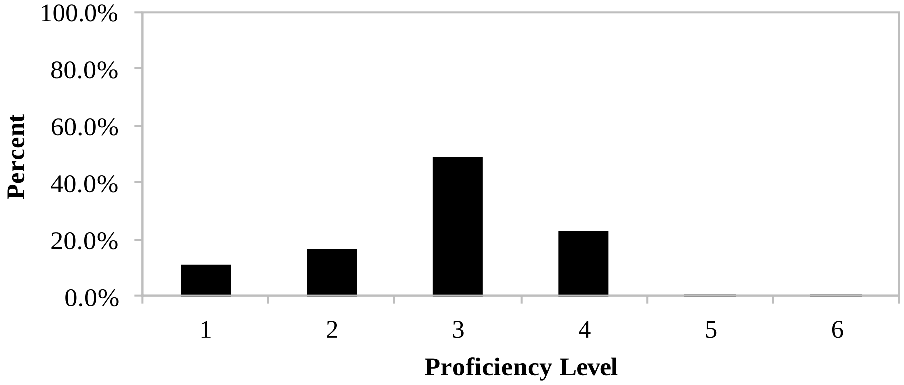


Table 2.5.3.4.3

Proficiency Level Distribution: Writ 6-8 S403 Online

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	8,977	9.14%	10,063	11.39%	11,256	13.71%	30,296	11.28%
2	15,302	15.57%	16,580	18.76%	12,356	15.05%	44,238	16.46%
3	51,861	52.79%	39,902	45.15%	39,385	47.99%	131,148	48.81%
4	21,970	22.36%	21,651	24.50%	18,834	22.95%	62,455	23.24%
5	132	0.13%	162	0.18%	232	0.28%	526	0.20%
6	6	0.01%	11	0.01%	14	0.02%	31	0.01%
Total	98,248	100.00%	88,369	100.00%	82,077	100.00%	268,694	100.00%

Figure 2.5.3.4.3
Proficiency Level: Writ 6-8 S403 Online



2.5.3.5 Grades 9–12

Table 2.5.3.5.1

Proficiency Level Distribution: Writ 9-12 A S403 Online

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	8,617	20.75%	5,218	18.13%	5,037	21.71%	4,090	26.62%	22,962	21.09%
2	10,754	25.89%	7,258	25.22%	5,530	23.83%	4,226	27.51%	27,768	25.51%
3	18,579	44.74%	13,225	45.96%	10,993	47.38%	5,883	38.29%	48,680	44.71%
4	3,558	8.57%	3,045	10.58%	1,638	7.06%	1,163	7.57%	9,404	8.64%
5	23	0.06%	30	0.10%	4	0.02%	1	0.01%	58	0.05%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	41,531	100.00%	28,776	100.00%	23,202	100.00%	15,363	100.00%	108,872	100.00%

Figure 2.5.3.5.1
Proficiency Level: Writ 9-12 AS403 Online

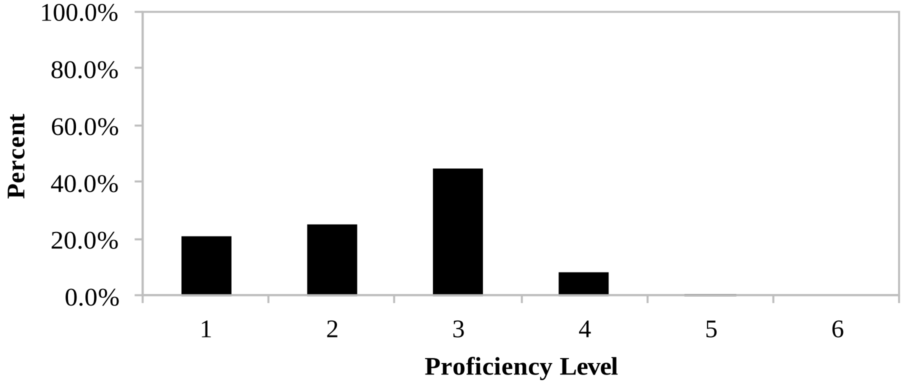


Table 2.5.3.5.2

Proficiency Level Distribution: Writ 9-12 B/C S403 Online

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	605	1.13%	905	1.83%	1,586	3.45%	2,440	6.82%	5,536	3.00%
2	2,866	5.35%	2,785	5.64%	2,797	6.08%	3,166	8.85%	11,614	6.29%
3	24,521	45.77%	24,575	49.75%	22,259	48.38%	17,756	49.64%	89,111	48.23%
4	24,580	45.88%	20,312	41.12%	19,000	41.30%	12,247	34.24%	76,139	41.21%
5	997	1.86%	817	1.65%	367	0.80%	163	0.46%	2,344	1.27%
6	9	0.02%	2	0.00%	0	0.00%	0	0.00%	11	0.01%
Total	53,578	100.00%	49,396	100.00%	46,009	100.00%	35,772	100.00%	184,755	100.00%

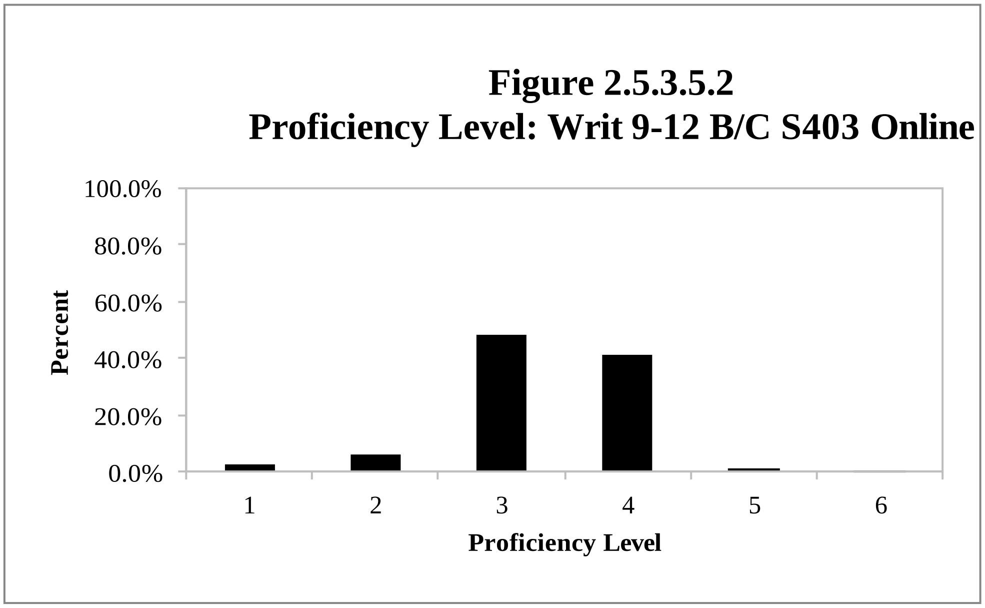


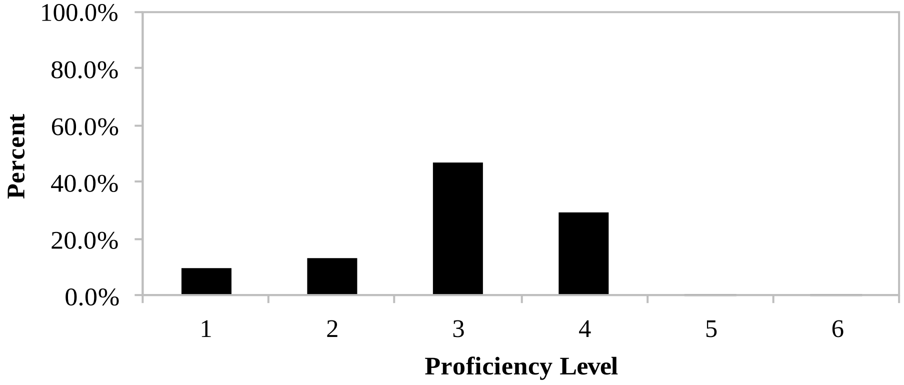
Figure 2.5.3.5.2
Proficiency Level: Writ 9-12 B/C S403 Online

Table 2.5.3.5.3

Proficiency Level Distribution: Writ 9-12 S403 Online

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	9,222	9.70%	6,123	7.83%	6,623	9.57%	6,530	12.77%	28,498	9.71%
2	13,620	14.32%	10,043	12.85%	8,327	12.03%	7,392	14.46%	39,382	13.41%
3	43,100	45.32%	37,800	48.35%	33,252	48.04%	23,639	46.23%	137,791	46.93%
4	28,138	29.59%	23,357	29.88%	20,638	29.82%	13,410	26.22%	85,543	29.13%
5	1,020	1.07%	847	1.08%	371	0.54%	164	0.32%	2,402	0.82%
6	9	0.01%	2	0.00%	0	0.00%	0	0.00%	11	0.00%
Total	95,109	100.00%	78,172	100.00%	69,211	100.00%	51,135	100.00%	293,627	100.00%

Figure 2.5.3.5.3
Proficiency Level: Writ 9-12 S403 Online



2.5.4 Speaking

2.5.4.1 *Grade 1*

Table 2.5.4.1.1

Proficiency Level Distribution: Spek 1 Pre-A S403 Online

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	5,158	100.00%	5,158	100.00%
Total	5,158	100.00%	5,158	100.00%

Figure 2.5.4.1.1

Proficiency Level: Spek 1 Pre-AS403 Online

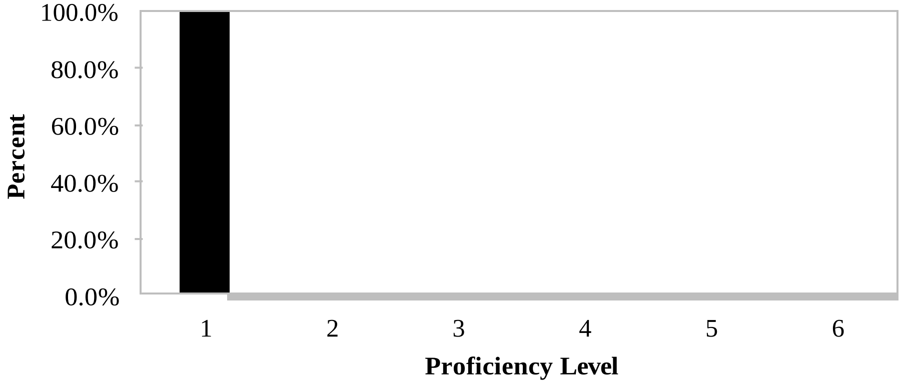


Table 2.5.4.1.2

Proficiency Level Distribution: Spek 1 A S403 Online

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	15,109	24.64%	15,109	24.64%
2	29,243	47.69%	29,243	47.69%
3	13,639	22.24%	13,639	22.24%
4	3,086	5.03%	3,086	5.03%
5	246	0.40%	246	0.40%
6	0	0.00%	0	0.00%
Total	61,323	100.00%	61,323	100.00%

Figure 2.5.4.1.2

Proficiency Level: Spek 1 A S403 Online

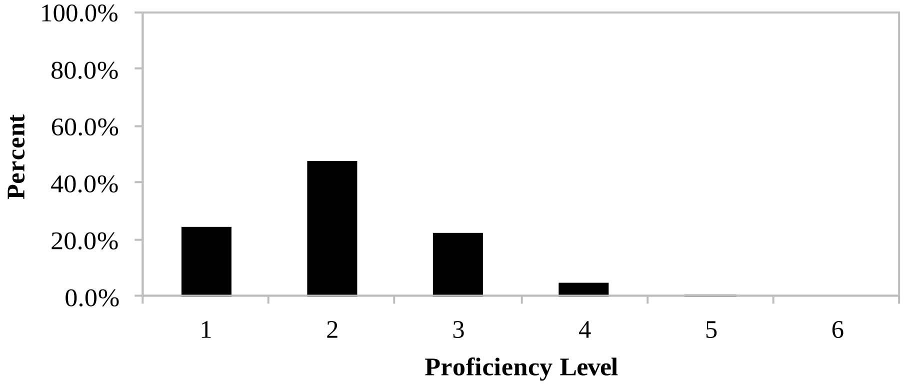


Table 2.5.4.1.3

Proficiency Level Distribution: Spek 1 B/C S403 Online

Level	Grade 1		Total	
	Count	Percent	Count	Percent
1	6,353	5.72%	6,353	5.72%
2	44,277	39.87%	44,277	39.87%
3	45,999	41.42%	45,999	41.42%
4	14,100	12.70%	14,100	12.70%
5	295	0.27%	295	0.27%
6	19	0.02%	19	0.02%
Total	111,043	100.00%	111,043	100.00%

Figure 2.5.4.1.3
Proficiency Level: Spek 1 B/C S403 Online

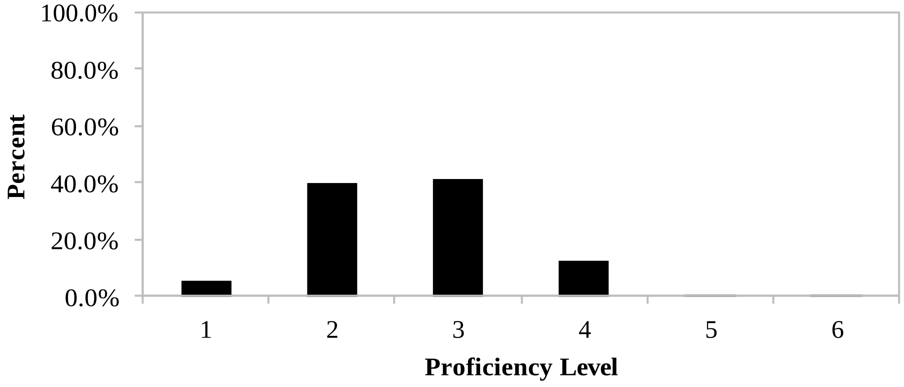
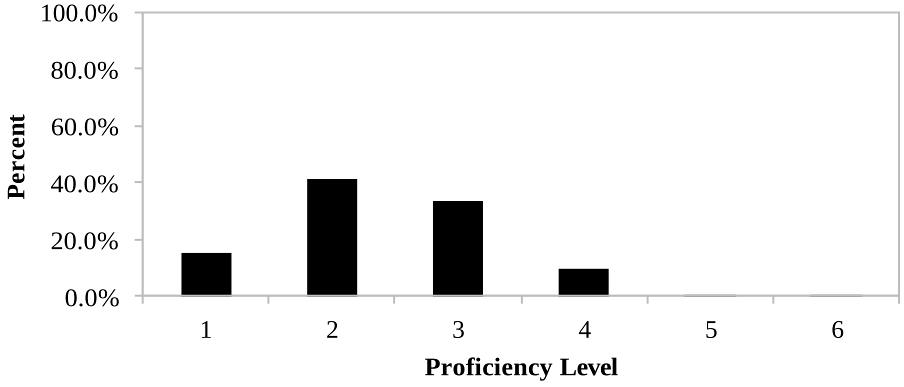


Table 2.5.4.1.4

Proficiency Level Distribution: Spek 1 S403 Online

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	26,620	15.00%	26,620	15.00%
2	73,520	41.41%	73,520	41.41%
3	59,638	33.59%	59,638	33.59%
4	17,186	9.68%	17,186	9.68%
5	541	0.30%	541	0.30%
6	19	0.01%	19	0.01%
Total	177,524	100.00%	177,524	100.00%

Figure 2.5.4.1.4
Proficiency Level: Spek 1 S403 Online



2.5.4.2 *Grades 2–3*

Table 2.5.4.2.1

Proficiency Level Distribution: Spek 2-3 Pre-A S403 Online

	Grade 2		**Grade 3**		**Total**	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	6,879	100.00%	9,480	100.00%	16,359	100.00%
Total	6,879	100.00%	9,480	100.00%	16,359	100.00%

Figure 2.5.4.2.1
Proficiency Level: Spek 2-3 Pre-AS403 Online

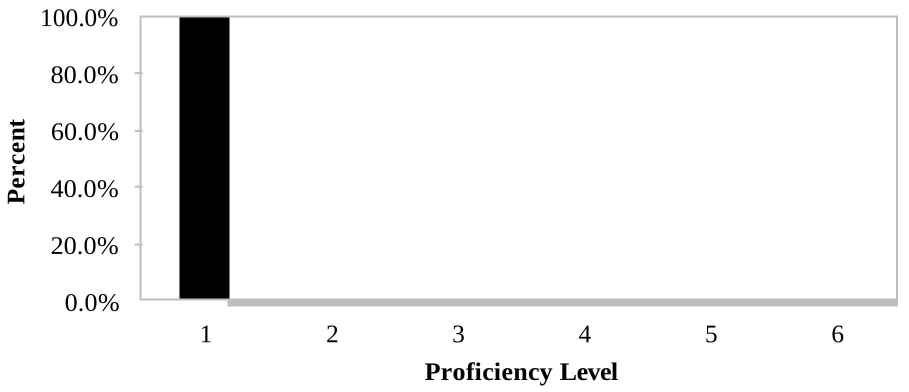


Table 2.5.4.2.2

Proficiency Level Distribution: Spek 2-3 A S403 Online

	Grade 2		Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	11,774	28.31%	11,071	29.58%	22,845	28.91%
2	18,299	44.00%	19,046	50.89%	37,345	47.26%
3	11,107	26.71%	6,444	17.22%	17,551	22.21%
4	392	0.94%	868	2.32%	1,260	1.59%
5	12	0.03%	0	0.00%	12	0.02%
6	0	0.00%	0	0.00%	0	0.00%
Total	41,584	100.00%	37,429	100.00%	79,013	100.00%

Figure 2.5.4.2.2
Proficiency Level: Spek 2-3 A S403 Online

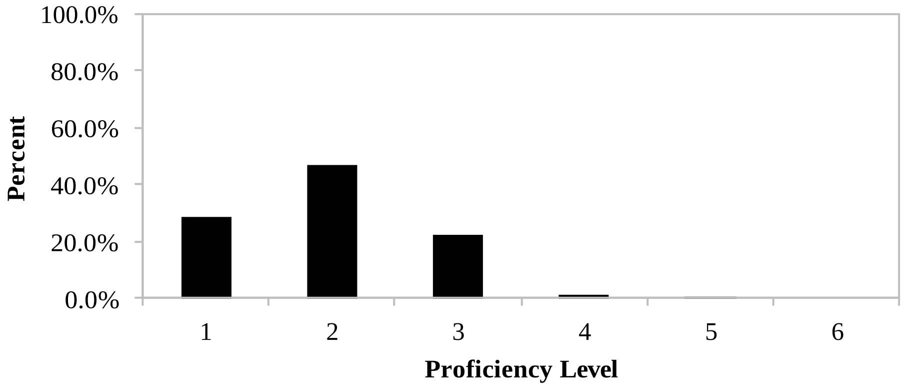


Table 2.5.4.2.3

Proficiency Level Distribution: Spek 2-3 B/C S403 Online

	Grade 2		Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	8,310	6.21%	10,558	7.72%	18,868	6.98%
2	61,888	46.26%	59,228	43.32%	121,116	44.78%
3	53,268	39.82%	56,541	41.36%	109,809	40.60%
4	9,906	7.40%	10,055	7.35%	19,961	7.38%
5	383	0.29%	250	0.18%	633	0.23%
6	30	0.02%	78	0.06%	108	0.04%
Total	133,785	100.00%	136,710	100.00%	270,495	100.00%

Figure 2.5.4.2.3
Proficiency Level: Spek 2-3 B/C S403 Online

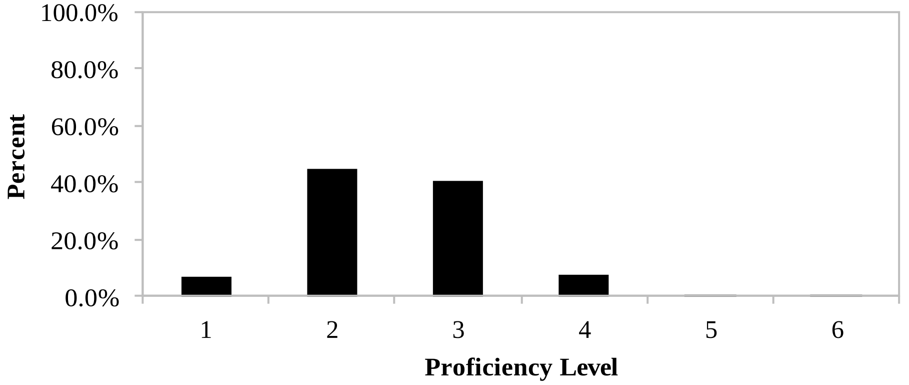
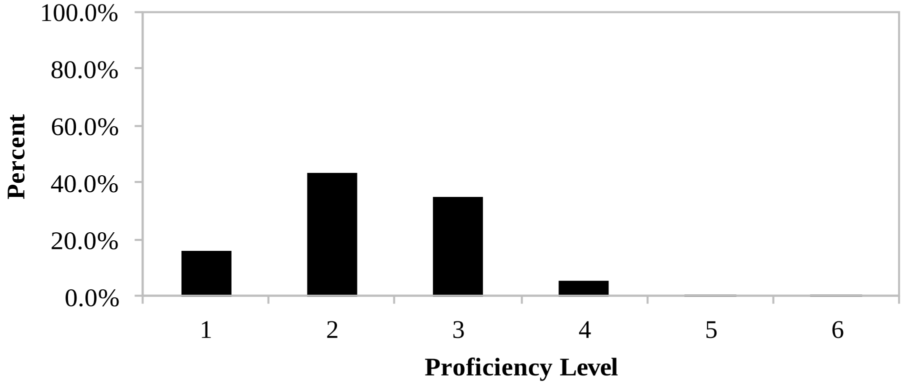


Table 2.5.4.2.4

Proficiency Level Distribution: Spek 2-3 S403 Online

Level	Grade 2		Grade 3		Total	
	Count	Percent	Count	Percent	Count	Percent
1	26,963	14.79%	31,109	16.94%	58,072	15.87%
2	80,187	44.00%	78,274	42.63%	158,461	43.31%
3	64,375	35.32%	62,985	34.30%	127,360	34.81%
4	10,298	5.65%	10,923	5.95%	21,221	5.80%
5	395	0.22%	250	0.14%	645	0.18%
6	30	0.02%	78	0.04%	108	0.03%
Total	182,248	100.00%	183,619	100.00%	365,867	100.00%

Figure 2.5.4.2.4
Proficiency Level: Spek 2-3 S403 Online



2.5.4.3 *Grades 4–5*

Table 2.5.4.3.1

Proficiency Level Distribution: Spek 4-5 Pre-A S403 Online

	Grade 4		**Grade 5**		**Total**	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,390	100.00%	3,621	100.00%	6,011	100.00%
Total	2,390	100.00%	3,621	100.00%	6,011	100.00%

Figure 2.5.4.3.1
Proficiency Level: Spek 4-5 Pre-AS403 Online

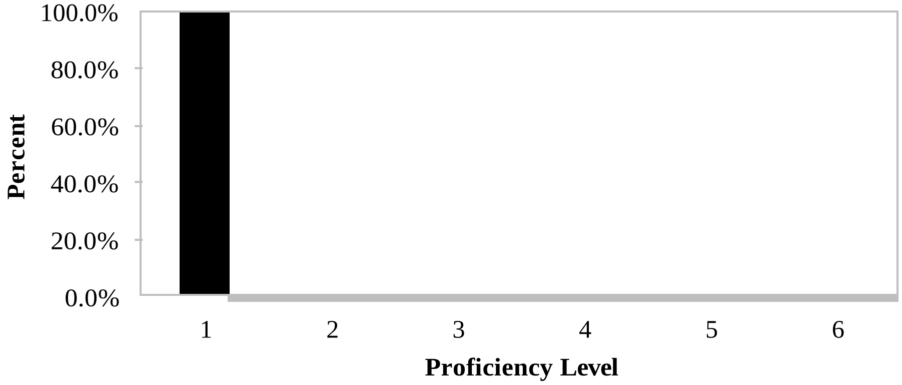


Table 2.5.4.3.2

Proficiency Level Distribution: Spek 4-5 A S403 Online

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	4,802	30.54%	4,989	40.64%	9,791	34.97%
2	5,213	33.15%	5,257	42.83%	10,470	37.39%
3	5,062	32.19%	1,857	15.13%	6,919	24.71%
4	644	4.10%	171	1.39%	815	2.91%
5	5	0.03%	1	0.01%	6	0.02%
6	0	0.00%	0	0.00%	0	0.00%
Total	15,726	100.00%	12,275	100.00%	28,001	100.00%

Figure 2.5.4.3.2
Proficiency Level: Spek 4-5 A S403 Online

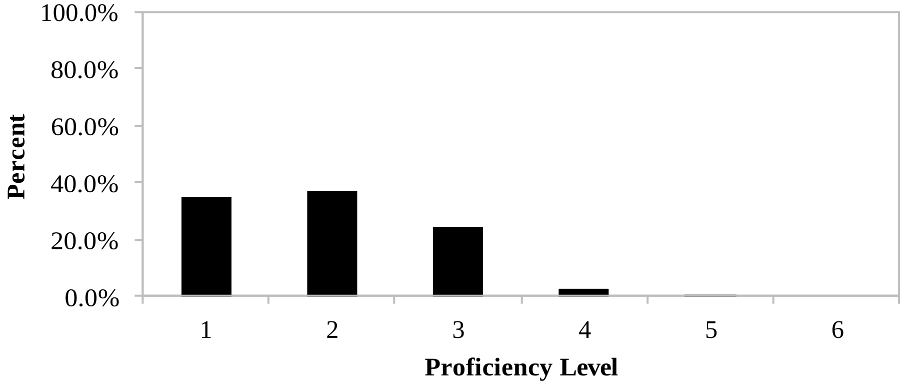


Table 2.5.4.3.3

Proficiency Level Distribution: Spek 4-5 B/C S403 Online

Level	Grade 4		Grade 5		Total	
	Count	Percent	Count	Percent	Count	Percent
1	5,366	3.45%	6,663	5.37%	12,029	4.30%
2	31,210	20.05%	31,192	25.15%	62,402	22.31%
3	75,006	48.19%	58,877	47.48%	133,883	47.88%
4	41,615	26.74%	26,076	21.03%	67,691	24.21%
5	2,336	1.50%	1,159	0.93%	3,495	1.25%
6	113	0.07%	33	0.03%	146	0.05%
Total	155,646	100.00%	124,000	100.00%	279,646	100.00%

Figure 2.5.4.3.3
Proficiency Level: Spek 4-5 B/C S403 Online

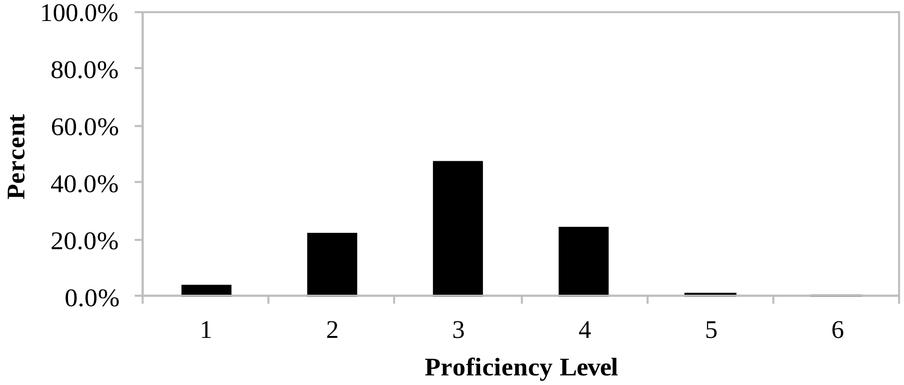


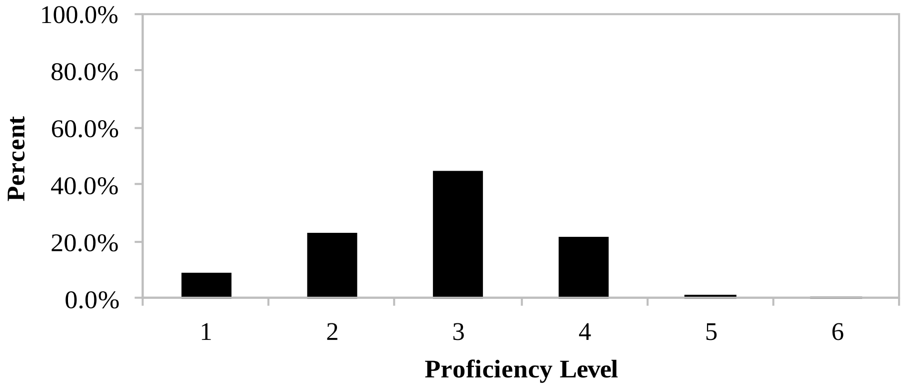
Table 2.5.4.3.4

Proficiency Level Distribution: Spek 4-5 S403 Online

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	12,558	7.23%	15,273	10.92%	27,831	8.87%
2	36,423	20.96%	36,449	26.05%	72,872	23.23%
3	80,068	46.08%	60,734	43.41%	140,802	44.89%
4	42,259	24.32%	26,247	18.76%	68,506	21.84%
5	2,341	1.35%	1,160	0.83%	3,501	1.12%
6	113	0.07%	33	0.02%	146	0.05%
Total	173,762	100.00%	139,896	100.00%	313,658	100.00%

Figure 2.5.4.3.4

Proficiency Level: Spek 4-5 S403 Online



2.5.4.4 *Grades 6–8*

Table 2.5.4.4.1

Proficiency Level Distribution: Spek 6-8 Pre-A S403 Online

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	2,811	100.00%	3,854	100.00%	4,829	100.00%	11,494	100.00%
Total	2,811	100.00%	3,854	100.00%	4,829	100.00%	11,494	100.00%

Figure 2.5.4.4.1
Proficiency Level: Spek 6-8 Pre-AS403 Online

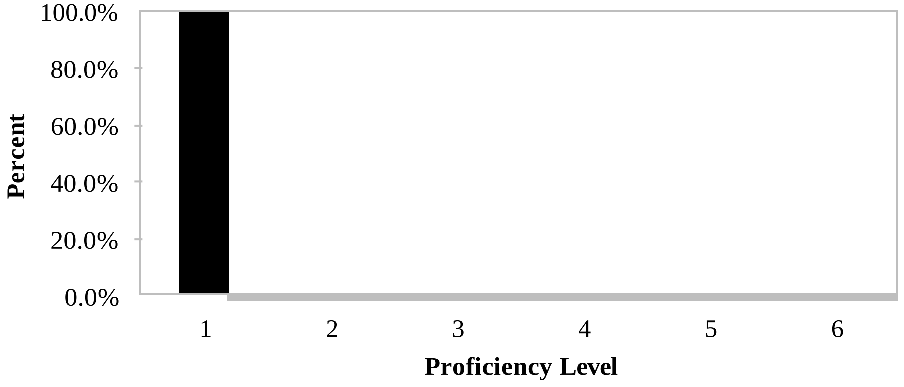


Table 2.5.4.4.2

Proficiency Level Distribution: Spek 6-8 A S403 Online

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	7,007	39.17%	6,570	43.70%	11,802	47.38%	25,379	43.89%
2	7,499	41.92%	5,971	39.72%	6,618	26.57%	20,088	34.74%
3	2,974	16.63%	2,179	14.49%	5,408	21.71%	10,561	18.26%
4	397	2.22%	310	2.06%	1,076	4.32%	1,783	3.08%
5	10	0.06%	4	0.03%	4	0.02%	18	0.03%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	17,887	100.00%	15,034	100.00%	24,908	100.00%	57,829	100.00%

Figure 2.5.4.4.2
Proficiency Level: Spek 6-8 A S403 Online

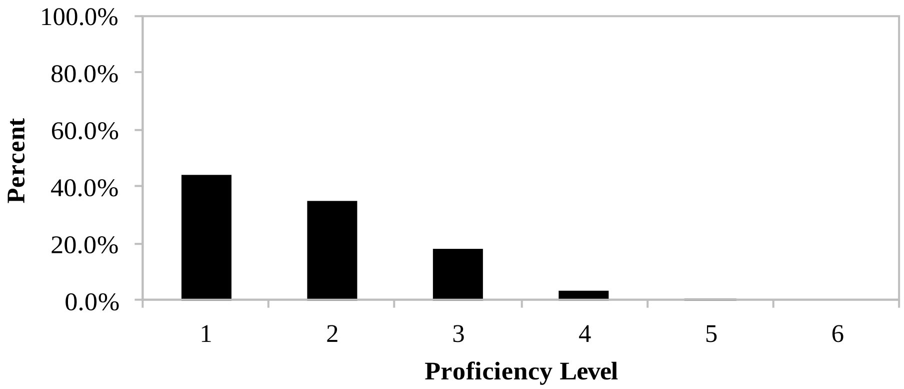


Table 2.5.4.4.3

Proficiency Level Distribution: Spek 6-8 B/C S403 Online

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	3,990	4.99%	5,478	7.69%	4,331	7.98%	13,799	6.71%
2	19,381	24.21%	22,022	30.92%	11,520	21.23%	52,923	25.75%
3	40,707	50.86%	32,833	46.10%	30,886	56.92%	104,426	50.81%
4	15,742	19.67%	10,544	14.81%	7,268	13.39%	33,554	16.33%
5	191	0.24%	331	0.46%	234	0.43%	756	0.37%
6	27	0.03%	8	0.01%	23	0.04%	58	0.03%
Total	80,038	100.00%	71,216	100.00%	54,262	100.00%	205,516	100.00%

Figure 2.5.4.4.3
Proficiency Level: Spek 6-8 B/C S403 Online

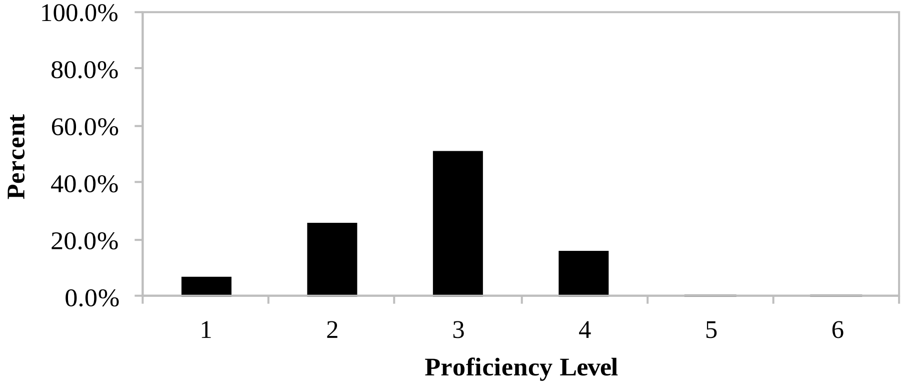
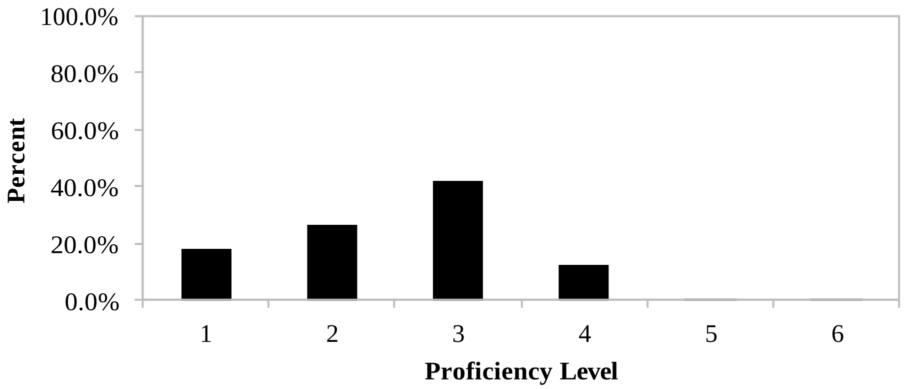


Table 2.5.4.4.4

Proficiency Level Distribution: Spek 6-8 S403 Online

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	13,808	13.71%	15,902	17.65%	20,962	24.96%	50,672	18.44%
2	26,880	26.68%	27,993	31.07%	18,138	21.59%	73,011	26.57%
3	43,681	43.36%	35,012	38.86%	36,294	43.21%	114,987	41.84%
4	16,139	16.02%	10,854	12.05%	8,344	9.93%	35,337	12.86%
5	201	0.20%	335	0.37%	238	0.28%	774	0.28%
6	27	0.03%	8	0.01%	23	0.03%	58	0.02%
Total	100,736	100.00%	90,104	100.00%	83,999	100.00%	274,839	100.00%

Figure 2.5.4.4.4
Proficiency Level: Spek 6-8 S403 Online



2.5.4.5 *Grades 9–12*

Table 2.5.4.5.1

Proficiency Level Distribution: Spek 9-12 Pre-A S403 Online

	Grade 9		**Grade 10**		**Grade 11**		**Grade 12**		**Total**	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	5,440	100.00%	6,455	100.00%	6,209	100.00%	4,998	100.00%	23,102	100.00%
Total	5,440	100.00%	6,455	100.00%	6,209	100.00%	4,998	100.00%	23,102	100.00%

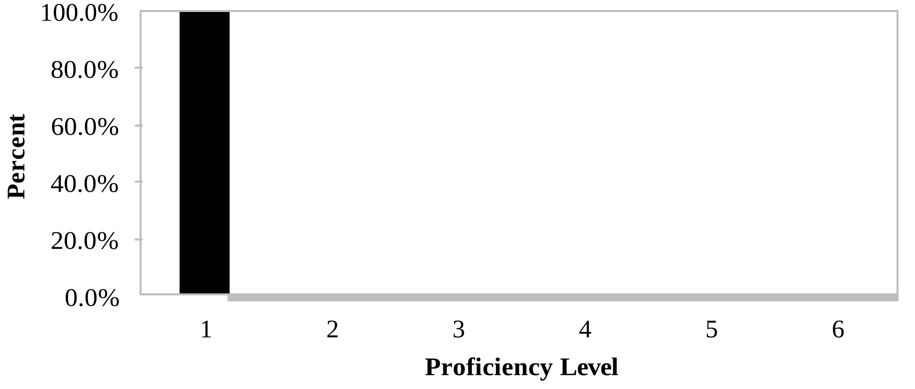


Figure 2.5.4.5.1
Proficiency Level: Spek 9-12 Pre-AS403 Online

Table 2.5.4.5.2

Proficiency Level Distribution: Spek 9-12 A S403 Online

Level	Grade 9		Grade 10		Grade 11		Grade 12		Total	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	21,729	39.88%	10,723	33.68%	7,924	56.43%	8,920	38.25%	49,296	39.85%
2	24,624	45.19%	15,861	49.81%	4,199	29.91%	7,929	34.00%	52,613	42.54%
3	7,564	13.88%	5,184	16.28%	1,904	13.56%	6,350	27.23%	21,002	16.98%
4	574	1.05%	72	0.23%	14	0.10%	117	0.50%	777	0.63%
5	1	0.00%	0	0.00%	0	0.00%	2	0.01%	3	0.00%
6	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Total	54,492	100.00%	31,840	100.00%	14,041	100.00%	23,318	100.00%	123,691	100.00%

Figure 2.5.4.5.2
Proficiency Level: Spek 9-12 AS403 Online

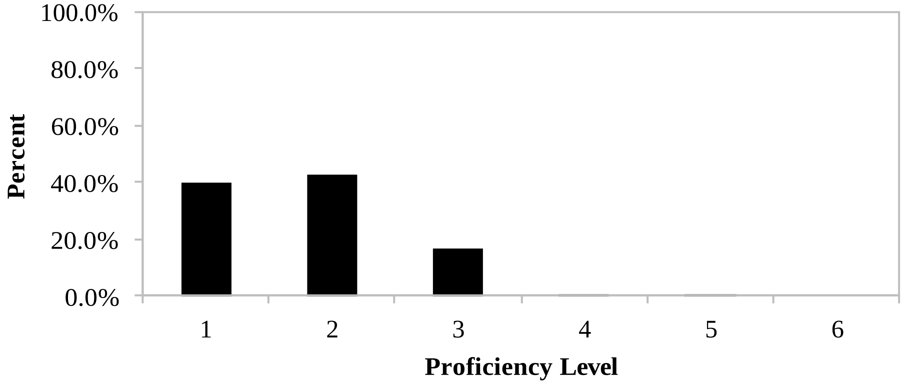


Table 2.5.4.5.3

Proficiency Level Distribution: Spek 9-12 B/C S403 Online

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	3,074	8.28%	5,586	13.46%	7,927	15.87%	2,833	11.84%	19,420	12.73%
2	8,496	22.88%	14,091	33.95%	16,676	33.39%	6,455	26.98%	45,718	29.98%
3	22,978	61.87%	20,223	48.72%	24,381	48.81%	13,973	58.41%	81,555	53.47%
4	2,554	6.88%	1,589	3.83%	941	1.88%	642	2.68%	5,726	3.75%
5	36	0.10%	16	0.04%	19	0.04%	12	0.05%	83	0.05%
6	1	0.00%	2	0.00%	4	0.01%	7	0.03%	14	0.01%
Total	37,139	100.00%	41,507	100.00%	49,948	100.00%	23,922	100.00%	152,516	100.00%

Figure 2.5.4.5.3
Proficiency Level: Spek 9-12 B/C S403 Online

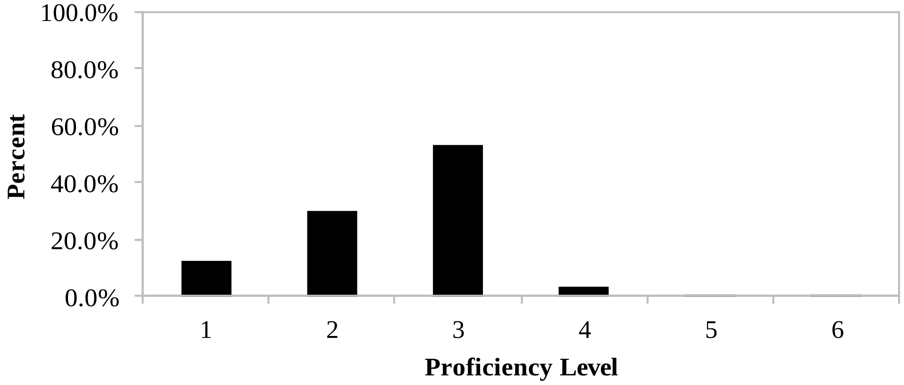
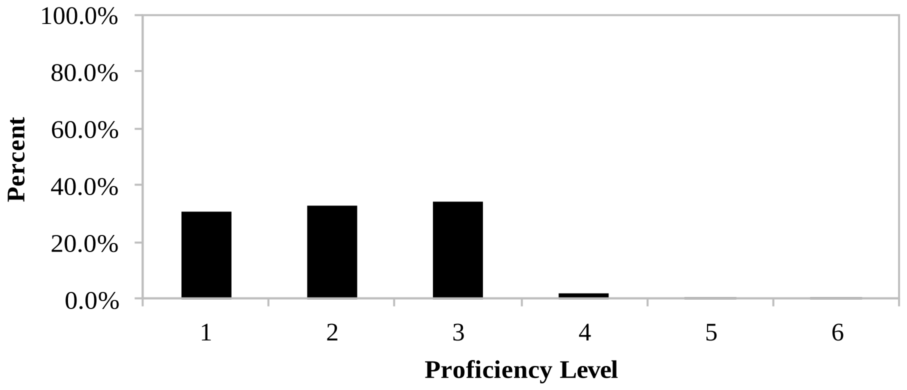


Table 2.5.4.5.4

Proficiency Level Distribution: Spek 9-12 S403 Online

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	30,243	31.16%	22,764	28.53%	22,060	31.43%	16,751	32.07%	91,818	30.68%
2	33,120	34.12%	29,952	37.53%	20,875	29.74%	14,384	27.54%	98,331	32.85%
3	30,542	31.46%	25,407	31.84%	26,285	37.44%	20,323	38.90%	102,557	34.26%
4	3,128	3.22%	1,661	2.08%	955	1.36%	759	1.45%	6,503	2.17%
5	37	0.04%	16	0.02%	19	0.03%	14	0.03%	86	0.03%
6	1	0.00%	2	0.00%	4	0.01%	7	0.01%	14	0.00%
Total	97,071	100.00%	79,802	100.00%	70,198	100.00%	52,238	100.00%	299,309	100.00%

Figure 2.5.4.5.4
Proficiency Level: Spek 9-12 S403 Online



2.6 Raw Score to Scale Score to Proficiency Level Conversion for Speaking and Writing

This section presents raw score to scale score conversions and associated proficiency levels for the test forms for Speaking and Writing.

The first column shows all possible raw scores. The following column shows the corresponding scale score. The next column shows the conditional standard error of measurement (CSEM) in the metric of the scale score, multiplied by 1.96. This is the confidence band as reported on students' score reports. Following the CSEM, columns provide the proficiency level interpretation for each grade in the grade-level cluster.

Note that raw scores of zero are truncated where necessary so that the lowest scale score given is the scale score corresponding to a proficiency level score of 1.0.

2.6.1 Listening

The ACCESS Online Listening test is a multistage adaptive assessment. As students do not all take the same set of items in the test, raw to scale score conversion tables are not presented.

2.6.2 Reading

The ACCESS Online Reading test is a multistage adaptive assessment. As students do not all take the same set of items in the test, raw to scale score conversion tables are not presented.

2.6.3 Writing

Table 2.6.3.1.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 1 A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	111	84	1.0
1	132	76	1.1
2	173	52	1.4
3	191	36	1.6
4	202	28	1.7
5	209	24	1.7
6	214	22	1.8
7	219	20	1.8
8	223	20	1.8
9	227	19	1.9
10	230	19	1.9
11	234	19	1.9
12	238	20	2.0
13	242	20	2.1
14	246	21	2.2
15	251	23	2.3
16	257	24	2.5
17	263	26	2.6
18	270	27	2.8
19	277	28	3.0
20	285	28	3.1
21	293	28	3.2
22	301	27	3.4
23	308	27	3.5
24	315	26	3.6
25	322	26	3.7
26	329	26	3.8
27	335	26	3.9
28	342	26	4.1
29	349	26	4.2
30	355	25	4.4
31	362	25	4.5
32	368	25	4.6
33	374	25	4.8
34	380	25	4.9
35	387	26	5.2
36	394	28	5.5
37	403	31	5.9
38	414	37	6.0
39	432	51	6.0
40	464	95	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.3.1.2

Raw Score to Scale Score to Proficiency Level Conversion: Writ 1 B/C S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	111	188	1.0
1	174	47	1.4
2	189	32	1.6
3	197	26	1.6
4	203	23	1.7
5	207	21	1.7
6	211	19	1.7
7	215	18	1.8
8	218	18	1.8
9	221	17	1.8
10	224	17	1.8
11	227	17	1.9
12	230	17	1.9
13	233	17	1.9
14	236	17	1.9
15	239	17	2.0
16	242	17	2.1
17	245	17	2.1
18	248	18	2.2
19	251	18	2.3
20	254	19	2.4
21	258	19	2.5
22	262	20	2.6
23	267	21	2.7
24	271	22	2.8
25	276	22	3.0
26	281	23	3.0
27	287	23	3.1
28	292	24	3.2
29	298	24	3.3
30	304	24	3.4
31	310	24	3.5
32	316	24	3.6
33	322	24	3.7
34	327	24	3.8
35	333	23	3.9
36	338	23	4.0
37	344	22	4.1
38	349	22	4.2
39	354	22	4.3
40	358	21	4.4
41	363	21	4.5
42	367	21	4.6
43	372	21	4.7
44	376	21	4.8
45	381	21	4.9
46	386	22	5.1
47	390	22	5.3
48	396	23	5.6
49	401	25	5.8
50	408	27	6.0
51	416	30	6.0
52	427	36	6.0
53	445	51	6.0
54	477	95	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.3.2.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 2-3 A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G2	PL for G3
0	133	217	1.0	1.0
1	201	45	1.6	1.5
2	214	32	1.7	1.7
3	223	26	1.8	1.7
4	229	24	1.8	1.8
5	234	23	1.9	1.8
6	239	22	1.9	1.9
7	244	23	2.0	1.9
8	250	23	2.2	2.0
9	256	25	2.3	2.2
10	263	27	2.5	2.4
11	270	29	2.7	2.6
12	280	31	3.0	2.9
13	290	33	3.1	3.1
14	301	34	3.3	3.2
15	313	35	3.5	3.4
16	325	34	3.7	3.6
17	337	34	3.9	3.8
18	348	33	4.1	4.0
19	358	31	4.3	4.2
20	368	30	4.5	4.4
21	377	29	4.7	4.6
22	385	29	4.9	4.8
23	394	30	5.2	5.0
24	404	32	5.6	5.4
25	416	37	6.0	5.9
26	434	51	6.0	6.0
27	466	94	6.0	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.3.2.2

Raw Score to Scale Score to Proficiency Level Conversion: Writ 2-3 B/C S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G2	PL for G3	Raw Score	Scale Score	CSEM x 1.96	PL for G2	PL for G3
0	133	223	1.0	1.0	34	356	24	4.3	4.2
1	205	47	1.6	1.6	35	362	23	4.4	4.3
2	220	32	1.8	1.7	36	367	23	4.5	4.4
3	228	26	1.8	1.8	37	373	23	4.6	4.5
4	234	22	1.9	1.8	38	378	22	4.7	4.6
5	238	20	1.9	1.9	39	382	22	4.8	4.7
6	242	18	2.0	1.9	40	387	21	4.9	4.8
7	245	17	2.0	1.9	41	392	21	5.1	4.9
8	248	17	2.1	2.0	42	396	21	5.3	5.0
9	251	16	2.2	2.1	43	400	21	5.5	5.2
10	254	16	2.3	2.1	44	405	21	5.7	5.4
11	256	16	2.3	2.2	45	409	21	5.9	5.6
12	259	16	2.4	2.3	46	413	21	6.0	5.7
13	261	16	2.5	2.3	47	418	22	6.0	6.0
14	264	16	2.5	2.4	48	423	23	6.0	6.0
15	266	16	2.6	2.5	49	429	24	6.0	6.0
16	269	16	2.7	2.6	50	435	26	6.0	6.0
17	272	17	2.8	2.6	51	443	30	6.0	6.0
18	275	17	2.8	2.7	52	453	36	6.0	6.0
19	278	18	2.9	2.8	53	471	51	6.0	6.0
20	282	19	3.0	2.9	54	503	95	6.0	6.0
21	286	20	3.1	3.0					
22	290	21	3.1	3.1					
23	294	21	3.2	3.1					
24	299	22	3.3	3.2					
25	304	23	3.4	3.3					
26	309	23	3.4	3.4					
27	315	24	3.5	3.5					
28	321	24	3.6	3.6					
29	327	24	3.7	3.6					
30	333	24	3.8	3.7					
31	338	24	3.9	3.8					
32	344	24	4.0	3.9					
33	350	24	4.1	4.0					

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.3.3.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 4-5 A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
0	155	243	1.0	1.0
1	229	45	1.6	1.6
2	242	32	1.8	1.7
3	251	26	1.8	1.8
4	257	24	1.9	1.9
5	262	23	1.9	1.9
6	268	22	2.0	2.0
7	273	23	2.3	2.2
8	278	23	2.5	2.4
9	284	25	2.8	2.6
10	291	27	3.0	2.9
11	299	29	3.1	3.0
12	308	31	3.3	3.2
13	318	33	3.4	3.3
14	330	34	3.6	3.5
15	341	35	3.8	3.7
16	353	34	4.0	3.9
17	365	34	4.2	4.1
18	376	33	4.5	4.3
19	386	31	4.7	4.5
20	396	30	4.9	4.7
21	405	29	5.1	4.9
22	414	29	5.5	5.2
23	422	30	5.8	5.5
24	432	32	6.0	5.9
25	444	37	6.0	6.0
26	463	51	6.0	6.0
27	494	94	6.0	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.3.3.2

Raw Score to Scale Score to Proficiency Level Conversion: Writ 4-5 B/C S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5	Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
0	155	256	1.0	1.0	34	386	24	4.7	4.5
1	234	47	1.7	1.7	35	392	23	4.8	4.7
2	249	32	1.8	1.8	36	397	23	4.9	4.8
3	257	26	1.9	1.9	37	403	23	5.0	4.9
4	263	23	1.9	1.9	38	408	22	5.2	5.0
5	267	20	2.0	2.0	39	413	22	5.5	5.2
6	271	19	2.2	2.1	40	417	21	5.6	5.3
7	275	18	2.4	2.3	41	422	21	5.8	5.5
8	278	17	2.5	2.4	42	426	21	6.0	5.7
9	281	17	2.6	2.5	43	431	21	6.0	5.9
10	284	16	2.8	2.6	44	435	21	6.0	6.0
11	287	16	2.9	2.7	45	440	21	6.0	6.0
12	289	16	3.0	2.8	46	444	21	6.0	6.0
13	292	16	3.0	2.9	47	449	22	6.0	6.0
14	294	16	3.0	3.0	48	454	23	6.0	6.0
15	297	16	3.1	3.0	49	460	24	6.0	6.0
16	300	17	3.1	3.1	50	466	26	6.0	6.0
17	303	17	3.2	3.1	51	474	30	6.0	6.0
18	306	17	3.2	3.2	52	485	36	6.0	6.0
19	309	18	3.3	3.2	53	503	51	6.0	6.0
20	313	19	3.3	3.3	54	534	95	6.0	6.0
21	316	20	3.4	3.3					
22	320	20	3.5	3.4					
23	325	21	3.5	3.5					
24	330	22	3.6	3.5					
25	335	23	3.7	3.6					
26	340	23	3.8	3.7					
27	345	24	3.9	3.8					
28	351	24	4.0	3.9					
29	357	24	4.1	4.0					
30	363	24	4.2	4.1					
31	369	24	4.3	4.2					
32	375	24	4.4	4.3					
33	381	24	4.6	4.4					

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.3.4.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 6-8 A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	188	101	1.2	1.1	1.0
1	219	45	1.5	1.4	1.3
2	233	32	1.6	1.5	1.4
3	242	27	1.7	1.6	1.5
4	249	25	1.8	1.7	1.6
5	255	24	1.8	1.8	1.7
6	261	23	1.9	1.8	1.7
7	266	24	1.9	1.9	1.8
8	272	24	2.1	1.9	1.9
9	278	25	2.3	2.1	1.9
10	285	27	2.5	2.3	2.1
11	293	29	2.8	2.6	2.4
12	302	31	3.0	2.9	2.7
13	312	33	3.2	3.1	3.0
14	323	34	3.3	3.2	3.1
15	335	34	3.5	3.4	3.3
16	346	34	3.7	3.6	3.5
17	358	34	3.9	3.8	3.7
18	369	33	4.1	4.0	3.9
19	379	31	4.3	4.2	4.1
20	389	30	4.5	4.4	4.3
21	398	30	4.7	4.5	4.5
22	407	30	4.8	4.7	4.6
23	416	31	5.1	4.9	4.8
24	426	33	5.4	5.2	5.0
25	438	38	5.8	5.6	5.4
26	457	52	6.0	6.0	5.9
27	489	94	6.0	6.0	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.3.4.2

Raw Score to Scale Score to Proficiency Level Conversion: Writ 6-8 B/C S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	188	102	1.2	1.1	1.0
1	221	47	1.5	1.4	1.3
2	236	32	1.6	1.6	1.5
3	244	26	1.7	1.6	1.6
4	250	22	1.8	1.7	1.6
5	254	20	1.8	1.8	1.7
6	258	19	1.9	1.8	1.7
7	262	18	1.9	1.8	1.7
8	265	17	1.9	1.9	1.8
9	267	16	1.9	1.9	1.8
10	270	16	2.0	1.9	1.8
11	273	16	2.1	2.0	1.9
12	275	16	2.2	2.0	1.9
13	278	16	2.3	2.1	1.9
14	280	16	2.4	2.2	1.9
15	283	16	2.5	2.3	2.0
16	286	16	2.6	2.4	2.1
17	288	17	2.6	2.4	2.2
18	292	17	2.8	2.5	2.3
19	295	18	2.9	2.6	2.4
20	298	19	3.0	2.7	2.5
21	302	20	3.0	2.9	2.7
22	306	21	3.1	3.0	2.8
23	311	21	3.2	3.0	3.0
24	316	22	3.2	3.1	3.0
25	321	23	3.3	3.2	3.1
26	326	23	3.4	3.3	3.2
27	332	24	3.5	3.4	3.3
28	337	24	3.6	3.5	3.4
29	343	24	3.7	3.6	3.5
30	349	24	3.8	3.7	3.6
31	355	24	3.9	3.8	3.7
32	361	24	4.0	3.9	3.8
33	367	24	4.1	4.0	3.9

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
34	373	24	4.2	4.1	4.0
35	378	23	4.3	4.2	4.1
36	384	23	4.4	4.3	4.2
37	389	23	4.5	4.4	4.3
38	394	22	4.6	4.5	4.4
39	399	22	4.7	4.6	4.5
40	404	21	4.8	4.7	4.6
41	408	21	4.9	4.7	4.6
42	413	21	5.0	4.8	4.7
43	417	21	5.1	4.9	4.8
44	421	21	5.2	5.0	4.9
45	426	21	5.4	5.2	5.0
46	430	21	5.6	5.3	5.1
47	435	22	5.7	5.5	5.3
48	440	23	5.9	5.6	5.4
49	445	24	6.0	5.8	5.6
50	452	26	6.0	6.0	5.8
51	460	30	6.0	6.0	6.0
52	470	36	6.0	6.0	6.0
53	488	51	6.0	6.0	6.0
54	520	95	6.0	6.0	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.3.5.1

Raw Score to Scale Score to Proficiency Level Conversion: Writ 9-12 A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
0	232	72	1.3	1.2	1.1	1.0
1	250	45	1.5	1.4	1.3	1.2
2	264	32	1.7	1.6	1.4	1.3
3	272	27	1.8	1.7	1.5	1.4
4	279	24	1.8	1.7	1.6	1.5
5	284	23	1.9	1.8	1.7	1.6
6	290	23	2.0	1.9	1.7	1.6
7	295	23	2.2	1.9	1.8	1.7
8	300	24	2.3	2.0	1.9	1.7
9	306	25	2.5	2.2	1.9	1.8
10	313	27	2.8	2.5	2.1	1.9
11	321	29	3.0	2.8	2.4	2.1
12	330	31	3.1	3.0	2.8	2.4
13	340	33	3.3	3.2	3.0	2.8
14	352	34	3.5	3.4	3.3	3.1
15	363	34	3.7	3.6	3.5	3.3
16	375	34	3.9	3.8	3.7	3.5
17	387	34	4.1	4.0	3.9	3.7
18	398	33	4.3	4.2	4.1	4.0
19	408	31	4.5	4.4	4.3	4.2
20	418	30	4.7	4.6	4.5	4.4
21	427	30	4.9	4.8	4.7	4.5
22	436	30	5.1	5.0	4.9	4.7
23	445	30	5.3	5.2	5.0	4.9
24	454	33	5.6	5.4	5.2	5.1
25	466	38	5.9	5.6	5.5	5.3
26	485	51	6.0	6.0	5.8	5.7
27	517	94	6.0	6.0	6.0	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.3.5.2

Raw Score to Scale Score to Proficiency Level Conversion: Writ 9-12 B/C S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
0	232	49	1.3	1.2	1.1	1.0
1	234	47	1.3	1.2	1.1	1.0
2	249	32	1.5	1.4	1.3	1.1
3	257	26	1.6	1.5	1.4	1.2
4	263	23	1.7	1.6	1.4	1.3
5	267	21	1.7	1.6	1.5	1.4
6	271	19	1.8	1.6	1.5	1.4
7	275	18	1.8	1.7	1.6	1.5
8	278	17	1.8	1.7	1.6	1.5
9	281	17	1.9	1.8	1.6	1.5
10	284	17	1.9	1.8	1.7	1.6
11	287	16	1.9	1.8	1.7	1.6
12	290	16	2.0	1.9	1.7	1.6
13	292	16	2.1	1.9	1.8	1.6
14	295	16	2.2	1.9	1.8	1.7
15	298	16	2.3	2.0	1.8	1.7
16	301	17	2.4	2.1	1.9	1.8
17	304	17	2.5	2.2	1.9	1.8
18	307	18	2.6	2.3	1.9	1.8
19	310	18	2.7	2.4	2.0	1.9
20	313	19	2.8	2.5	2.1	1.9
21	317	20	2.9	2.6	2.3	1.9
22	321	20	3.0	2.8	2.4	2.1
23	326	21	3.1	3.0	2.6	2.3
24	330	22	3.1	3.0	2.8	2.4
25	335	23	3.2	3.1	3.0	2.6
26	341	23	3.3	3.2	3.1	2.8
27	346	23	3.4	3.3	3.1	3.0
28	352	24	3.5	3.4	3.3	3.1
29	358	24	3.6	3.5	3.4	3.2
30	363	24	3.7	3.6	3.5	3.3
31	369	24	3.8	3.7	3.6	3.4
32	375	24	3.9	3.8	3.7	3.5
33	381	24	4.0	3.9	3.8	3.6

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
34	387	24	4.1	4.0	3.9	3.7
35	392	23	4.2	4.1	4.0	3.8
36	398	23	4.3	4.2	4.1	4.0
37	403	23	4.4	4.3	4.2	4.1
38	408	22	4.5	4.4	4.3	4.2
39	413	22	4.6	4.5	4.4	4.3
40	418	21	4.7	4.6	4.5	4.4
41	422	21	4.8	4.7	4.6	4.4
42	427	21	4.9	4.8	4.7	4.5
43	431	21	5.0	4.9	4.8	4.6
44	436	21	5.1	5.0	4.9	4.7
45	440	21	5.2	5.0	4.9	4.8
46	445	21	5.3	5.2	5.0	4.9
47	450	22	5.5	5.3	5.1	5.0
48	455	23	5.6	5.4	5.2	5.1
49	460	24	5.7	5.5	5.3	5.2
50	467	27	5.9	5.7	5.5	5.3
51	475	30	6.0	5.9	5.6	5.5
52	486	36	6.0	6.0	5.9	5.7
53	504	51	6.0	6.0	6.0	6.0
54	535	95	6.0	6.0	6.0	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

2.6.4 Speaking

Table 2.6.4.1.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 1 Pre-A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	106	50	1.0
1	106	50	1.0
2	122	40	1.1
3	136	37	1.3
4	149	40	1.4
5	162	48	1.5
6	175	61	1.6

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.1.2

Raw Score to Scale Score to Proficiency Level Conversion: Spek 1 A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G1
0	106	46	1.0
1	106	46	1.0
2	117	38	1.1
3	129	34	1.2
4	139	33	1.3
5	148	33	1.4
6	159	34	1.5
7	170	36	1.6
8	182	37	1.7
9	195	39	1.8
10	210	42	2.0
11	228	48	2.4
12	253	55	2.8
13	279	52	3.3
14	302	48	3.8
15	321	47	4.2
16	342	50	4.6
17	363	59	5.0
18	384	75	5.5

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.1.3

Raw Score to Scale Score to Proficiency Level Conversion: Spek 1 B/C S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G1
6	106	43	1.0
7	159	29	1.5
8	167	29	1.6
9	175	29	1.6
10	182	29	1.7
11	190	29	1.8
12	198	29	1.9
13	205	29	2.0
14	214	30	2.1
15	223	32	2.3
16	233	34	2.5
17	244	37	2.6
18	258	38	2.9
19	271	38	3.2
20	284	37	3.4
21	296	35	3.7
22	307	33	3.9
23	317	33	4.1
24	327	33	4.3
25	337	33	4.5
26	348	35	4.7
27	360	38	4.9
28	372	43	5.2
29	384	49	5.5
30	403	64	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.2.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 2-3 Pre-A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G2	PL for G3
0	118	38	1.0	1.0
1	118	38	1.0	1.0
2	118	38	1.0	1.0
3	127	37	1.1	1.0
4	141	40	1.2	1.1
5	155	49	1.3	1.3
6	169	64	1.5	1.4

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.2.2

Raw Score to Scale Score to Proficiency Level Conversion: Spek 2-3 A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G2	PL for G3
0	118	36	1.0	1.0
1	118	36	1.0	1.0
2	118	36	1.0	1.0
3	122	34	1.0	1.0
4	132	33	1.1	1.1
5	143	34	1.2	1.2
6	154	36	1.3	1.3
7	166	37	1.5	1.4
8	179	38	1.6	1.5
9	193	40	1.7	1.6
10	208	42	1.8	1.7
11	227	49	2.1	1.9
12	251	55	2.5	2.3
13	278	52	3.1	2.8
14	300	48	3.5	3.3
15	320	47	3.9	3.7
16	341	50	4.3	4.1
17	362	59	4.7	4.5
18	383	76	5.2	4.9

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.2.3

Raw Score to Scale Score to Proficiency Level Conversion: Spek 2-3 B/C S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G2	PL for G3
6	118	33	1.0	1.0
7	153	30	1.3	1.3
8	161	31	1.4	1.3
9	170	31	1.5	1.4
10	179	30	1.6	1.5
11	187	30	1.6	1.5
12	196	30	1.7	1.6
13	205	31	1.8	1.7
14	213	32	1.9	1.8
15	223	33	2.0	1.9
16	233	34	2.2	1.9
17	245	36	2.4	2.2
18	257	37	2.6	2.4
19	270	37	2.9	2.7
20	282	36	3.1	2.9
21	294	35	3.4	3.2
22	305	34	3.6	3.4
23	315	34	3.8	3.6
24	326	34	4.0	3.8
25	337	34	4.2	4.0
26	348	36	4.5	4.2
27	361	40	4.7	4.5
28	374	44	5.0	4.7
29	387	51	5.3	5.0
30	425	85	6.0	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.3.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 4-5 Pre-A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
0	130	46	1.0	1.0
1	130	46	1.0	1.0
2	141	40	1.1	1.0
3	155	38	1.2	1.1
4	168	40	1.3	1.2
5	181	48	1.4	1.3
6	194	60	1.5	1.5

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.3.2

Raw Score to Scale Score to Proficiency Level Conversion: Spek 4-5 A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
0	130	44	1.0	1.0
1	130	44	1.0	1.0
2	138	40	1.1	1.0
3	151	36	1.2	1.1
4	162	35	1.3	1.2
5	174	37	1.4	1.3
6	187	40	1.5	1.4
7	202	42	1.6	1.5
8	218	42	1.7	1.6
9	234	42	1.9	1.8
10	251	44	2.1	1.9
11	271	49	2.5	2.2
12	295	53	3.0	2.8
13	320	52	3.5	3.3
14	343	48	4.0	3.8
15	364	48	4.4	4.2
16	385	51	4.7	4.6
17	406	60	5.2	4.9
18	427	75	5.7	5.5

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.3.3

Raw Score to Scale Score to Proficiency Level Conversion: Spek 4-5 B/C S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G4	PL for G5
6	130	43	1.0	1.0
7	194	36	1.5	1.5
8	206	36	1.6	1.5
9	217	34	1.7	1.6
10	227	32	1.8	1.7
11	236	32	1.9	1.8
12	245	31	1.9	1.8
13	254	31	2.1	1.9
14	263	32	2.3	2.1
15	272	33	2.5	2.3
16	282	34	2.7	2.5
17	294	36	3.0	2.8
18	306	37	3.2	3.0
19	319	37	3.5	3.3
20	332	37	3.7	3.6
21	344	35	4.0	3.8
22	355	34	4.2	4.0
23	365	34	4.4	4.2
24	376	34	4.6	4.4
25	387	34	4.8	4.6
26	398	36	5.0	4.8
27	411	39	5.3	5.1
28	424	44	5.7	5.4
29	437	51	6.0	5.8
30	450	60	6.0	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.4.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 6-8 Pre-A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	148	40	1.0	1.0	1.0
1	148	40	1.0	1.0	1.0
2	148	40	1.0	1.0	1.0
3	159	38	1.1	1.1	1.0
4	173	41	1.2	1.2	1.1
5	187	49	1.3	1.3	1.2
6	201	63	1.4	1.4	1.3

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.4.2

Raw Score to Scale Score to Proficiency Level Conversion: Spek 6-8 A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
0	148	38	1.0	1.0	1.0
1	148	38	1.0	1.0	1.0
2	148	38	1.0	1.0	1.0
3	156	37	1.1	1.1	1.0
4	168	37	1.2	1.1	1.1
5	181	40	1.3	1.2	1.2
6	197	44	1.4	1.4	1.3
7	215	44	1.5	1.5	1.4
8	232	42	1.7	1.6	1.6
9	248	41	1.8	1.7	1.7
10	264	44	1.9	1.9	1.8
11	283	49	2.3	2.1	1.9
12	308	55	2.9	2.7	2.6
13	335	52	3.5	3.3	3.2
14	358	48	3.9	3.7	3.6
15	377	47	4.2	4.1	4.0
16	398	50	4.6	4.5	4.3
17	419	59	5.0	4.8	4.7
18	440	75	5.6	5.4	5.2

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.4.3

Raw Score to Scale Score to Proficiency Level Conversion: Spek 6-8 B/C S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G6	PL for G7	PL for G8
6	148	38	1.0	1.0	1.0
7	208	40	1.5	1.4	1.4
8	221	37	1.6	1.5	1.5
9	233	34	1.7	1.6	1.6
10	244	33	1.8	1.7	1.7
11	253	32	1.8	1.8	1.7
12	262	31	1.9	1.8	1.8
13	271	31	2.0	1.9	1.9
14	280	32	2.2	2.0	1.9
15	290	33	2.5	2.3	2.1
16	300	34	2.7	2.5	2.4
17	312	36	3.0	2.8	2.7
18	324	37	3.2	3.1	3.0
19	337	37	3.5	3.3	3.2
20	350	36	3.8	3.6	3.5
21	362	35	4.0	3.8	3.7
22	372	34	4.2	4.0	3.9
23	383	34	4.4	4.2	4.1
24	394	34	4.5	4.4	4.3
25	404	34	4.7	4.6	4.4
26	415	36	4.9	4.8	4.6
27	428	39	5.3	5.0	4.9
28	441	44	5.7	5.5	5.2
29	454	51	6.0	5.9	5.7
30	467	59	6.0	6.0	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.5.1

Raw Score to Scale Score to Proficiency Level Conversion: Spek 9-12 Pre-A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
0	172	40	1.1	1.0	1.0	1.0
1	172	40	1.1	1.0	1.0	1.0
2	172	40	1.1	1.0	1.0	1.0
3	183	38	1.2	1.1	1.1	1.0
4	197	40	1.3	1.2	1.2	1.1
5	211	49	1.4	1.3	1.3	1.3
6	225	64	1.5	1.4	1.4	1.4

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.5.2

Raw Score to Scale Score to Proficiency Level Conversion: Spek 9-12 A S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
0	172	37	1.1	1.0	1.0	1.0
1	172	37	1.1	1.0	1.0	1.0
2	172	37	1.1	1.0	1.0	1.0
3	179	36	1.1	1.1	1.0	1.0
4	191	36	1.2	1.2	1.1	1.1
5	203	37	1.3	1.3	1.2	1.2
6	217	40	1.4	1.4	1.3	1.3
7	232	42	1.5	1.5	1.4	1.4
8	248	41	1.6	1.6	1.6	1.5
9	263	41	1.8	1.7	1.7	1.7
10	279	43	1.9	1.8	1.8	1.8
11	298	49	2.2	2.0	1.9	1.9
12	323	55	2.8	2.7	2.6	2.5
13	350	52	3.3	3.2	3.2	3.1
14	372	48	3.7	3.6	3.5	3.4
15	392	47	4.1	3.9	3.8	3.7
16	413	50	4.5	4.3	4.2	4.1
17	434	59	4.8	4.7	4.6	4.5
18	455	75	5.5	5.3	5.1	5.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

Table 2.6.4.5.3

Raw Score to Scale Score to Proficiency Level Conversion: Spek 9-12 B/C S403 Online

Raw Score	Scale Score	CSEM x 1.96	PL for G9	PL for G10	PL for G11	PL for G12
6	172	37	1.1	1.0	1.0	1.0
7	221	35	1.4	1.4	1.4	1.3
8	232	34	1.5	1.5	1.4	1.4
9	242	33	1.6	1.6	1.5	1.5
10	251	31	1.7	1.6	1.6	1.6
11	260	30	1.7	1.7	1.7	1.6
12	268	30	1.8	1.8	1.7	1.7
13	277	30	1.9	1.8	1.8	1.8
14	285	31	1.9	1.9	1.8	1.8
15	294	32	2.1	1.9	1.9	1.9
16	305	34	2.3	2.2	2.1	2.0
17	316	37	2.6	2.5	2.4	2.3
18	330	38	3.0	2.9	2.8	2.7
19	343	38	3.2	3.1	3.0	3.0
20	356	37	3.4	3.3	3.3	3.2
21	368	35	3.7	3.5	3.4	3.4
22	379	34	3.8	3.7	3.6	3.5
23	389	33	4.0	3.9	3.8	3.7
24	399	33	4.2	4.1	3.9	3.8
25	409	34	4.4	4.3	4.1	4.0
26	420	35	4.6	4.5	4.3	4.2
27	432	38	4.8	4.7	4.6	4.5
28	444	43	5.1	4.9	4.8	4.7
29	456	49	5.5	5.4	5.2	5.0
30	476	64	6.0	6.0	6.0	6.0

Note: Score reports provided to students include the CSEM value multiplied by 1.96.

2.7 Equating Summary

Each year a certain percentage of items on each ACCESS for ELLs test form are new, as determined by the refreshment plan for that series. For Series 403, all four domains were refreshed. An equating procedure known as common-item equating is used to equate the results on new forms to the older forms. In this procedure, the difficulty measures for items that appear on both the new and the old forms are kept constant across both forms. Thus, performances on the newer form may be interpreted with the same frame of reference. Many items appearing on ACCESS Online Series 403 also appeared on Series 402. All items common to both forms were anchored to their 402 values in the first equating run. In addition, for the Speaking domain, difficulty measures for the new tasks were anchored to their initial calibrated values from the Speaking field test analysis. After the first equating run, some items that were originally anchored, either to their operational or to their field test value, proved to have changed in their difficulty measure. This change is measured by the "Displacement" statistic. This statistic shows the difference between the difficulty value of the anchored item and what its difficulty value would have been had it not been anchored. Typically, displacements of less than 0.5 logits are unlikely to have much impact on measurement in a test instrument (Linacre, n.d.). For Listening and Reading items and for Writing tasks, if this value was large (i.e., usually above 0.30 or below -0.30), that item was unanchored in the final equating run (i.e., it was treated as if it were a new item). For Speaking tasks, a slightly different displacement criterion (above 0.50 or below -0.50) was used since anchored tasks from the Speaking domain have been shown to be less stable than items and tasks from other domains.

A pre-equating design was used to conduct the annual equating for Listening and Reading domains. This design allows for Listening and Reading item parameters to be available for setting up the computer adaptive engine prior to operational administration. For the Listening and Reading domains, student data collected from the Series 403 embedded field test were used to conduct the equating analyses. All available student data at the time the equating analyses were conducted were included in the analyses.

For the Writing domain, the equating analysis was conducted using Series 403 operational data collected during the early testing window. The Writing equating study was conducted with a random sample drawn with a target sample size of 3,000 students per grade-level cluster. The Writing equating sample was drawn so that it was proportional to the Series 401 and Series 402 population means for the Writing domain, by grade and tiered form.

For the Speaking domain, student data from the Series 403 appended Speaking field test administration were used to conduct the initial common-item equating. These initial item parameters were then verified using Series 403 operational data collected during the early testing window. The Speaking verification study was conducted using a random sample drawn with a target sample size of 3,000 students per grade-level cluster. The Speaking verification sample was drawn such that it was proportional to the Series 401 and Series 402 population means for the Speaking domain, by grade and tier.

Tables in this section present a summary of the equating and verification procedures. The first section of the table compares the current test (i.e., the Series 403 version of that test form) to the previous year's test (i.e., the Series 402 version of that test form). The number of items, the average item difficulty, the standard deviation of the item difficulty values, and the difficulty value of the easiest and hardest item on each test form are shown. These values are in terms of logits used in the Rasch measurement model.

The second section of the table presents information on the anchoring items. The total number of possible anchors that were initially anchored to their previous values is shown, as well as the standard deviation of those items. Next, the number of items that were anchored in the final equating run is shown, again with the average item difficulty and standard deviation. Finally, the percentage of items that served as anchors and their average displacement value is given. Generally speaking, the greater the number of tasks anchored and the closer their average displacement is to 0.00, the more trustworthy the equating results will be.

The third section of the tables gives information about the anchor items or tasks, both by order of displacement statistics and by order of item difficulty. The displacement statistics provide information on the difference between the difficulty value of the anchored item and what that difficulty value would have been had the item not been anchored. Smaller displacement statistics indicate more consistency between the item's difficulty value on the Series 403 test form and on the Series 402 test form. It is desirable that the anchor items represent a wide range of difficulties across the entire spectrum of the item difficulty values on a test form.

In general, and for longer tests such as Listening and Reading, the greater the representation across the difficulty range of anchor items, the more trustworthy the equating results will be. For the Writing and Speaking domains, which are shorter and performance based, and which have additional content and exposure considerations in terms of item refreshment, this rule of thumb may not always apply. In addition, the number of anchors is also a function of the targeted refreshment plan, which can differ by series and by domains.

For the Writing and Speaking tasks, this table has a fourth section, which provides the anchored step measures for the score on each task. For the ACCESS Writing and Speaking tasks, a Rasch-grouped rating scale model is used (see Part 2, Chapter 2 "Analysis of Domains"). For Writing, the step difficulty values are the same for all the tasks that are scored on the 0–9 raw score scale. These constant step difficulty values help to provide anchors in the calibration of new Writing tasks onto the common WIDA score scale each year. For Speaking, the step difficulty values for all PL 1 tasks are the same and the step difficulty values for all PL 3 and PL 5 tasks are the same. As with Writing, these constant step difficulty values help to provide anchors in the calibration of new Speaking tasks onto the common WIDA score scale each year.

2.7.1 Listening

2.7.1.1 Grade 1

Table 2.7.1.1

Equating Summary: List 1 S403 Online

Comparison of Forms	Form 403		Form 402	
	No. of Items	**Average Difficulty (Std. Dev.)**	**No. of Items**	**Average Difficulty (Std. Dev.)**
	54	-1.16 (1.33)	54	-1.25 (1.33)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	-4.38	1.34	-4.38	1.46
Anchoring Items	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	42	-1.48 (1.27)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	35	-1.48 (1.39)		
	Percentage Anchors	**Average Displacement**		
	65%	0.00		

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	13803	-1.67	-0.29	12445	-4.38	0.18
	12692	-3.17	-0.23	12690	-3.39	-0.02
	14952	-3.03	-0.20	13889	-3.27	0.05
	14837	-2.11	-0.19	12692	-3.17	-0.23
	13816	-1.89	-0.16	14952	-3.03	-0.20
	12691	-2.96	-0.15	13810	-2.97	0.00
	13898	-0.94	-0.09	12691	-2.96	-0.15
	13877	1.34	-0.09	13891	-2.55	0.05
	11667	0.44	-0.08	12446	-2.54	0.19
	13878	0.31	-0.06	13808	-2.41	0.01
	11671	-2.02	-0.04	13802	-2.40	0.03
	12746	-0.86	-0.04	13890	-2.23	0.08
	13879	-0.64	-0.04	14837	-2.11	-0.19
	12754	-0.47	-0.02	11671	-2.02	-0.04
	12690	-3.39	-0.02	13809	-1.90	0.03

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	11668	-0.81	-0.01	13816	-1.89	-0.16
	12447	1.15	-0.01	13900	-1.82	0.19
	13810	-2.97	0.00	13803	-1.67	-0.29
	13808	-2.41	0.01	14899	-1.50	0.16
	13802	-2.40	0.03	12745	-1.35	0.08
	13809	-1.90	0.03	13899	-1.15	0.04
	13899	-1.15	0.04	13804	-1.03	0.17
	13891	-2.55	0.05	13898	-0.94	-0.09
	13889	-3.27	0.05	12746	-0.86	-0.04
	12745	-1.35	0.08	11668	-0.81	-0.01
	13890	-2.23	0.08	13879	-0.64	-0.04
	12848	0.10	0.12	12754	-0.47	-0.02
	12846	0.58	0.12	13815	-0.38	0.21
	12847	0.02	0.14	12847	0.02	0.14
	14899	-1.50	0.16	12848	0.10	0.12
	13804	-1.03	0.17	13878	0.31	-0.06
	12445	-4.38	0.18	11667	0.44	-0.08
	12446	-2.54	0.19	12846	0.58	0.12
	13900	-1.82	0.19	12447	1.15	-0.01
	13815	-0.38	0.21	13877	1.34	-0.09

2.7.1.2 *Grades 2–3*

Table 2.7.1.2

Equating Summary: List 2-3 S403 Online

Comparison of Forms	Form 403		Form 402	
	No. of Items	Average Difficulty (Std. Dev.)	No. of Items	Average Difficulty (Std. Dev.)
	54	-1.25 (1.56)	54	-1.15 (1.63)
	Easiest	Hardest	Easiest	Hardest
	-4.12	1.64	-3.89	1.84
Anchoring Items	No. of Possible Anchors	Average Difficulty (Std. Dev.)		
	45	-1.48 (1.53)		
	No. of Anchors Used	Average Difficulty (Std. Dev.)		
	42	-1.48 (1.49)		
	Percentage Anchors	Average Displacement		
	78%	-0.04		

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	Item Difficulty	Displacement	Item ID	Item Difficulty	Displacement
	13790	-2.68	-0.26	12731	-3.89	-0.05
	12734	0.26	-0.24	13788	-3.85	-0.11
	12706	-1.64	-0.24	13797	-3.68	-0.07
	13799	-2.87	-0.22	11544	-3.56	-0.06
	11546	-1.50	-0.21	13905	-3.26	0.13
	12705	-0.17	-0.20	12825	-3.26	-0.12
	13798	-2.72	-0.17	13904	-3.24	-0.02
	12828	-2.29	-0.14	12956	-3.09	0.05
	12813	-1.83	-0.12	13799	-2.87	-0.22
	11545	-1.90	-0.12	13798	-2.72	-0.17
	12825	-3.26	-0.12	13790	-2.68	-0.26
	13788	-3.85	-0.11	14884	-2.64	-0.11
	14884	-2.64	-0.11	12732	-2.36	0.08
	12787	-0.11	-0.11	13910	-2.33	0.07
	12988	-1.07	-0.10	12828	-2.29	-0.14

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	Item Difficulty	Displacement	Item ID	Item Difficulty	Displacement
	13906	-2.24	-0.10	13906	-2.24	-0.10
	13797	-3.68	-0.07	12730	-2.07	-0.02
	11544	-3.56	-0.06	12954	-1.97	0.04
	12815	-0.47	-0.05	11545	-1.90	-0.12
	12785	0.00	-0.05	12813	-1.83	-0.12
	12786	1.56	-0.05	12733	-1.69	0.03
	12731	-3.89	-0.05	12706	-1.64	-0.24
	12707	-0.30	-0.05	11546	-1.50	-0.21
	12830	-1.17	-0.04	12814	-1.26	-0.03
	12814	-1.26	-0.03	12830	-1.17	-0.04
	12730	-2.07	-0.02	12957	-1.12	0.10
	13904	-3.24	-0.02	12988	-1.07	-0.10
	13912	-0.24	0.01	13911	-0.87	0.26
	12733	-1.69	0.03	12815	-0.47	-0.05
	12954	-1.97	0.04	12707	-0.30	-0.05
	12956	-3.09	0.05	13912	-0.24	0.01
	13910	-2.33	0.07	12705	-0.17	-0.20
	12732	-2.36	0.08	12787	-0.11	-0.11
	12971	0.35	0.08	12991	-0.07	0.18
	12735	1.28	0.10	12785	0.00	-0.05
	12957	-1.12	0.10	12953	0.13	0.12
	12955	1.23	0.11	12734	0.26	-0.24
	12953	0.13	0.12	12990	0.29	0.16
	13905	-3.26	0.13	12971	0.35	0.08
	12990	0.29	0.16	12955	1.23	0.11
	12991	-0.07	0.18	12735	1.28	0.10
	13911	-0.87	0.26	12786	1.56	-0.05

2.7.1.3 *Grades 4–5*

Table 2.7.1.3

Equating Summary: List 4-5 S403 Online

<table>
<tr><td rowspan="5">Comparison of Forms</td><td colspan="3">Form 403</td><td colspan="3">Form 402</td></tr>
<tr><td colspan="2">No. of Items</td><td>Average Difficulty (Std. Dev.)</td><td>No. of Items</td><td colspan="2">Average Difficulty (Std. Dev.)</td></tr>
<tr><td colspan="2">54</td><td>0.41 (1.72)</td><td>54</td><td colspan="2">0.45 (1.80)</td></tr>
<tr><td colspan="2">Easiest</td><td>Hardest</td><td>Easiest</td><td colspan="2">Hardest</td></tr>
<tr><td colspan="2">-2.65</td><td>4.50</td><td>-2.76</td><td colspan="2">4.50</td></tr>
<tr><td rowspan="6">Anchoring Items</td><td colspan="2">No. of Possible Anchors</td><td>Average Difficulty (Std. Dev.)</td><td colspan="3" rowspan="6"></td></tr>
<tr><td colspan="2">42</td><td>0.24 (1.79)</td></tr>
<tr><td colspan="2">No. of Anchors Used</td><td>Average Difficulty (Std. Dev.)</td></tr>
<tr><td colspan="2">27</td><td>0.48 (1.98)</td></tr>
<tr><td colspan="2">Percentage Anchors</td><td>Average Displacement</td></tr>
<tr><td colspan="2">50%</td><td>0.01</td></tr>
<tr><td rowspan="17">Displacement of Anchor Items</td><td colspan="3">Anchor Items by Displacement</td><td colspan="3">Anchor Items by Item Difficulty</td></tr>
<tr><td>Item ID</td><td>Item Difficulty</td><td>Displacement</td><td>Item ID</td><td>Item Difficulty</td><td>Displacement</td></tr>
<tr><td>14945</td><td>0.10</td><td>-0.28</td><td>12919</td><td>-2.65</td><td>-0.01</td></tr>
<tr><td>14214</td><td>4.50</td><td>-0.25</td><td>12793</td><td>-2.36</td><td>0.10</td></tr>
<tr><td>12917</td><td>-0.39</td><td>-0.19</td><td>12792</td><td>-1.97</td><td>-0.09</td></tr>
<tr><td>13028</td><td>3.20</td><td>-0.18</td><td>13024</td><td>-1.78</td><td>-0.04</td></tr>
<tr><td>12794</td><td>-0.22</td><td>-0.13</td><td>12571</td><td>-1.61</td><td>0.10</td></tr>
<tr><td>13027</td><td>3.33</td><td>-0.11</td><td>12416</td><td>-1.40</td><td>0.09</td></tr>
<tr><td>12792</td><td>-1.97</td><td>-0.09</td><td>12394</td><td>-0.89</td><td>0.17</td></tr>
<tr><td>12918</td><td>-0.47</td><td>-0.05</td><td>13026</td><td>-0.76</td><td>-0.02</td></tr>
<tr><td>13024</td><td>-1.78</td><td>-0.04</td><td>12918</td><td>-0.47</td><td>-0.05</td></tr>
<tr><td>13025</td><td>-0.38</td><td>-0.04</td><td>12917</td><td>-0.39</td><td>-0.19</td></tr>
<tr><td>14845</td><td>0.17</td><td>-0.03</td><td>13025</td><td>-0.38</td><td>-0.04</td></tr>
<tr><td>13026</td><td>-0.76</td><td>-0.02</td><td>12794</td><td>-0.22</td><td>-0.13</td></tr>
<tr><td>12919</td><td>-2.65</td><td>-0.01</td><td>14833</td><td>-0.21</td><td>0.05</td></tr>
<tr><td>14212</td><td>2.49</td><td>0.02</td><td>14945</td><td>0.10</td><td>-0.28</td></tr>
<tr><td>14833</td><td>-0.21</td><td>0.05</td><td>14845</td><td>0.17</td><td>-0.03</td></tr>
</table>

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	Item Difficulty	Displacement	Item ID	Item Difficulty	Displacement
	13029	1.83	0.06	12597	0.43	0.08
	12597	0.43	0.08	14213	1.13	0.12
	14846	3.22	0.08	12396	1.26	0.13
	12416	-1.40	0.09	12942	1.44	0.29
	12793	-2.36	0.10	12946	1.46	0.18
	12571	-1.61	0.10	13029	1.83	0.06
	14213	1.13	0.12	14212	2.49	0.02
	12396	1.26	0.13	13028	3.20	-0.18
	12394	-0.89	0.17	14846	3.22	0.08
	12946	1.46	0.18	13027	3.33	-0.11
	12943	3.60	0.23	12943	3.60	0.23
	12942	1.44	0.29	14214	4.50	-0.25

2.7.1.4 Grades 6–8

Table 2.7.1.4

Equating Summary: List 6-8 S403 Online

Comparison of Forms	**Form 403**		**Form 402**	
	No. of Items	**Average Difficulty (Std. Dev.)**	**No. of Items**	**Average Difficulty (Std. Dev.)**
	54	1.22 (1.35)	54	1.16 (1.36)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	-2.40	3.55	-2.40	3.95
Anchoring Items	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	42	1.07 (1.41)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	27	0.75 (1.33)		
	Percentage Anchors	**Average Displacement**		
	50%	-0.04		

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	13062	-0.13	-0.28	12884	-2.40	-0.17
	11474	1.71	-0.25	13061	-0.79	-0.09
	14863	1.43	-0.24	12380	-0.74	0.15
	12849	-0.68	-0.22	13060	-0.69	-0.22
	13060	-0.69	-0.22	12849	-0.68	-0.22
	12850	0.90	-0.21	12885	-0.67	-0.12
	14915	1.08	-0.18	12922	-0.51	0.04
	14917	1.05	-0.18	12454	-0.49	0.13
	12884	-2.40	-0.17	13062	-0.13	-0.28
	12886	0.11	-0.13	12455	-0.01	0.27
	12885	-0.67	-0.12	12886	0.11	-0.13
	12899	2.09	-0.10	11467	0.84	0.18
	13061	-0.79	-0.09	12850	0.90	-0.21
	13838	2.51	-0.07	14917	1.05	-0.18
	12901	2.25	-0.06	14915	1.08	-0.18

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	Item Difficulty	Displacement	Item ID	Item Difficulty	Displacement
	12905	2.50	-0.05	12456	1.26	0.22
	13840	2.02	0.02	12381	1.30	0.16
	13839	2.15	0.03	14863	1.43	-0.24
	12382	2.70	0.03	11472	1.61	0.11
	12922	-0.51	0.04	11474	1.71	-0.25
	11472	1.61	0.11	13840	2.02	0.02
	12454	-0.49	0.13	12899	2.09	-0.10
	12380	-0.74	0.15	13839	2.15	0.03
	12381	1.30	0.16	12901	2.25	-0.06
	11467	0.84	0.18	12905	2.50	-0.05
	12456	1.26	0.22	13838	2.51	-0.07
	12455	-0.01	0.27	12382	2.70	0.03

2.7.1.5 *Grades 9–12*

Table 2.7.1.5

Equating Summary: List 9-12 S403 Online

Comparison of Forms	Form 403		Form 402	
	No. of Items	**Average Difficulty (Std. Dev.)**	**No. of Items**	**Average Difficulty (Std. Dev.)**
	54	1.43 (1.19)	54	1.46 (1.16)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	-0.80	3.91	-0.52	3.91
Anchoring Items	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	45	1.37 (1.26)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	39	1.57 (1.21)		
	Percentage Anchors	**Average Displacement**		
	72%	-0.01		

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	14888	1.28	-0.24	11035	-0.52	0.00
	14885	2.16	-0.23	12869	-0.48	-0.05
	15651	2.46	-0.20	12360	-0.42	0.01
	12870	0.37	-0.15	11034	-0.40	-0.06
	14886	1.96	-0.13	12893	0.07	0.09
	13867	3.88	-0.13	12444	0.11	0.09
	11198	1.98	-0.12	12870	0.37	-0.15
	11344	3.91	-0.10	11199	0.40	0.03
	12871	1.19	-0.09	11036	0.51	-0.07
	13865	2.68	-0.08	14954	0.53	0.12
	14955	1.98	-0.08	15649	0.54	0.10
	11036	0.51	-0.07	13036	1.07	0.21
	12443	2.22	-0.07	12442	1.10	0.10
	14956	1.86	-0.07	12871	1.19	-0.09
	11034	-0.40	-0.06	11245	1.21	0.11

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	12869	-0.48	-0.05	15650	1.23	0.09
	13037	2.66	-0.04	14888	1.28	-0.24
	13038	3.60	-0.02	12903	1.40	0.08
	11035	-0.52	0.00	11215	1.44	0.05
	12360	-0.42	0.01	12441	1.47	0.13
	12895	2.83	0.01	12720	1.60	0.06
	12907	2.18	0.02	14956	1.86	-0.07
	11199	0.40	0.03	14886	1.96	-0.13
	11215	1.44	0.05	14955	1.98	-0.08
	11343	2.65	0.05	11198	1.98	-0.12
	12720	1.60	0.06	12361	2.15	0.10
	12903	1.40	0.08	14885	2.16	-0.23
	12893	0.07	0.09	12907	2.18	0.02
	15650	1.23	0.09	12443	2.22	-0.07
	12444	0.11	0.09	15651	2.46	-0.20
	15649	0.54	0.10	11343	2.65	0.05
	12361	2.15	0.10	13037	2.66	-0.04
	12442	1.10	0.10	13865	2.68	-0.08
	11245	1.21	0.11	12894	2.82	0.13
	13866	3.49	0.12	12895	2.83	0.01
	14954	0.53	0.12	13866	3.49	0.12
	12441	1.47	0.13	13038	3.60	-0.02
	12894	2.82	0.13	13867	3.88	-0.13
	13036	1.07	0.21	11344	3.91	-0.10

2.7.2 Reading

2.7.2.1 Grade 1

Table 2.7.2.1

Equating Summary: Read 1 S403 Online

<table>
<tr><td rowspan="6">Comparison of Forms</td><td colspan="3">Form 403</td><td colspan="3">Form 402</td></tr>
<tr><td>No. of Items</td><td colspan="2">Average Difficulty (Std. Dev.)</td><td>No. of Items</td><td colspan="2">Average Difficulty (Std. Dev.)</td></tr>
<tr><td>72</td><td colspan="2">-1.00 (1.14)</td><td>72</td><td colspan="2">-0.95 (1.10)</td></tr>
<tr><td>Easiest</td><td colspan="2">Hardest</td><td>Easiest</td><td colspan="2">Hardest</td></tr>
<tr><td>-4.46</td><td colspan="2">1.01</td><td>-4.46</td><td colspan="2">0.79</td></tr>
<tr><td></td><td colspan="2"></td><td></td><td colspan="2"></td></tr>
<tr><td rowspan="6">Anchoring Items</td><td>No. of Possible Anchors</td><td colspan="2">Average Difficulty (Std. Dev.)</td><td colspan="3" rowspan="6"></td></tr>
<tr><td>60</td><td colspan="2">-1.13 (1.19)</td></tr>
<tr><td>No. of Anchors Used</td><td colspan="2">Average Difficulty (Std. Dev.)</td></tr>
<tr><td>50</td><td colspan="2">-1.13 (1.09)</td></tr>
<tr><td>Percentage Anchors</td><td colspan="2">Average Displacement</td></tr>
<tr><td>69%</td><td colspan="2">-0.01</td></tr>
<tr><td rowspan="17">Displacement of Anchor Items</td><td colspan="3">Anchor Items by Displacement</td><td colspan="3">Anchor Items by Item Difficulty</td></tr>
<tr><td>Item ID</td><td>Item Difficulty</td><td>Displacement</td><td>Item ID</td><td>Item Difficulty</td><td>Displacement</td></tr>
<tr><td>16040</td><td>-0.61</td><td>-0.29</td><td>13313</td><td>-4.46</td><td>-0.24</td></tr>
<tr><td>13246</td><td>-0.45</td><td>-0.27</td><td>13274</td><td>-4.07</td><td>-0.14</td></tr>
<tr><td>13313</td><td>-4.46</td><td>-0.24</td><td>13315</td><td>-2.58</td><td>0.12</td></tr>
<tr><td>13201</td><td>-1.09</td><td>-0.24</td><td>13314</td><td>-2.42</td><td>0.08</td></tr>
<tr><td>13212</td><td>-1.67</td><td>-0.23</td><td>13211</td><td>-2.30</td><td>0.05</td></tr>
<tr><td>13200</td><td>-2.17</td><td>-0.19</td><td>13213</td><td>-2.28</td><td>-0.05</td></tr>
<tr><td>13257</td><td>0.61</td><td>-0.17</td><td>13204</td><td>-2.19</td><td>-0.08</td></tr>
<tr><td>13284</td><td>-1.71</td><td>-0.16</td><td>13200</td><td>-2.17</td><td>-0.19</td></tr>
<tr><td>13258</td><td>0.14</td><td>-0.16</td><td>13193</td><td>-2.11</td><td>0.04</td></tr>
<tr><td>13203</td><td>-1.75</td><td>-0.14</td><td>13194</td><td>-2.06</td><td>0.10</td></tr>
<tr><td>13274</td><td>-4.07</td><td>-0.14</td><td>13199</td><td>-1.94</td><td>0.07</td></tr>
<tr><td>13256</td><td>0.57</td><td>-0.14</td><td>13202</td><td>-1.81</td><td>-0.11</td></tr>
<tr><td>13202</td><td>-1.81</td><td>-0.11</td><td>13203</td><td>-1.75</td><td>-0.14</td></tr>
<tr><td>13220</td><td>-0.83</td><td>-0.09</td><td>13284</td><td>-1.71</td><td>-0.16</td></tr>
<tr><td>13204</td><td>-2.19</td><td>-0.08</td><td>13212</td><td>-1.67</td><td>-0.23</td></tr>
</table>

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	13240	-0.95	-0.07	16039	-1.61	0.00
	13245	-0.81	-0.07	13238	-1.58	0.05
	13243	0.47	-0.06	13197	-1.54	0.06
	13241	0.50	-0.06	13195	-1.52	0.00
	13213	-2.28	-0.05	16035	-1.29	0.26
	13198	-0.64	-0.05	13283	-1.23	0.17
	13242	0.57	-0.01	13276	-1.13	0.02
	16039	-1.61	0.00	13196	-1.11	0.03
	13195	-1.52	0.00	13275	-1.10	0.05
	13222	-1.04	0.01	13201	-1.09	-0.24
	13276	-1.13	0.02	16036	-1.05	0.23
	13196	-1.11	0.03	13222	-1.04	0.01
	13285	-0.62	0.03	13217	-1.03	0.08
	14722	0.08	0.03	13221	-1.02	0.05
	14721	-0.08	0.03	13240	-0.95	-0.07
	14723	0.10	0.04	13220	-0.83	-0.09
	13193	-2.11	0.04	13245	-0.81	-0.07
	13238	-1.58	0.05	13244	-0.80	0.09
	13239	-0.60	0.05	13198	-0.64	-0.05
	13211	-2.30	0.05	13285	-0.62	0.03
	13221	-1.02	0.05	16040	-0.61	-0.29
	13275	-1.10	0.05	13239	-0.60	0.05
	13197	-1.54	0.06	13246	-0.45	-0.27
	13199	-1.94	0.07	13218	-0.45	0.20
	13217	-1.03	0.08	14721	-0.08	0.03
	13314	-2.42	0.08	14620	0.03	0.29
	13244	-0.80	0.09	13219	0.05	0.29
	13194	-2.06	0.10	14722	0.08	0.03

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	13315	-2.58	0.12	14723	0.10	0.04
	13283	-1.23	0.17	13258	0.14	-0.16
	13218	-0.45	0.20	13243	0.47	-0.06
	16036	-1.05	0.23	13241	0.50	-0.06
	16035	-1.29	0.26	13256	0.57	-0.14
	14620	0.03	0.29	13242	0.57	-0.01
	13219	0.05	0.29	13257	0.61	-0.17

2.7.2.2 *Grades 2–3*

Table 2.7.2.2

Equating Summary: Read 2-3 S403 Online

Comparison of Forms	**Form 403**		**Form 402**	
	No. of Items	**Average Difficulty (Std. Dev.)**	**No. of Items**	**Average Difficulty (Std. Dev.)**
	72	-0.37 (1.03)	72	-0.44 (1.06)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	-2.99	1.34	-3.26	1.34
Anchoring Items	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	63	-0.46 (1.06)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	57	-0.53 (1.04)		
	Percentage Anchors	**Average Displacement**		
	79%	-0.02		

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	13401	-1.23	-0.30	13292	-2.99	0.04
	13355	-0.57	-0.29	13353	-2.97	-0.27
	13353	-2.97	-0.27	13278	-2.72	0.19
	13346	0.60	-0.26	13325	-2.62	0.03
	13403	-0.92	-0.25	13293	-2.03	-0.18
	13421	0.70	-0.23	13294	-1.88	0.05
	15716	0.36	-0.19	13392	-1.78	0.14
	13279	-0.51	-0.18	15626	-1.69	0.03
	13393	-1.25	-0.18	13419	-1.67	-0.09
	13293	-2.03	-0.18	13331	-1.45	-0.10
	13375	1.34	-0.17	13277	-1.42	-0.03
	13344	0.33	-0.16	13372	-1.27	0.04
	13415	-0.14	-0.14	13326	-1.25	-0.08
	13413	-0.47	-0.12	13393	-1.25	-0.18
	13331	-1.45	-0.10	13394	-1.25	0.14

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	13338	0.80	-0.09	13401	-1.23	-0.30
	13419	-1.67	-0.09	13354	-1.20	0.11
	13376	0.66	-0.08	13327	-1.01	-0.02
	13326	-1.25	-0.08	13403	-0.92	-0.25
	13340	-0.25	-0.07	13420	-0.90	0.13
	13358	0.54	-0.06	15700	-0.83	0.09
	14584	0.92	-0.05	15628	-0.79	0.03
	13412	-0.11	-0.05	13414	-0.76	0.23
	13373	-0.63	-0.05	13373	-0.63	-0.05
	13333	-0.34	-0.04	13355	-0.57	-0.29
	13277	-1.42	-0.03	13279	-0.51	-0.18
	13327	-1.01	-0.02	13413	-0.47	-0.12
	13345	1.24	-0.01	13332	-0.43	0.14
	13932	-0.08	0.00	14589	-0.34	0.15
	14590	-0.21	0.01	13333	-0.34	-0.04
	14583	0.05	0.01	13410	-0.34	0.02
	13410	-0.34	0.02	13374	-0.30	0.05
	15628	-0.79	0.03	13340	-0.25	-0.07
	15626	-1.69	0.03	14590	-0.21	0.01
	13325	-2.62	0.03	13371	-0.21	0.21
	13372	-1.27	0.04	13415	-0.14	-0.14
	13292	-2.99	0.04	13357	-0.13	0.15
	13339	0.38	0.04	13412	-0.11	-0.05
	13411	0.21	0.04	13932	-0.08	0.00
	13374	-0.30	0.05	14583	0.05	0.01
	13294	-1.88	0.05	13356	0.07	0.14
	15700	-0.83	0.09	13411	0.21	0.04
	13354	-1.20	0.11	13344	0.33	-0.16
	13937	0.42	0.12	15716	0.36	-0.19
	13420	-0.90	0.13	13339	0.38	0.04

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	Item Difficulty	Displacement	Item ID	Item Difficulty	Displacement
	13394	-1.25	0.14	13937	0.42	0.12
	13392	-1.78	0.14	13936	0.45	0.21
	13332	-0.43	0.14	13358	0.54	-0.06
	13356	0.07	0.14	13346	0.60	-0.26
	13357	-0.13	0.15	14585	0.64	0.20
	14589	-0.34	0.15	13376	0.66	-0.08
	14591	0.81	0.16	13421	0.70	-0.23
	13278	-2.72	0.19	13338	0.80	-0.09
	14585	0.64	0.20	14591	0.81	0.16
	13371	-0.21	0.21	14584	0.92	-0.05
	13936	0.45	0.21	13345	1.24	-0.01
	13414	-0.76	0.23	13375	1.34	-0.17

2.7.2.3 *Grades 4–5*

Table 2.7.2.3

Equating Summary: Read 4-5 S403 Online

Comparison of Forms	**Form 403**		**Form 402**	
	No. of Items	**Average Difficulty (Std. Dev.)**	**No. of Items**	**Average Difficulty (Std. Dev.)**
	69	0.78 (1.30)	69	0.77 (1.29)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	-3.19	3.31	-3.19	3.18
Anchoring Items	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	63	0.70 (1.32)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	57	0.61 (1.28)		
	Percentage Anchors	**Average Displacement**		
	83%	-0.01		

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	15993	0.01	-0.30	13509	-3.19	-0.15
	13469	1.15	-0.26	13518	-1.94	0.00
	16019	1.59	-0.22	15991	-1.81	-0.10
	15706	0.21	-0.18	13464	-1.61	-0.04
	15708	1.38	-0.18	13519	-1.12	0.04
	13509	-3.19	-0.15	13510	-1.06	-0.03
	13533	1.43	-0.15	13465	-0.93	0.03
	13480	0.75	-0.14	14715	-0.75	0.13
	13407	-0.72	-0.13	13407	-0.72	-0.13
	13479	-0.69	-0.12	13479	-0.69	-0.12
	16011	0.12	-0.11	13539	-0.63	-0.03
	13504	2.57	-0.10	13466	-0.43	-0.08
	16018	1.28	-0.10	14714	-0.37	0.00
	15991	-1.81	-0.10	13520	-0.31	-0.03
	13540	1.92	-0.09	14626	-0.26	0.19

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	13466	-0.43	-0.08	13409	-0.17	0.12
	13505	1.96	-0.08	13451	-0.11	-0.04
	15707	1.23	-0.08	16009	-0.10	-0.04
	13511	1.08	-0.07	15993	0.01	-0.30
	13528	1.64	-0.06	13408	0.06	0.20
	16017	0.09	-0.06	16017	0.09	-0.06
	13530	1.35	-0.06	16011	0.12	-0.11
	13527	1.08	-0.06	14716	0.20	0.09
	13464	-1.61	-0.04	15706	0.21	-0.18
	16009	-0.10	-0.04	13449	0.25	0.01
	13535	2.38	-0.04	13488	0.40	0.17
	13451	-0.11	-0.04	13467	0.47	-0.02
	13531	2.11	-0.03	13480	0.75	-0.14
	13520	-0.31	-0.03	13529	0.98	0.18
	13510	-1.06	-0.03	13534	0.99	0.12
	13539	-0.63	-0.03	13511	1.08	-0.07
	13484	2.41	-0.02	13527	1.08	-0.06
	13467	0.47	-0.02	13469	1.15	-0.26
	14714	-0.37	0.00	15707	1.23	-0.08
	13518	-1.94	0.00	16018	1.28	-0.10
	13449	0.25	0.01	13468	1.31	0.05
	13465	-0.93	0.03	13489	1.33	0.13
	13481	1.68	0.03	13530	1.35	-0.06
	13532	1.51	0.03	15708	1.38	-0.18
	13519	-1.12	0.04	13533	1.43	-0.15
	13482	2.46	0.05	13926	1.50	0.27
	13468	1.31	0.05	13532	1.51	0.03
	13450	1.77	0.06	14627	1.59	0.16
	14716	0.20	0.09	16019	1.59	-0.22
	13534	0.99	0.12	13528	1.64	-0.06

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	Item Difficulty	Displacement	Item ID	Item Difficulty	Displacement
	13409	-0.17	0.12	13481	1.68	0.03
	13541	2.13	0.13	13450	1.77	0.06
	14715	-0.75	0.13	13540	1.92	-0.09
	13489	1.33	0.13	13505	1.96	-0.08
	13503	2.15	0.13	13531	2.11	-0.03
	14627	1.59	0.16	13483	2.13	0.22
	13488	0.40	0.17	13541	2.13	0.13
	13529	0.98	0.18	13503	2.15	0.13
	14626	-0.26	0.19	13535	2.38	-0.04
	13408	0.06	0.20	13484	2.41	-0.02
	13483	2.13	0.22	13482	2.46	0.05
	13926	1.50	0.27	13504	2.57	-0.10

2.7.2.4 Grades 6–8

Table 2.7.2.4

Equating Summary: Read 6-8 S403 Online

Comparison of Forms	**Form 403**		**Form 402**	
	No. of Items	**Average Difficulty (Std. Dev.)**	**No. of Items**	**Average Difficulty (Std. Dev.)**
	72	1.18 (1.50)	72	1.10 (1.46)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	-2.04	3.66	-2.04	3.78
Anchoring Items	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	60	0.89 (1.46)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	57	0.87 (1.44)		
	Percentage Anchors	**Average Displacement**		
	79%	0.00		

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	15713	-1.16	-0.21	13554	-2.04	-0.07
	13630	1.84	-0.19	13575	-1.58	-0.14
	13555	-1.30	-0.18	13566	-1.37	-0.15
	13604	1.87	-0.16	13555	-1.30	-0.18
	13566	-1.37	-0.15	13565	-1.25	0.13
	13963	1.94	-0.15	14641	-1.23	-0.04
	13629	0.78	-0.14	15713	-1.16	-0.21
	13575	-1.58	-0.14	13573	-1.03	-0.05
	14491	1.72	-0.13	13564	-0.98	0.09
	13657	1.62	-0.13	13572	-0.68	0.05
	13962	1.39	-0.12	14640	-0.55	0.04
	13585	0.57	-0.12	13563	-0.52	0.02
	16023	0.24	-0.12	13574	-0.52	-0.02
	13631	0.84	-0.11	13576	-0.48	0.09
	16022	0.19	-0.11	13567	-0.34	-0.03

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	13554	-2.04	-0.07	13577	-0.09	0.14
	13573	-1.03	-0.05	13556	-0.09	-0.04
	13587	0.31	-0.05	16024	-0.03	-0.02
	13618	3.13	-0.05	13568	0.15	0.19
	13658	3.59	-0.05	16022	0.19	-0.11
	13556	-0.09	-0.04	16023	0.24	-0.12
	14641	-1.23	-0.04	13587	0.31	-0.05
	13588	0.78	-0.03	14642	0.52	0.08
	13614	1.95	-0.03	13585	0.57	-0.12
	13650	1.09	-0.03	13586	0.61	0.13
	13567	-0.34	-0.03	13588	0.78	-0.03
	13574	-0.52	-0.02	13629	0.78	-0.14
	16024	-0.03	-0.02	13631	0.84	-0.11
	13589	1.93	0.01	13603	1.01	0.07
	13656	2.35	0.02	13602	1.09	0.15
	13616	2.67	0.02	13650	1.09	-0.03
	13563	-0.52	0.02	13661	1.22	0.10
	13651	2.53	0.02	13584	1.26	0.21
	13615	2.21	0.03	13659	1.26	0.06
	14640	-0.55	0.04	13962	1.39	-0.12
	14618	2.75	0.04	13657	1.62	-0.13
	13617	3.33	0.04	14491	1.72	-0.13
	13572	-0.68	0.05	14492	1.78	0.11
	13659	1.26	0.06	13630	1.84	-0.19
	13603	1.01	0.07	13604	1.87	-0.16
	14642	0.52	0.08	13660	1.89	0.11
	13564	-0.98	0.09	13589	1.93	0.01
	13576	-0.48	0.09	13963	1.94	-0.15
	13964	2.02	0.09	13614	1.95	-0.03
	14493	2.61	0.09	13964	2.02	0.09

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	Item Difficulty	Displacement	Item ID	Item Difficulty	Displacement
	13661	1.22	0.10	13615	2.21	0.03
	13660	1.89	0.11	13656	2.35	0.02
	14492	1.78	0.11	13652	2.41	0.12
	13652	2.41	0.12	14616	2.52	0.15
	13586	0.61	0.13	13651	2.53	0.02
	13565	-1.25	0.13	14493	2.61	0.09
	13577	-0.09	0.14	13616	2.67	0.02
	13602	1.09	0.15	14618	2.75	0.04
	14616	2.52	0.15	13619	3.03	0.19
	13568	0.15	0.19	13618	3.13	-0.05
	13619	3.03	0.19	13617	3.33	0.04
	13584	1.26	0.21	13658	3.59	-0.05

2.7.2.5 Grades 9–12

Table 2.7.2.5

Equating Summary: Read 9-12 S403 Online

Comparison of Forms	Form 403		Form 402	
	No. of Items	Average Difficulty (Std. Dev.)	No. of Items	Average Difficulty (Std. Dev.)
	72	1.94 (1.34)	72	1.91 (1.33)
	Easiest	Hardest	Easiest	Hardest
	-1.77	4.30	-1.77	4.58
Anchoring Items	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	60	1.71 (1.32)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	58	1.64 (1.29)		
	Percentage Anchors	**Average Displacement**		
	81%	0.00		

Displacement of Anchor Items	Anchor Items by Displacement			Anchor Items by Item Difficulty		
	Item ID	Item Difficulty	Displacement	Item ID	Item Difficulty	Displacement
	13733	2.87	-0.24	13674	-1.77	-0.15
	13693	1.19	-0.22	13767	-0.89	0.19
	16064	1.45	-0.20	13680	-0.46	-0.16
	13689	0.26	-0.20	13677	-0.32	-0.03
	13950	1.41	-0.20	13690	0.10	-0.02
	13676	1.32	-0.18	13675	0.12	0.08
	13680	-0.46	-0.16	13692	0.12	-0.13
	13674	-1.77	-0.15	13678	0.16	0.11
	13744	0.60	-0.14	13743	0.26	0.06
	16072	3.26	-0.14	13689	0.26	-0.20
	13692	0.12	-0.13	16062	0.53	-0.12
	13723	2.82	-0.12	13681	0.55	-0.06
	16062	0.53	-0.12	13744	0.60	-0.14
	16063	0.98	-0.12	16058	0.62	-0.08
	13969	3.47	-0.11	16059	0.67	-0.02

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	16058	0.62	-0.08	13682	0.77	0.12
	13679	0.86	-0.08	13704	0.80	0.21
	13722	2.50	-0.08	13679	0.86	-0.08
	16070	2.65	-0.07	16063	0.98	-0.12
	13769	2.14	-0.06	13745	1.04	0.03
	13681	0.55	-0.06	13694	1.14	0.05
	13968	2.61	-0.04	13693	1.19	-0.22
	13786	4.30	-0.03	13698	1.20	0.21
	13677	-0.32	-0.03	13705	1.29	0.20
	13708	2.16	-0.03	13676	1.32	-0.18
	16059	0.67	-0.02	13950	1.41	-0.20
	13690	0.10	-0.02	16064	1.45	-0.20
	14635	3.35	-0.01	13699	1.48	0.02
	14975	1.87	0.00	13691	1.49	0.11
	13952	1.92	0.01	13707	1.60	0.17
	13731	3.95	0.01	14975	1.87	0.00
	16071	2.36	0.02	13952	1.92	0.01
	13699	1.48	0.02	13706	1.94	0.20
	13951	2.69	0.02	16060	2.00	0.29
	14636	3.11	0.02	13768	2.02	0.11
	13745	1.04	0.03	13700	2.04	0.18
	13724	3.96	0.05	14634	2.09	0.24
	13694	1.14	0.05	13769	2.14	-0.06
	13743	0.26	0.06	13708	2.16	-0.03
	13709	2.72	0.07	16071	2.36	0.02
	13970	2.85	0.07	13722	2.50	-0.08
	13787	3.69	0.07	13785	2.54	0.26
	13675	0.12	0.08	13968	2.61	-0.04
	13678	0.16	0.11	16070	2.65	-0.07
	13768	2.02	0.11	13951	2.69	0.02

Displacement of Anchor Items	**Anchor Items by Displacement**			**Anchor Items by Item Difficulty**		
	Item ID	**Item Difficulty**	**Displacement**	**Item ID**	**Item Difficulty**	**Displacement**
	13691	1.49	0.11	13709	2.72	0.07
	13682	0.77	0.12	13723	2.82	-0.12
	13732	2.92	0.14	13970	2.85	0.07
	13707	1.60	0.17	13733	2.87	-0.24
	13700	2.04	0.18	13732	2.92	0.14
	13767	-0.89	0.19	14636	3.11	0.02
	13706	1.94	0.20	16072	3.26	-0.14
	13705	1.29	0.20	14635	3.35	-0.01
	13698	1.20	0.21	13969	3.47	-0.11
	13704	0.80	0.21	13787	3.69	0.07
	14634	2.09	0.24	13731	3.95	0.01
	13785	2.54	0.26	13724	3.96	0.05
	16060	2.00	0.29	13786	4.30	-0.03

2.7.3 Writing

2.7.3.1 *Grade 1*

Table 2.7.3.1.1

Equating Summary: Writ 1 A S403 Online

Comparison of Forms*	**Form 403**			**Form 402**		
	No. of Tasks		**Average Difficulty (Std. Dev.)**	**No. of Tasks**		**Average Difficulty (Std. Dev.)**
	4		-2.17 (3.40)	4		-2.17 (3.40)
	Easiest		**Hardest**	**Easiest**		**Hardest**
	-7.22		-0.07	-7.22		-0.07
Anchoring Tasks	**No. of Possible Anchors**		**Average Difficulty (Std. Dev.)**			
	4		-2.17 (3.40)			
	No. of Anchors Used		**Average Difficulty (Std. Dev.)**			
	4		-2.17 (3.40)			
	Percentage Anchors		**Average Displacement**			
	100%		-0.06			
Rating Scale Step Measures by Task	**Anchored Scale Steps**					
	Task	**Step**	**Measure**			
	Task 2	1	-3.80			
		2	3.89			
		3	-0.09			
	Task 3/Task 4	1	-2.47			
		2	-2.78			
		3	-2.61			
		4	-1.68			
		5	-0.48			
		6	0.97			
		7	2.25			
		8	3.21			
		9	3.59			
Displacement of Anchor Tasks	**Anchor Tasks by Displacement**			**Anchor Tasks by Task Difficulty**		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	14775	-1.11	-0.27	14245	-7.22	0.07
	15919	-0.07	-0.06	14775	-1.11	-0.27
	14248	-0.27	0.03	14248	-0.27	0.03
	14245	-7.22	0.07	15919	-0.07	-0.06

Table 2.7.3.1.2

Equating Summary: Writ 1 B/C S403 Online

Comparison of Forms*	**Form 403**		**Form 402**	
	No. of Tasks	**Average Difficulty (Std. Dev.)**	**No. of Tasks**	**Average Difficulty (Std. Dev.)**
	3	-0.10 (0.45)	3	0.12 (0.47)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	-0.60	0.25	-0.40	0.52
Anchoring Tasks	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	1	0.25 (N/A)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	1	0.25 (N/A)		
	Percentage Anchors	**Average Displacement**		
	33%	0.00		
Common Rating Scale Step Measures	**Anchored Scale Steps**			
	Step	**Measure**		
	1	-2.47		
	2	-2.78		
	3	-2.61		
	4	-1.68		
	5	-0.48		
	6	0.97		
	7	2.25		
	8	3.21		
	9	3.59		

Displacement of Anchor Tasks	**Anchor Tasks by Displacement**			**Anchor Tasks by Task Difficulty**		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	14960	0.25	0.00	14960	0.25	0.00

2.7.3.2 *Grades 2–3*

Table 2.7.3.2.1

Equating Summary: Writ 2-3 A S403 Online

<table>
<tr><th rowspan="6">Comparison of Forms*</th><th colspan="3">Form 403</th><th colspan="3">Form 402</th></tr>
<tr><th>No. of Tasks</th><th colspan="2">Average Difficulty (Std. Dev.)</th><th>No. of Tasks</th><th colspan="2">Average Difficulty (Std. Dev.)</th></tr>
<tr><td>3</td><td colspan="2">0.14 (0.13)</td><td>3</td><td colspan="2">-0.02 (0.26)</td></tr>
<tr><th>Easiest</th><th colspan="2">Hardest</th><th>Easiest</th><th colspan="2">Hardest</th></tr>
<tr><td>0.01</td><td colspan="2">0.27</td><td>-0.23</td><td colspan="2">0.27</td></tr>
<tr></tr>
<tr><th rowspan="6">Anchoring Tasks</th><th>No. of Possible Anchors</th><th colspan="2">Average Difficulty (Std. Dev.)</th><td colspan="3" rowspan="6"></td></tr>
<tr><td>1</td><td colspan="2">0.27 (N/A)</td></tr>
<tr><th>No. of Anchors Used</th><th colspan="2">Average Difficulty (Std. Dev.)</th></tr>
<tr><td>1</td><td colspan="2">0.27 (N/A)</td></tr>
<tr><th>Percentage Anchors</th><th colspan="2">Average Displacement</th></tr>
<tr><td>33%</td><td colspan="2">0.00</td></tr>
<tr><th rowspan="11">Common Rating Scale Step Measures</th><th colspan="3">Anchored Scale Steps</th><td colspan="3" rowspan="11"></td></tr>
<tr><th>Step</th><th colspan="2">Measure</th></tr>
<tr><td>1</td><td colspan="2">-2.47</td></tr>
<tr><td>2</td><td colspan="2">-2.78</td></tr>
<tr><td>3</td><td colspan="2">-2.61</td></tr>
<tr><td>4</td><td colspan="2">-1.68</td></tr>
<tr><td>5</td><td colspan="2">-0.48</td></tr>
<tr><td>6</td><td colspan="2">0.97</td></tr>
<tr><td>7</td><td colspan="2">2.25</td></tr>
<tr><td>8</td><td colspan="2">3.21</td></tr>
<tr><td>9</td><td colspan="2">3.59</td></tr>
<tr><th rowspan="3">Displacement of Anchor Tasks</th><th colspan="3">Anchor Tasks by Displacement</th><th colspan="3">Anchor Tasks by Task Difficulty</th></tr>
<tr><th>Task ID</th><th>Task Difficulty</th><th>Displacement</th><th>Task ID</th><th>Task Difficulty</th><th>Displacement</th></tr>
<tr><td>14792</td><td>0.27</td><td>0.00</td><td>14792</td><td>0.27</td><td>0.00</td></tr>
</table>

Table 2.7.3.2.2

Equating Summary: Writ 2-3 B/C S403 Online

<table>
<tr><th rowspan="6">Comparison of Forms*</th><th colspan="6">Form 403</th><th colspan="6">Form 402</th></tr>
<tr><th colspan="3">No. of Tasks</th><th colspan="3">Average Difficulty (Std. Dev.)</th><th colspan="3">No. of Tasks</th><th colspan="3">Average Difficulty (Std. Dev.)</th></tr>
<tr><td colspan="3">3</td><td colspan="3">0.86 (0.07)</td><td colspan="3">3</td><td colspan="3">1.02 (0.14)</td></tr>
<tr><th colspan="3">Easiest</th><th colspan="3">Hardest</th><th colspan="3">Easiest</th><th colspan="3">Hardest</th></tr>
<tr><td colspan="3">0.80</td><td colspan="3">0.94</td><td colspan="3">0.94</td><td colspan="3">1.19</td></tr>
<tr></tr>
<tr><th rowspan="6">Anchoring Tasks</th><th colspan="3">No. of Possible Anchors</th><th colspan="3">Average Difficulty (Std. Dev.)</th><td colspan="6" rowspan="6"></td></tr>
<tr><td colspan="3">1</td><td colspan="3">0.94 (N/A)</td></tr>
<tr><th colspan="3">No. of Anchors Used</th><th colspan="3">Average Difficulty (Std. Dev.)</th></tr>
<tr><td colspan="3">1</td><td colspan="3">0.94 (N/A)</td></tr>
<tr><th colspan="3">Percentage Anchors</th><th colspan="3">Average Displacement</th></tr>
<tr><td colspan="3">33%</td><td colspan="3">0.00</td></tr>
<tr><th rowspan="11">Common Rating Scale Step Measures</th><th colspan="6">Anchored Scale Steps</th><td colspan="6" rowspan="11"></td></tr>
<tr><th colspan="3">Step</th><th colspan="3">Measure</th></tr>
<tr><td colspan="3">1</td><td colspan="3">-2.47</td></tr>
<tr><td colspan="3">2</td><td colspan="3">-2.78</td></tr>
<tr><td colspan="3">3</td><td colspan="3">-2.61</td></tr>
<tr><td colspan="3">4</td><td colspan="3">-1.68</td></tr>
<tr><td colspan="3">5</td><td colspan="3">-0.48</td></tr>
<tr><td colspan="3">6</td><td colspan="3">0.97</td></tr>
<tr><td colspan="3">7</td><td colspan="3">2.25</td></tr>
<tr><td colspan="3">8</td><td colspan="3">3.21</td></tr>
<tr><td colspan="3">9</td><td colspan="3">3.59</td></tr>
<tr><th rowspan="3">Displacement of Anchor Tasks</th><th colspan="6">Anchor Tasks by Displacement</th><th colspan="6">Anchor Tasks by Task Difficulty</th></tr>
<tr><th colspan="2">Task ID</th><th colspan="2">Task Difficulty</th><th colspan="2">Displacement</th><th colspan="2">Task ID</th><th colspan="2">Task Difficulty</th><th colspan="2">Displacement</th></tr>
<tr><td colspan="2">15549</td><td colspan="2">0.94</td><td colspan="2">0.00</td><td colspan="2">15549</td><td colspan="2">0.94</td><td colspan="2">0.00</td></tr>
</table>

2.7.3.3 *Grades 4–5*

Table 2.7.3.3.1

Equating Summary: Writ 4-5 A S403 Online

Comparison of Forms*	Form 403		Form 402	
	No. of Tasks	**Average Difficulty (Std. Dev.)**	**No. of Tasks**	**Average Difficulty (Std. Dev.)**
	3	1.18 (0.15)	3	1.31 (0.12)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	1.05	1.34	1.19	1.41
Anchoring Tasks	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	1	1.34 (N/A)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	1	1.34 (N/A)		
	Percentage Anchors	**Average Displacement**		
	33%	0.00		
Common Rating Scale Step Measures	**Anchored Scale Steps**			
	Step	**Measure**		
	1	-2.47		
	2	-2.78		
	3	-2.61		
	4	-1.68		
	5	-0.48		
	6	0.97		
	7	2.25		
	8	3.21		
	9	3.59		

Displacement of Anchor Tasks	Anchor Tasks by Displacement			Anchor Tasks by Task Difficulty		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	15397_17212	1.34	0.00	15397_17212	1.34	0.00

Table 2.7.3.3.2

Equating Summary: Writ 4-5 B/C S403 Online

Comparison of Forms*	Form 403		Form 402	
	No. of Tasks	**Average Difficulty (Std. Dev.)**	**No. of Tasks**	**Average Difficulty (Std. Dev.)**
	3	1.85 (0.41)	3	1.81 (0.27)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	1.39	2.17	1.51	2.00
Anchoring Tasks	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	1	2.00 (N/A)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	1	2.00 (N/A)		
	Percentage Anchors	**Average Displacement**		
	33%	0.00		
Common Rating Scale Step Measures	**Anchored Scale Steps**			
	Step	**Measure**		
	1	-2.47		
	2	-2.78		
	3	-2.61		
	4	-1.68		
	5	-0.48		
	6	0.97		
	7	2.25		
	8	3.21		
	9	3.59		

Displacement of Anchor Tasks	Anchor Tasks by Displacement			Anchor Tasks by Task Difficulty		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	15416_17207	2.00	0.00	15416_17207	2.00	0.00

2.7.3.4 *Grades 6–8*

Table 2.7.3.4.1

Equating Summary: Writ 6-8 A S403 Online

Comparison of Forms*	**Form 403**		**Form 402**	
	No. of Tasks	**Average Difficulty (Std. Dev.)**	**No. of Tasks**	**Average Difficulty (Std. Dev.)**
	3	0.93 (0.47)	3	0.97 (0.40)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	0.48	1.41	0.64	1.41
Anchoring Tasks	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	1	1.41 (N/A)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	1	1.41 (N/A)		
	Percentage Anchors	**Average Displacement**		
	33%	0.00		
Common Rating Scale Step Measures	**Anchored Scale Steps**			
	Step	**Measure**		
	1	-2.47		
	2	-2.78		
	3	-2.61		
	4	-1.68		
	5	-0.48		
	6	0.97		
	7	2.25		
	8	3.21		
	9	3.59		

Displacement of Anchor Tasks	**Anchor Tasks by Displacement**			**Anchor Tasks by Task Difficulty**		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	14786_17188	1.41	0.00	14786_17188	1.41	0.00

Table 2.7.3.4.2

Equating Summary: Writ 6-8 B/C S403 Online

<table>
<tr><td rowspan="5">Comparison of Forms*</td><td colspan="3">Form 403</td><td colspan="3">Form 402</td></tr>
<tr><td colspan="2">No. of Tasks</td><td>Average Difficulty (Std. Dev.)</td><td colspan="2">No. of Tasks</td><td>Average Difficulty (Std. Dev.)</td></tr>
<tr><td colspan="2">3</td><td>1.40 (0.23)</td><td colspan="2">3</td><td>1.30 (0.20)</td></tr>
<tr><td colspan="2">Easiest</td><td>Hardest</td><td colspan="2">Easiest</td><td>Hardest</td></tr>
<tr><td colspan="2">1.13</td><td>1.56</td><td colspan="2">1.10</td><td>1.49</td></tr>
<tr><td rowspan="6">Anchoring Tasks</td><td colspan="2">No. of Possible Anchors</td><td>Average Difficulty (Std. Dev.)</td><td colspan="3"></td></tr>
<tr><td colspan="2">1</td><td>1.49 (N/A)</td><td colspan="3"></td></tr>
<tr><td colspan="2">No. of Anchors Used</td><td>Average Difficulty (Std. Dev.)</td><td colspan="3"></td></tr>
<tr><td colspan="2">1</td><td>1.49 (N/A)</td><td colspan="3"></td></tr>
<tr><td colspan="2">Percentage Anchors</td><td>Average Displacement</td><td colspan="3"></td></tr>
<tr><td colspan="2">33%</td><td>0.00</td><td colspan="3"></td></tr>
<tr><td rowspan="11">Common Rating Scale Step Measures</td><td colspan="3">Anchored Scale Steps</td><td colspan="3"></td></tr>
<tr><td colspan="2">Step</td><td>Measure</td><td colspan="3"></td></tr>
<tr><td colspan="2">1</td><td>-2.47</td><td colspan="3"></td></tr>
<tr><td colspan="2">2</td><td>-2.78</td><td colspan="3"></td></tr>
<tr><td colspan="2">3</td><td>-2.61</td><td colspan="3"></td></tr>
<tr><td colspan="2">4</td><td>-1.68</td><td colspan="3"></td></tr>
<tr><td colspan="2">5</td><td>-0.48</td><td colspan="3"></td></tr>
<tr><td colspan="2">6</td><td>0.97</td><td colspan="3"></td></tr>
<tr><td colspan="2">7</td><td>2.25</td><td colspan="3"></td></tr>
<tr><td colspan="2">8</td><td>3.21</td><td colspan="3"></td></tr>
<tr><td colspan="2">9</td><td>3.59</td><td colspan="3"></td></tr>
<tr><td rowspan="3">Displacement of Anchor Tasks</td><td colspan="3">Anchor Tasks by Displacement</td><td colspan="3">Anchor Tasks by Task Difficulty</td></tr>
<tr><td>Task ID</td><td>Task Difficulty</td><td>Displacement</td><td>Task ID</td><td>Task Difficulty</td><td>Displacement</td></tr>
<tr><td>15418_17176</td><td>1.49</td><td>0.00</td><td>15418_17176</td><td>1.49</td><td>0.00</td></tr>
</table>

2.7.3.5 *Grades 9–12*

Table 2.7.3.5.1

Equating Summary: Writ 9-12 A S403 Online

Comparison of Forms*	Form 403		Form 402	
	No. of Tasks	**Average Difficulty (Std. Dev.)**	**No. of Tasks**	**Average Difficulty (Std. Dev.)**
	3	2.00 (0.31)	3	2.07 (0.36)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	1.64	2.19	1.68	2.36
Anchoring Tasks	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	1	2.17 (N/A)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	1	2.17 (N/A)		
	Percentage Anchors	**Average Displacement**		
	33%	0.00		
Common Rating Scale Step Measures	**Anchored Scale Steps**			
	Step	**Measure**		
	1	-2.47		
	2	-2.78		
	3	-2.61		
	4	-1.68		
	5	-0.48		
	6	0.97		
	7	2.25		
	8	3.21		
	9	3.59		

Displacement of Anchor Tasks	Anchor Tasks by Displacement			Anchor Tasks by Task Difficulty		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	15555_17166	2.17	0.00	15555_17166	2.17	0.00

Table 2.7.3.5.2

Equating Summary: Writ 9-12 B/C S403 Online

<table>
<tr><td rowspan="5">Comparison of Forms*</td><td colspan="3">Form 403</td><td colspan="3">Form 402</td></tr>
<tr><td>No. of Tasks</td><td colspan="2">Average Difficulty (Std. Dev.)</td><td>No. of Tasks</td><td colspan="2">Average Difficulty (Std. Dev.)</td></tr>
<tr><td>3</td><td colspan="2">1.85 (0.45)</td><td>3</td><td colspan="2">2.03 (0.34)</td></tr>
<tr><td>Easiest</td><td colspan="2">Hardest</td><td>Easiest</td><td colspan="2">Hardest</td></tr>
<tr><td>1.42</td><td colspan="2">2.32</td><td>1.83</td><td colspan="2">2.43</td></tr>
<tr><td rowspan="6">Anchoring Tasks</td><td>No. of Possible Anchors</td><td colspan="2">Average Difficulty (Std. Dev.)</td><td colspan="3" rowspan="6"></td></tr>
<tr><td>1</td><td colspan="2">1.83 (N/A)</td></tr>
<tr><td>No. of Anchors Used</td><td colspan="2">Average Difficulty (Std. Dev.)</td></tr>
<tr><td>1</td><td colspan="2">1.83 (N/A)</td></tr>
<tr><td>Percentage Anchors</td><td colspan="2">Average Displacement</td></tr>
<tr><td>33%</td><td colspan="2">0.00</td></tr>
<tr><td rowspan="11">Common Rating Scale Step Measures</td><td colspan="3">Anchored Scale Steps</td><td colspan="3" rowspan="11"></td></tr>
<tr><td>Step</td><td colspan="2">Measure</td></tr>
<tr><td>1</td><td colspan="2">-2.47</td></tr>
<tr><td>2</td><td colspan="2">-2.78</td></tr>
<tr><td>3</td><td colspan="2">-2.61</td></tr>
<tr><td>4</td><td colspan="2">-1.68</td></tr>
<tr><td>5</td><td colspan="2">-0.48</td></tr>
<tr><td>6</td><td colspan="2">0.97</td></tr>
<tr><td>7</td><td colspan="2">2.25</td></tr>
<tr><td>8</td><td colspan="2">3.21</td></tr>
<tr><td>9</td><td colspan="2">3.59</td></tr>
<tr><td rowspan="3">Displacement of Anchor Tasks</td><td colspan="3">Anchor Tasks by Displacement</td><td colspan="3">Anchor Tasks by Task Difficulty</td></tr>
<tr><td>Task ID</td><td>Task Difficulty</td><td>Displacement</td><td>Task ID</td><td>Task Difficulty</td><td>Displacement</td></tr>
<tr><td>14281_14652</td><td>1.83</td><td>0.00</td><td>14281_14652</td><td>1.83</td><td>0.00</td></tr>
</table>

2.7.4 Speaking

2.7.4.1 *Grade 1*

Table 2.7.4.1

Equating Summary: Spek 1 S403 Online

Comparison of Forms	**Form 403**			**Form 402**		
	No. of Tasks	**Average Difficulty (Std. Dev.)**		**No. of Tasks**	**Average Difficulty (Std. Dev.)**	
	9	-1.55 (2.16)		15	-1.30 (2.13)	
	Easiest	**Hardest**		**Easiest**	**Hardest**	
	-4.62	0.27		-4.43	0.35	
Anchoring Tasks	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**				
	9	-1.55 (2.16)				
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**				
	9	-1.55 (2.16)				
	Percentage Anchors	**Average Displacement**				
	100%	-0.02				
Rating Scale Step Measures by Task	**Anchored Scale Steps**					
	Task	**Step**	**Measure**			
	PL 1 Tasks	1	0.56			
		2	-0.56			
	PL 3/PL 5 Tasks	1	-2.65			
		2	-1.80			
		3	1.46			
		4	2.98			
Displacement of Anchor Tasks	**Anchor Tasks by Displacement**			**Anchor Tasks by Task Difficulty**		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	15011	-0.07	-0.33	15004	-4.62	0.00
	16137	-4.20	-0.26	16140	-4.42	-0.03
	16268	-0.20	-0.18	16137	-4.20	-0.26
	16140	-4.42	-0.03	16149	-0.52	0.29
	15004	-4.62	0.00	16262	-0.30	0.02
	16262	-0.30	0.02	16268	-0.20	-0.18
	15019	0.13	0.03	15011	-0.07	-0.33
	16158	0.27	0.25	15019	0.13	0.03
	16149	-0.52	0.29	16158	0.27	0.25

2.7.4.2 *Grades 2–3*

Table 2.7.4.2

Equating Summary: Spek 2-3 S403 Online

Comparison of Forms	**Form 403**			**Form 402**		
	No. of Tasks	**Average Difficulty (Std. Dev.)**		**No. of Tasks**	**Average Difficulty (Std. Dev.)**	
	9	-1.66 (2.33)		15	-1.90 (2.42)	
	Easiest	**Hardest**		**Easiest**	**Hardest**	
	-4.75	0.66		-5.24	0.66	
Anchoring Tasks	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**				
	9	-1.66 (2.33)				
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**				
	9	-1.66 (2.33)				
	Percentage Anchors	**Average Displacement**				
	100%	0.02				
Rating Scale Step Measures by Task	**Anchored Scale Steps**					
	Task	**Step**	**Measure**			
	PL 1 Tasks	1	0.56			
		2	-0.56			
	PL 3/PL 5 Tasks	1	-2.65			
		2	-1.80			
		3	1.46			
		4	2.98			
Displacement of Anchor Tasks	**Anchor Tasks by Displacement**			**Anchor Tasks by Task Difficulty**		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	15112	-0.24	-0.38	15106	-4.75	0.09
	15102	0.66	-0.20	16207	-4.71	0.10
	15092	-0.31	-0.12	15085	-4.60	0.00
	15085	-4.60	0.00	15119	-1.16	0.04
	15119	-1.16	0.04	16211	-0.49	0.35
	15106	-4.75	0.09	15092	-0.31	-0.12
	16207	-4.71	0.10	15112	-0.24	-0.38
	16215	0.63	0.34	16215	0.63	0.34
	16211	-0.49	0.35	15102	0.66	-0.20

2.7.4.3 *Grades 4–5*

Table 2.7.4.3

Equating Summary: Spek 4-5 S403 Online

Comparison of Forms	**Form 403**		**Form 402**	
	No. of Tasks	**Average Difficulty (Std. Dev.)**	**No. of Tasks**	**Average Difficulty (Std. Dev.)**
	9	-0.22 (2.71)	15	-0.25 (2.68)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	-4.06	2.22	-4.04	2.20

Anchoring Tasks		
	No. of Possible Anchors	**Average Difficulty (Std. Dev.)**
	9	-0.22 (2.71)
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**
	9	-0.22 (2.71)
	Percentage Anchors	**Average Displacement**
	100%	-0.09

Rating Scale Step Measures by Task	**Anchored Scale Steps**		
	Task	**Step**	**Measure**
	PL 1 Tasks	1	0.56
		2	-0.56
	PL 3/PL 5 Tasks	1	-2.65
		2	-1.80
		3	1.46
		4	2.98

Displacement of Anchor Tasks	**Anchor Tasks by Displacement**			**Anchor Tasks by Task Difficulty**		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	16301	-3.47	-0.47	16174	-4.06	0.00
	15077	-3.76	-0.38	15077	-3.76	-0.38
	16183	2.22	-0.04	16301	-3.47	-0.47
	16321	2.15	-0.04	15174	0.74	-0.01
	16179	0.82	-0.02	16179	0.82	-0.02
	16315	1.89	-0.02	15254	1.52	0.19
	15174	0.74	-0.01	16315	1.89	-0.02
	16174	-4.06	0.00	16321	2.15	-0.04
	15254	1.52	0.19	16183	2.22	-0.04

2.7.4.4 Grades 6–8

Table 2.7.4.4

Equating Summary: Spek 6-8 S403 Online

Comparison of Forms*	**Form 403**			**Form 402**		
	No. of Tasks	**Average Difficulty (Std. Dev.)**		**No. of Tasks**	**Average Difficulty (Std. Dev.)**	
	9	0.24 (2.94)		15	0.33 (2.91)	
	Easiest	**Hardest**		**Easiest**	**Hardest**	
	-3.90	3.02		-3.73	3.02	
Anchoring Tasks	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**				
	9	0.24 (2.94)				
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**				
	9	0.24 (2.94)				
	Percentage Anchors	**Average Displacement**				
	100%	0.01				
Rating Scale Step Measures by Task	**Anchored Scale Steps**					
	Task	**Step**	**Measure**			
	PL 1 Tasks	1	0.56			
		2	-0.56			
	PL 3/PL 5 Tasks	1	-2.65			
		2	-1.80			
		3	1.46			
		4	2.98			
Displacement of Anchor Tasks	**Anchor Tasks by Displacement**			**Anchor Tasks by Task Difficulty**		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	15163	-3.20	-0.43	16218	-3.90	0.10
	15326	1.86	-0.22	15319	-3.73	0.35
	15340	3.02	-0.18	15163	-3.20	-0.43
	15175	1.50	-0.14	15175	1.50	-0.14
	16218	-3.90	0.10	16223	1.51	0.34
	16228	2.72	0.10	15326	1.86	-0.22
	15181	2.37	0.13	15181	2.37	0.13
	16223	1.51	0.34	16228	2.72	0.10
	15319	-3.73	0.35	15340	3.02	-0.18

2.7.4.5 *Grades 9–12*

Table 2.7.4.5

Equating Summary: Spek 9-12 S403 Online

Comparison of Forms*	**Form 403**		**Form 402**	
	No. of Tasks	**Average Difficulty (Std. Dev.)**	**No. of Tasks**	**Average Difficulty (Std. Dev.)**
	9	0.64 (2.59)	15	0.71 (2.48)
	Easiest	**Hardest**	**Easiest**	**Hardest**
	-3.13	2.84	-2.61	2.84
Anchoring Tasks	**No. of Possible Anchors**	**Average Difficulty (Std. Dev.)**		
	9	0.64 (2.59)		
	No. of Anchors Used	**Average Difficulty (Std. Dev.)**		
	9	0.64 (2.59)		
	Percentage Anchors	**Average Displacement**		
	100%	-0.01		

Rating Scale Step Measures by Task	**Anchored Scale Steps**					
	Task	**Step**	**Measure**			
	PL 1 Tasks	1	0.56			
		2	-0.56			
	PL 3/PL 5 Tasks	1	-2.65			
		2	-1.80			
		3	1.46			
		4	2.98			
Displacement of Anchor Tasks	**Anchor Tasks by Displacement**			**Anchor Tasks by Task Difficulty**		
	Task ID	**Task Difficulty**	**Displacement**	**Task ID**	**Task Difficulty**	**Displacement**
	15479	-2.61	-0.32	16231	-3.13	0.43
	16116	2.08	-0.28	16111	-2.63	-0.15
	16111	-2.63	-0.15	15479	-2.61	-0.32
	16122	2.55	-0.13	16236	1.88	0.19
	16242	2.32	-0.07	16116	2.08	-0.28
	15495	2.84	0.08	16242	2.32	-0.07
	15487	2.42	0.15	15487	2.42	0.15
	16236	1.88	0.19	16122	2.55	-0.13

2.8 Test Characteristic Curve

Test characteristic curves graphically show the relationship between the ability measure (in logits) on the horizontal axis and the expected raw score on the vertical axis. Five vertical lines indicate the five cut scores for the highest grade in the cluster for the test form, dividing the figure into six sections for each of the WIDA proficiency levels (PLs 1–6) for the domain being tested. (Note that for Speaking Tier Pre-A or Speaking or Writing Tier A tests, it may not be possible to place into all six language proficiency levels.) 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 raw score points must be earned to be placed into a WIDA language proficiency level.

As the Listening and Reading assessments are multistage adaptive tests, raw scores are not a meaningful aspect of these tests, so no test characteristic curve is presented for these domains.

2.8.1 Listening

The ACCESS Online Listening test is a multistage adaptive assessment. As students do not all take the same set of items in the test, no test characteristic curve is presented.

2.8.2 Reading

The ACCESS Online Reading test is a multistage adaptive assessment. As students do not all take the same set of items in the test, no test characteristic curve is presented.

2.8.3 Writing

2.8.3.1 Grade 1

Figure 2.8.3.1.1
Test Characteristic Curve: Writ 1 A S403 Online

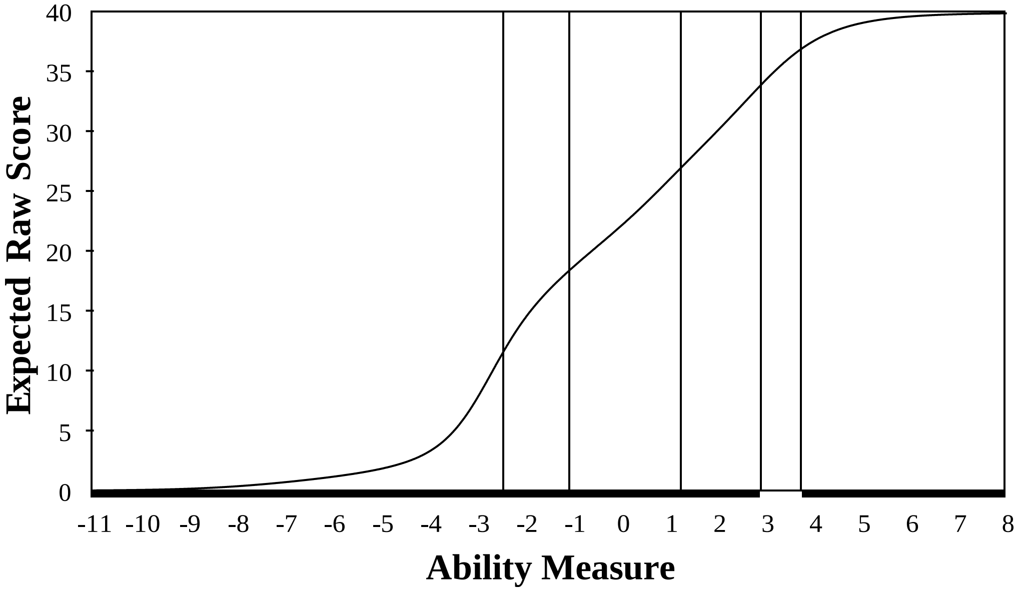


Figure 2.8.3.1.2
Test Characteristic Curve: Writ 1 B/C S403 Online

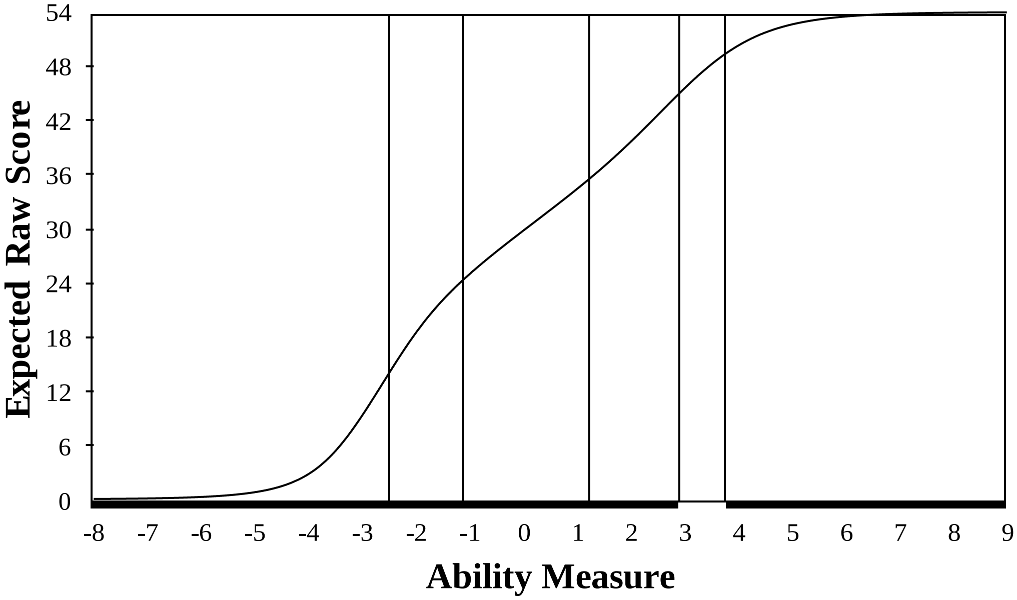
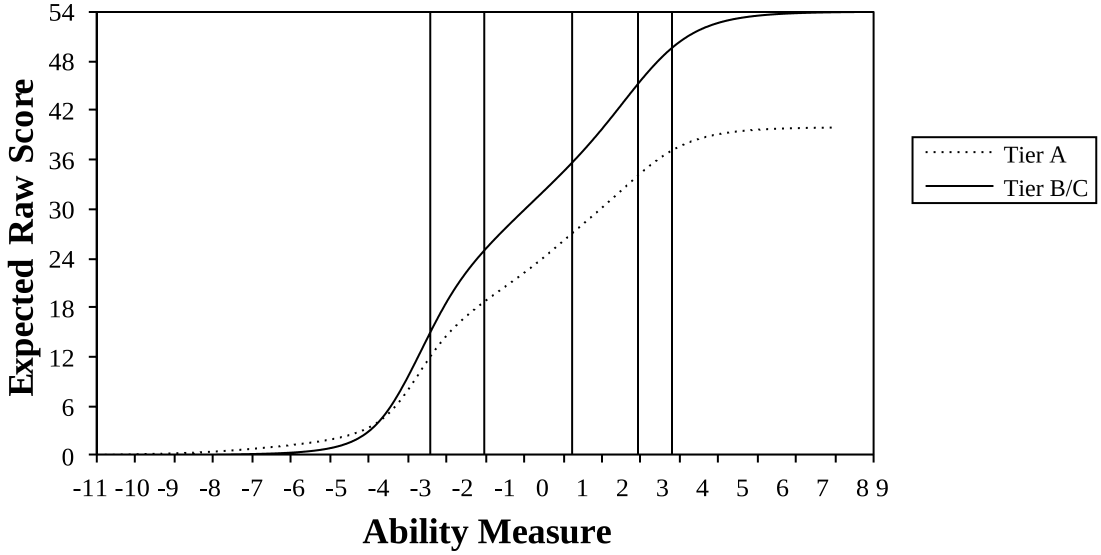


Figure 2.8.3.1.3

Test Characteristic Curve: Writ 1 S403 Online



2.8.3.2 Grades 2–3

Figure 2.8.3.2.1

Test Characteristic Curve: Writ 2-3 A S403 Online

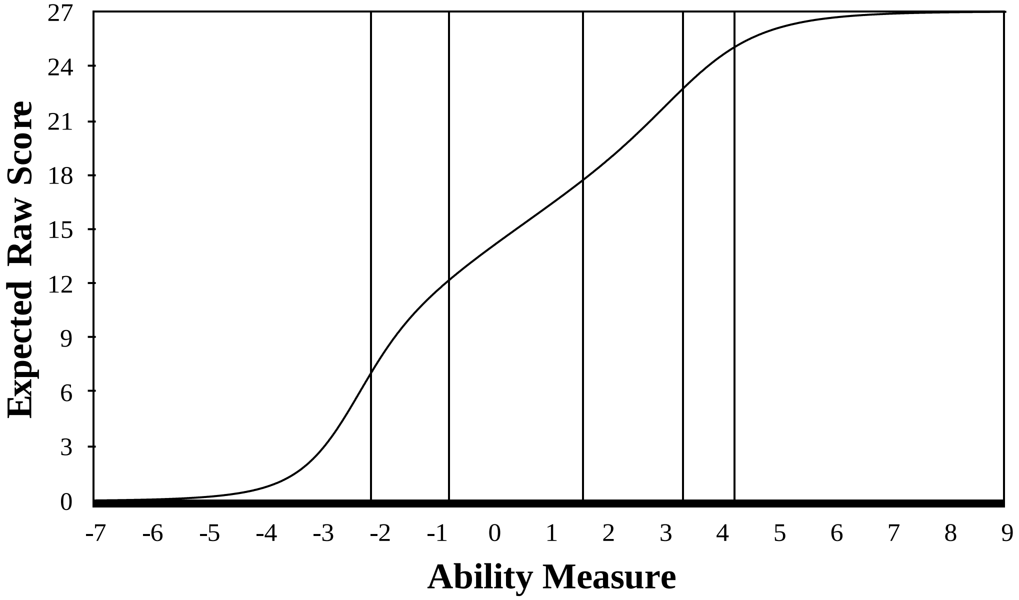


Figure 2.8.3.2.2
Test Characteristic Curve: Writ 2-3 B/C S403 Online

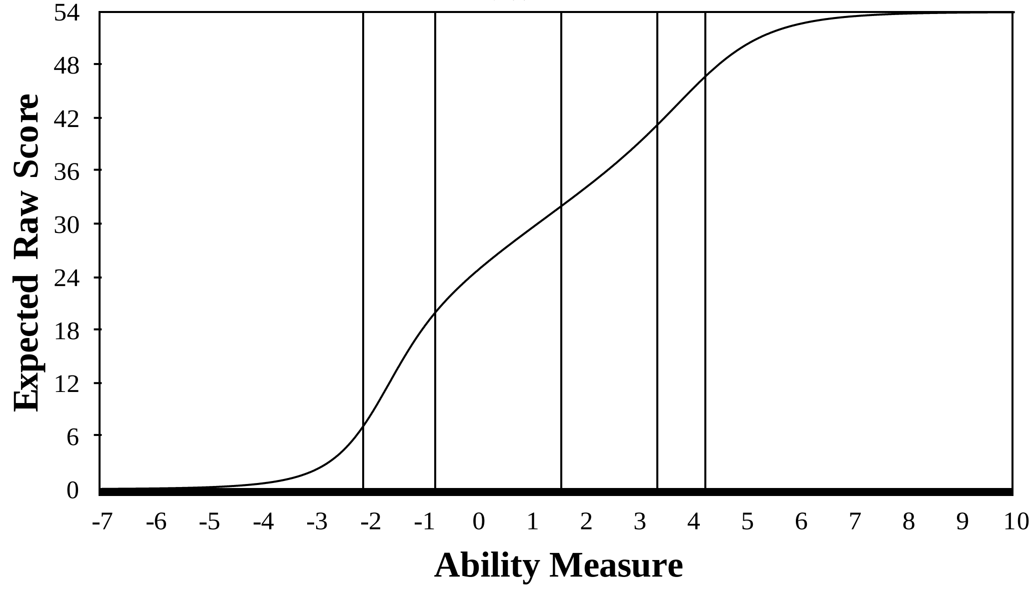
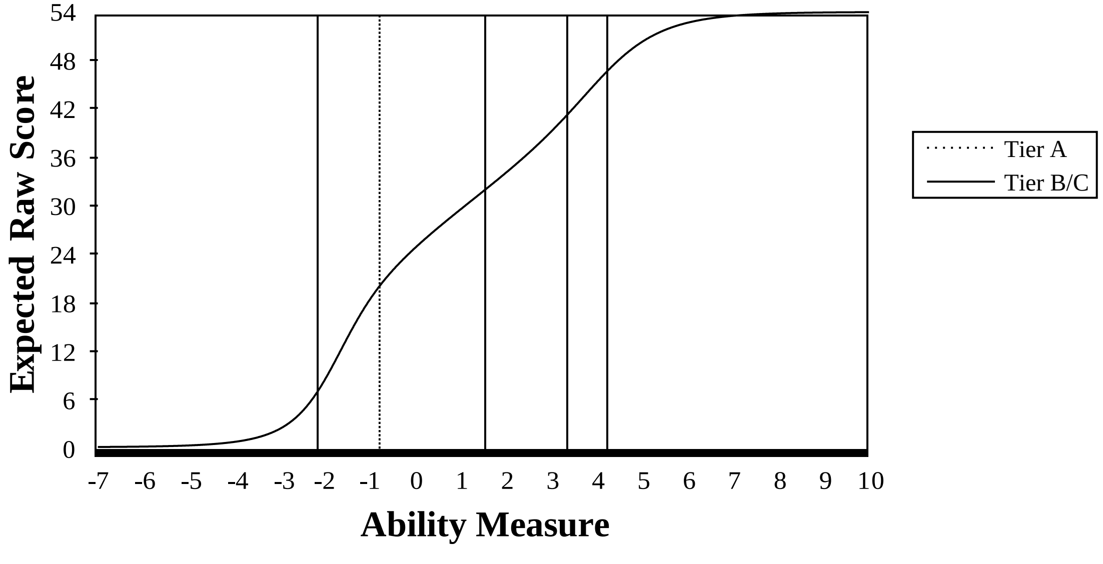


Figure 2.8.3.2.3
Test Characteristic Curve: Writ 2-3 S403 Online



2.8.3.3 *Grades 4–5*

Figure 2.8.3.3.1
Test Characteristic Curve: Writ 4-5 A S403 Online

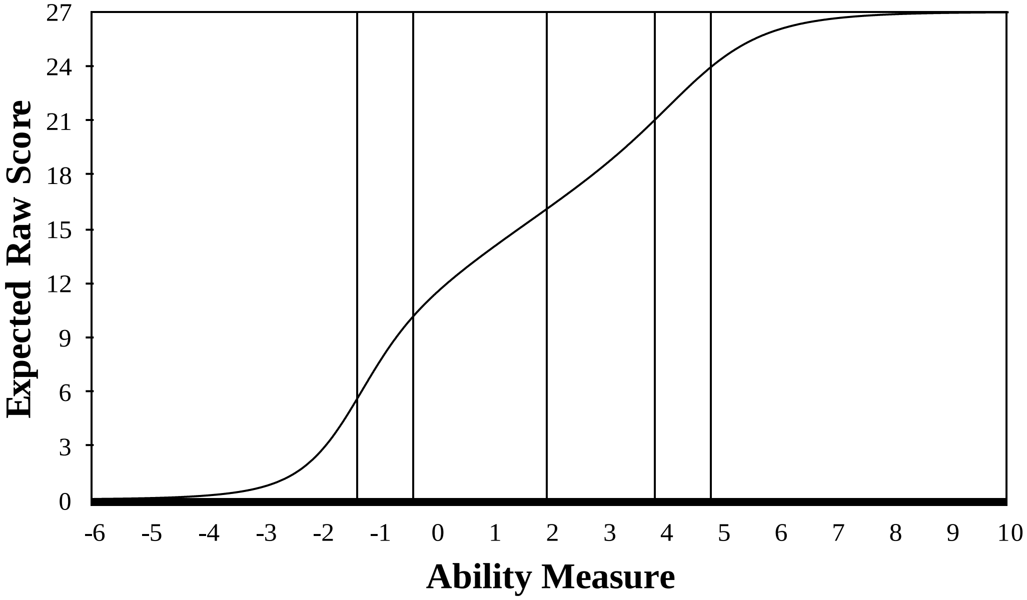


Figure 2.8.3.3.2
Test Characteristic Curve: Writ 4-5 B/C S403 Online

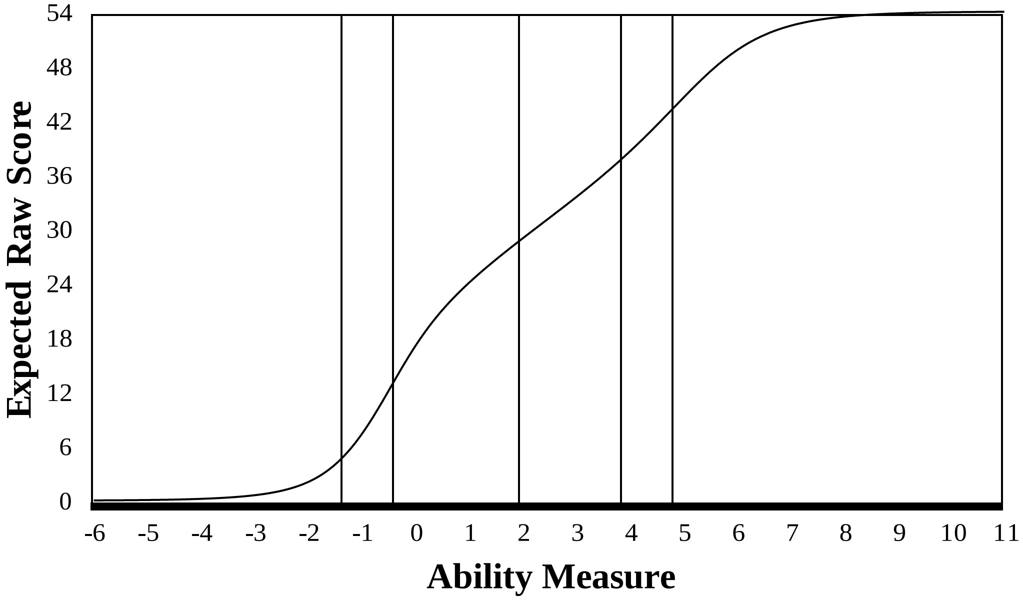
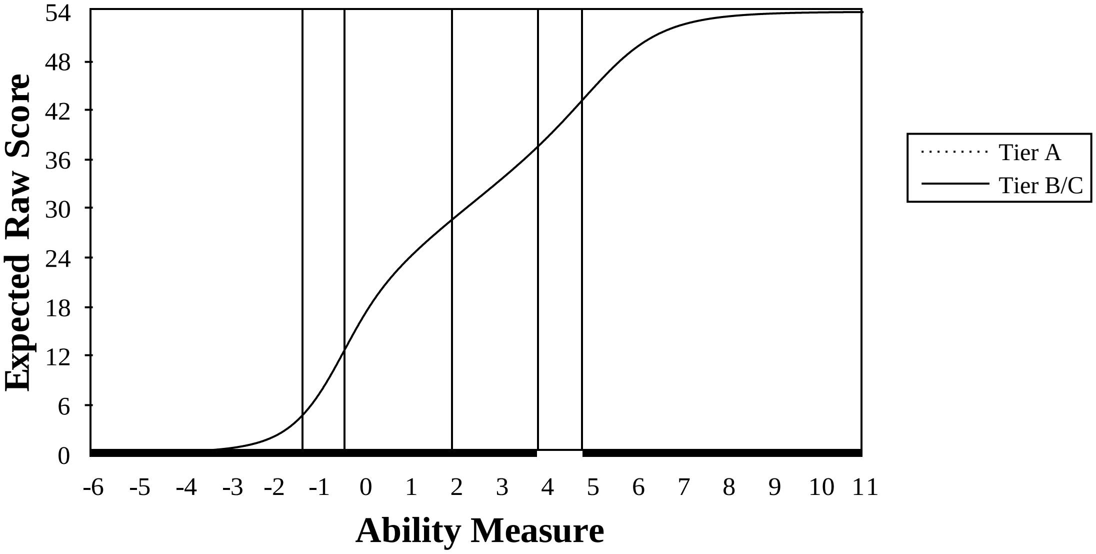


Figure 2.8.3.3.3

Test Characteristic Curve: Writ 4-5 S403 Online



2.8.3.4 *Grades 6–8*

Figure 2.8.3.4.1

Test Characteristic Curve: Writ 6-8 A S403 Online

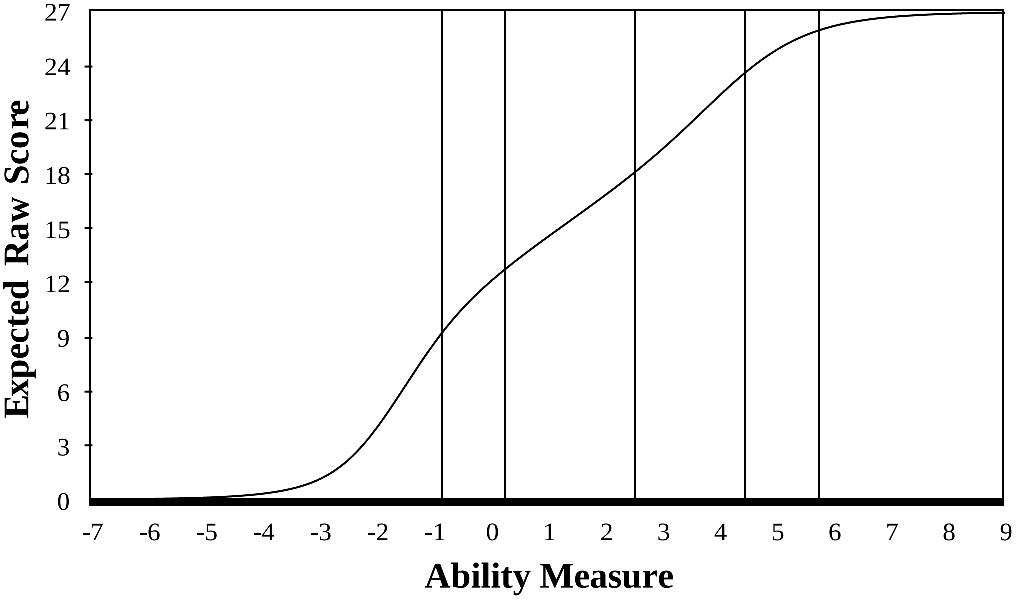


Figure 2.8.3.4.2
Test Characteristic Curve: Writ 6-8 B/C S403 Online

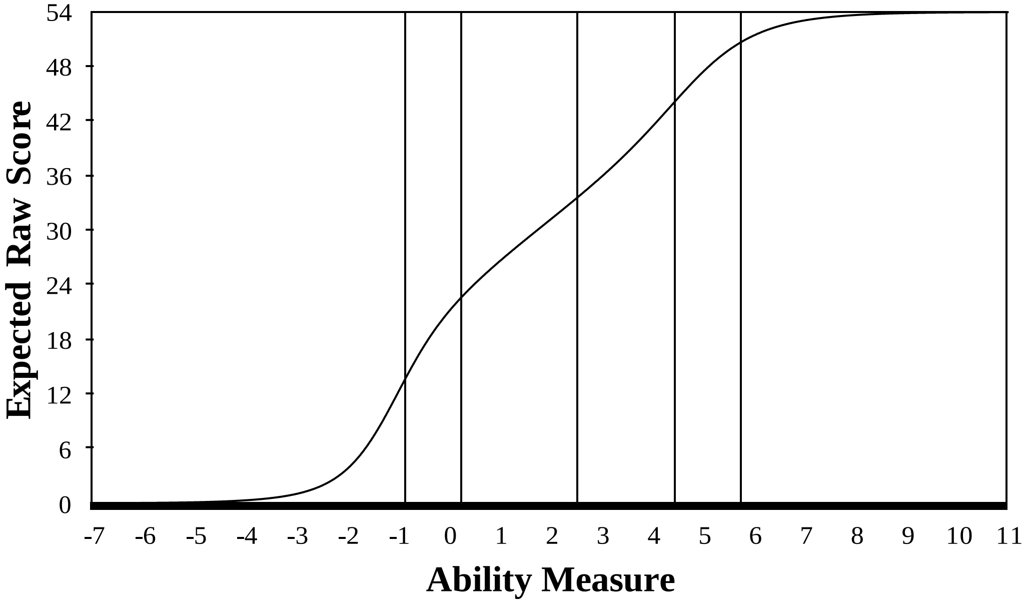
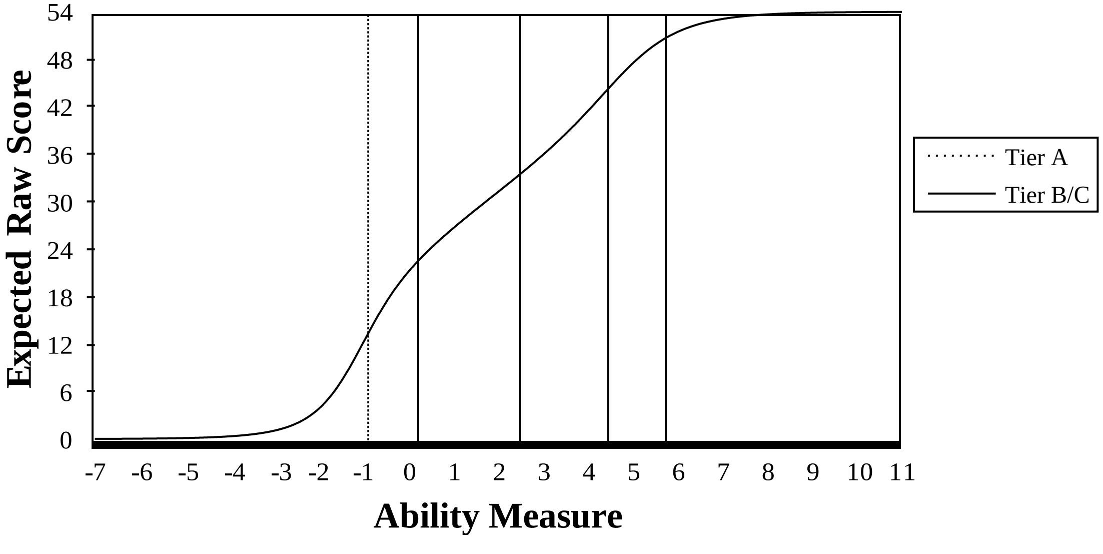


Figure 2.8.3.4.3
Test Characteristic Curve: Writ 6-8 S403 Online



2.8.3.5 *Grades 9–12*

Figure 2.8.3.5.1
Test Characteristic Curve: Writ 9-12 A S403 Online

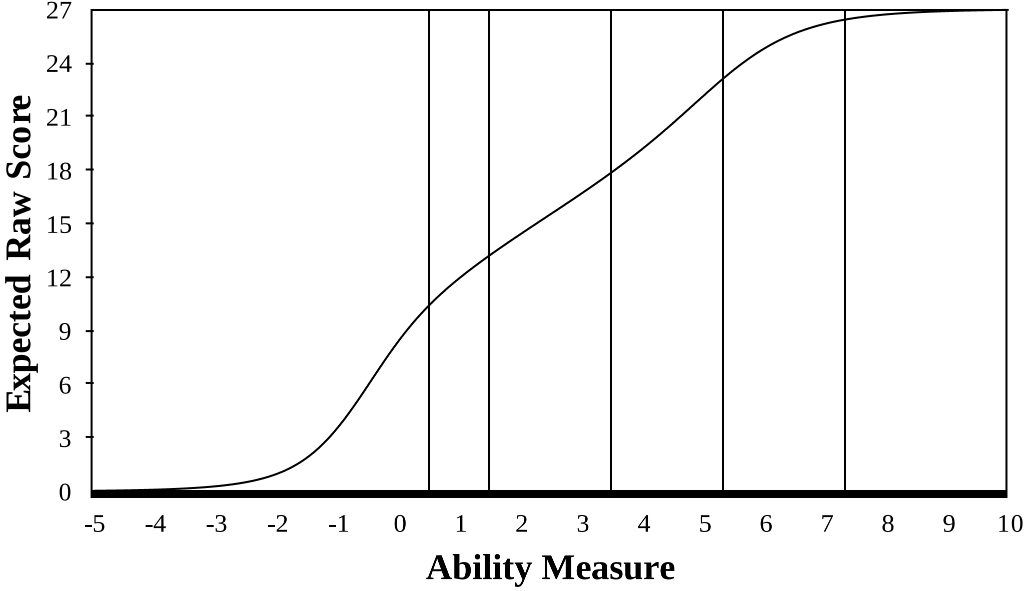


Figure 2.8.3.5.2
Test Characteristic Curve: Writ 9-12 B/C S403 Online

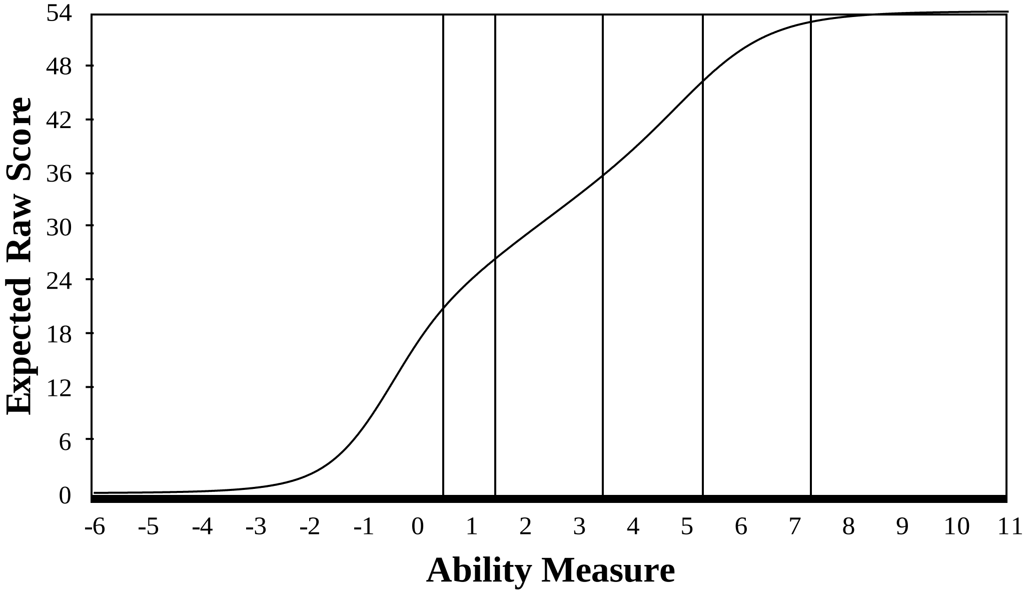
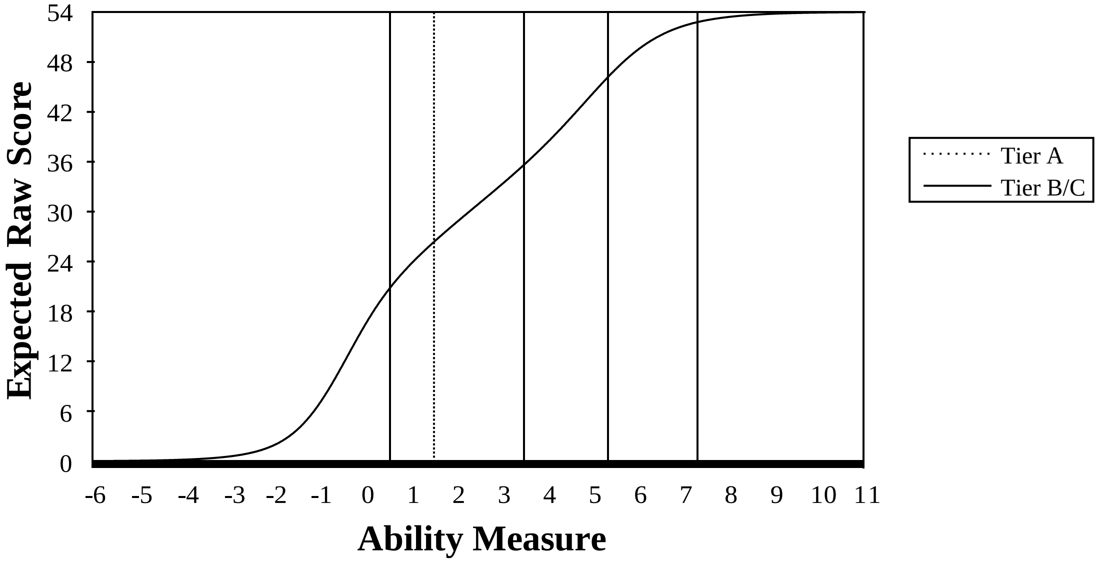


Figure 2.8.3.5.3
Test Characteristic Curve: Writ 9-12 S403 Online



2.8.4 Speaking

2.8.4.1 Grade 1

Figure 2.8.4.1.1

Test Characteristic Curve: Spek 1 Pre-A S403 Online

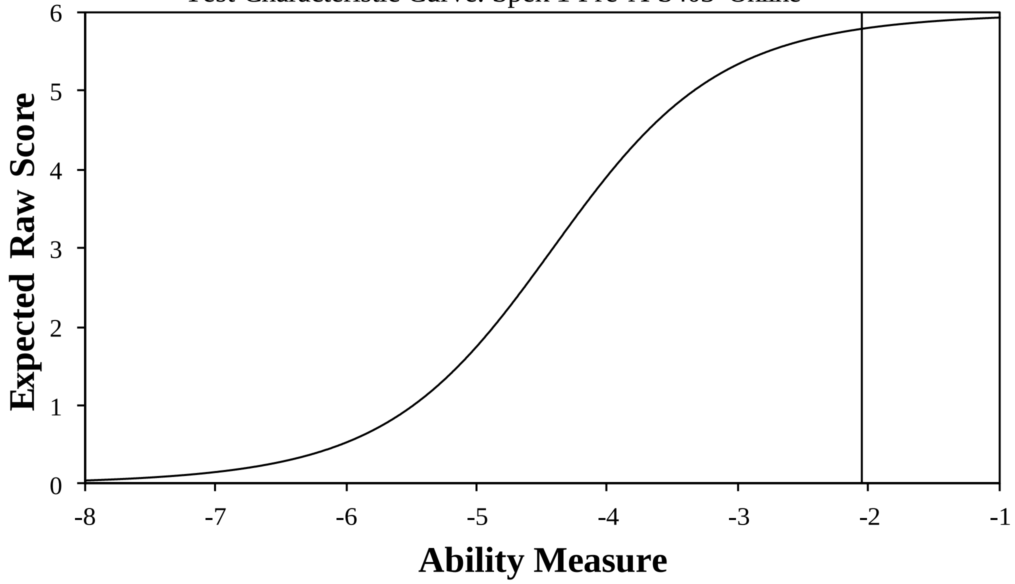


Figure 2.8.4.1.2

Test Characteristic Curve: Spek 1 A S403 Online

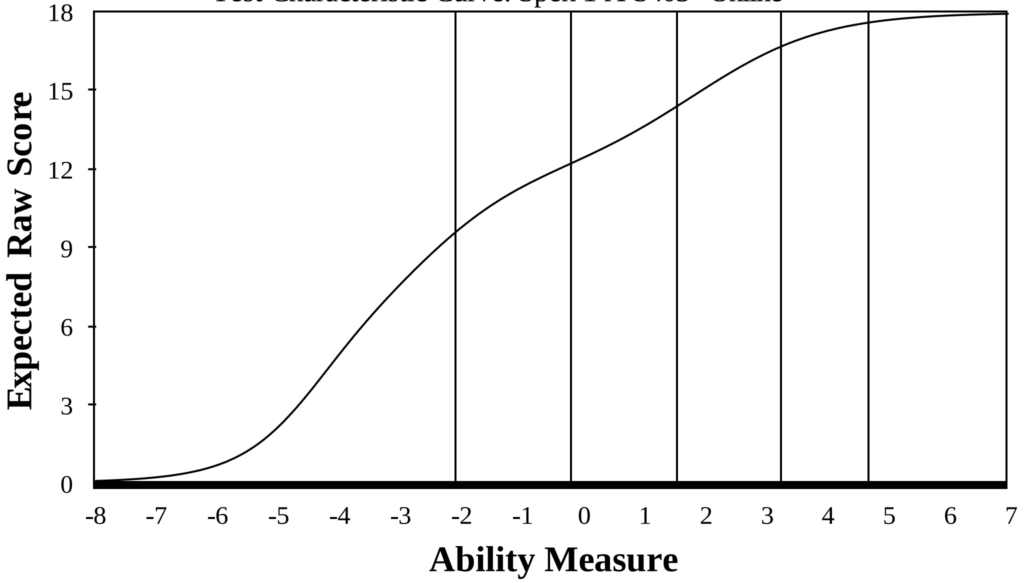


Figure 2.8.4.1.3
Test Characteristic Curve: Spek 1 B/C S403 Online

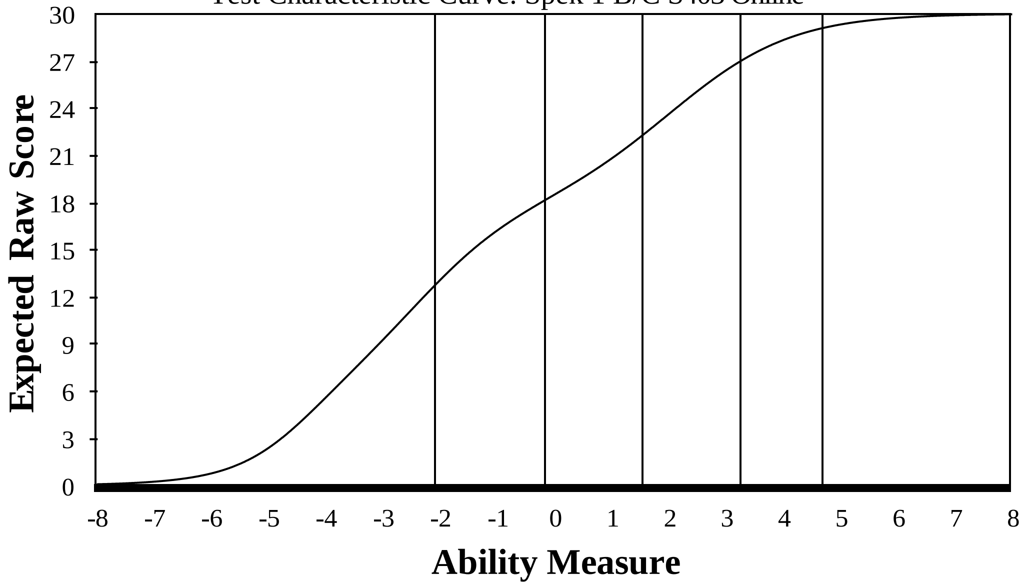
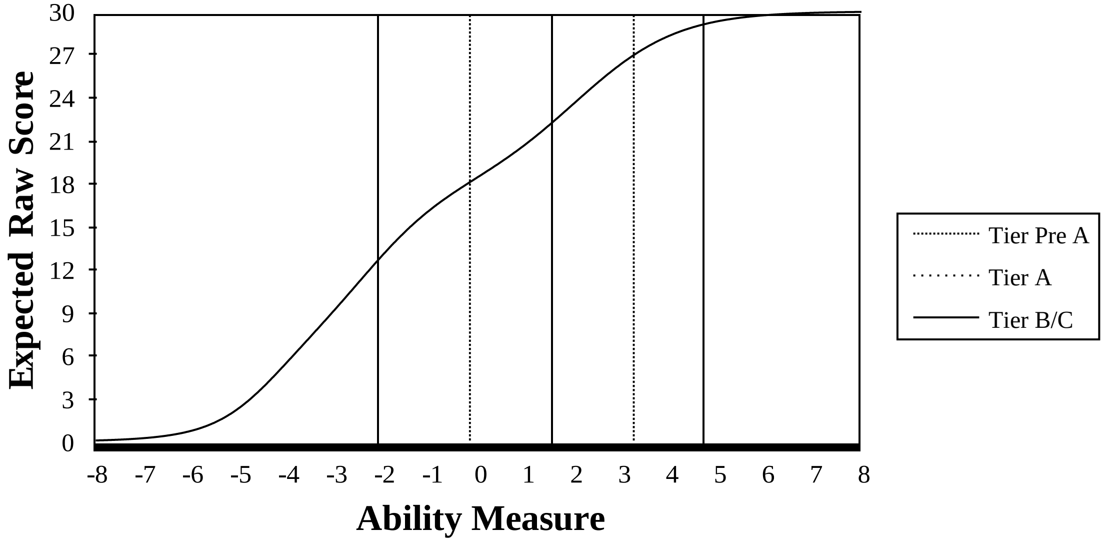


Figure 2.8.4.1.4
Test Characteristic Curve: Spek 1 S403 Online



2.8.4.2 *Grades 2–3*

Figure 2.8.4.2.1
Test Characteristic Curve: Spek 2-3 Pre-A S403 Online

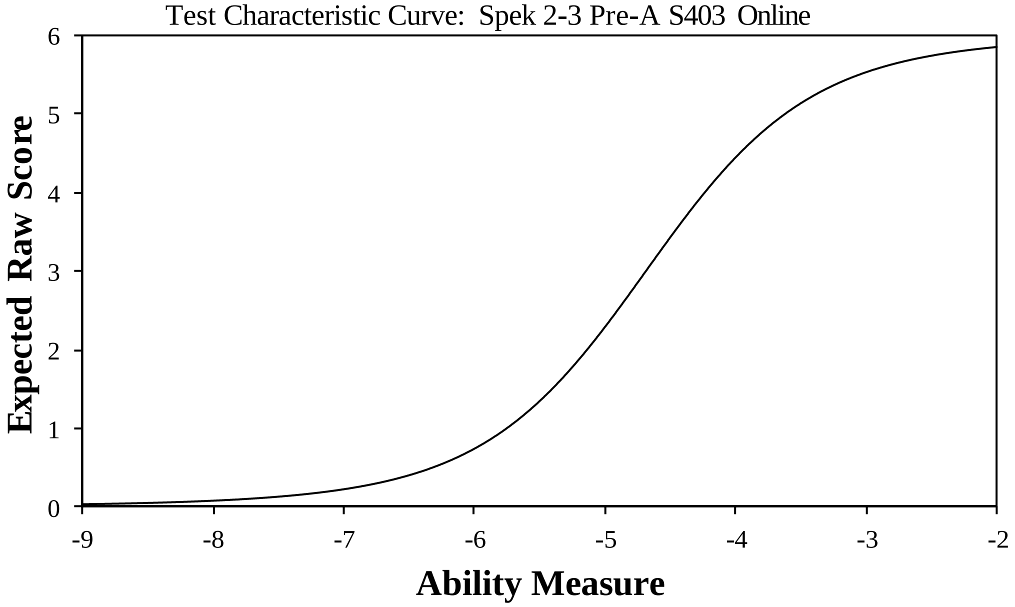


Figure 2.8.4.2.2
Test Characteristic Curve: Spek 2-3 A S403 Online

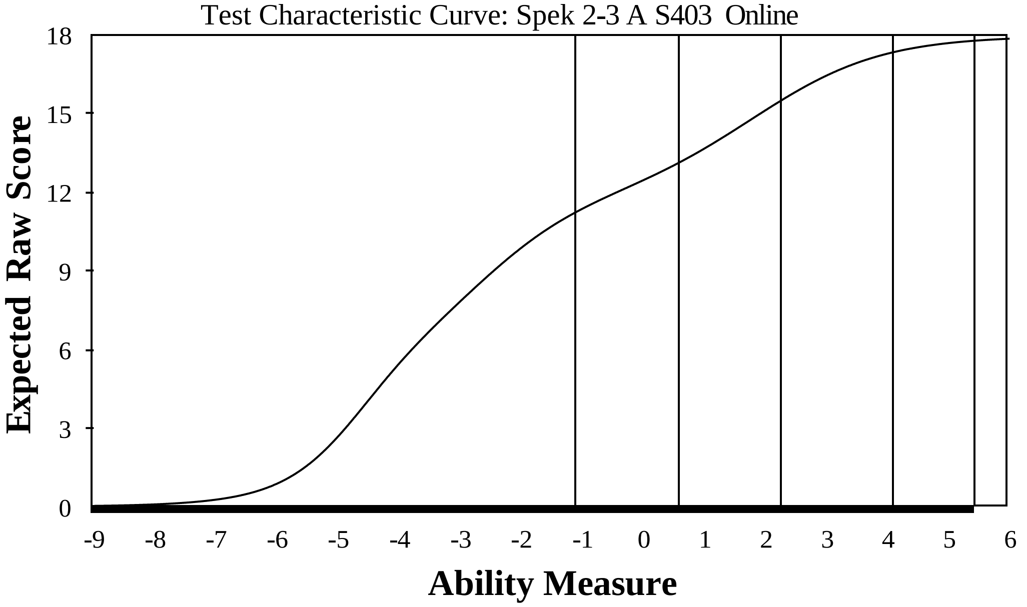


Figure 2.8.4.2.3

Test Characteristic Curve: Spek 2-3 B/C S403 Online

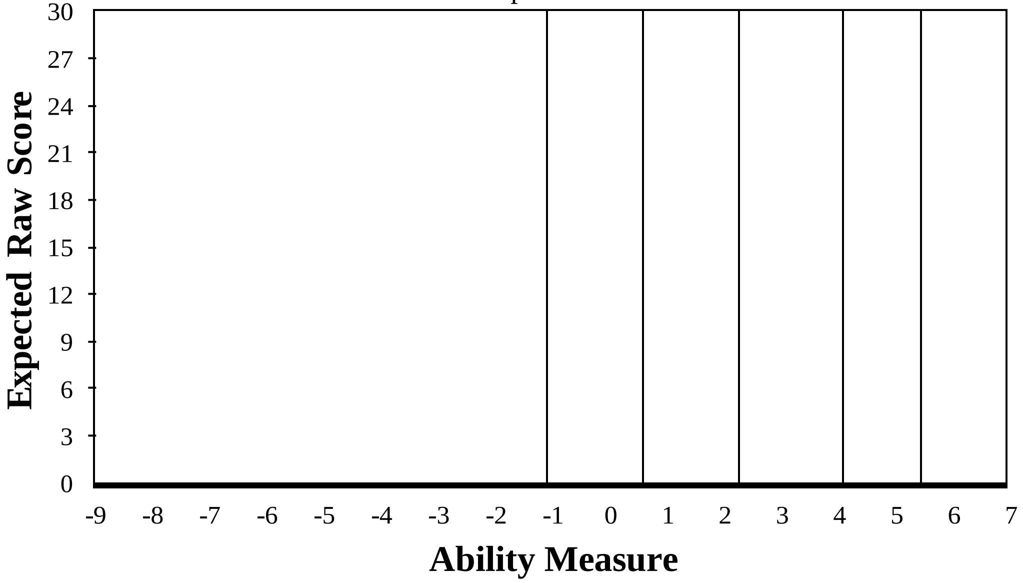
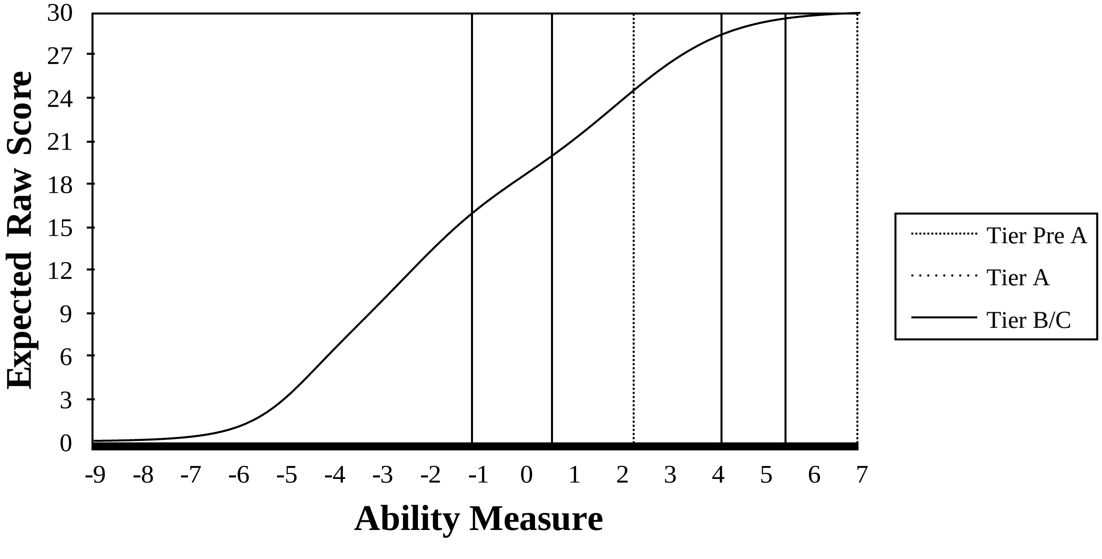


Figure 2.8.4.2.4

Test Characteristic Curve: Spek 2-3 S403 Online



2.8.4.3 *Grades 4–5*

Figure 2.8.4.3.1

Test Characteristic Curve: Spek 4-5 Pre-A S403 Online

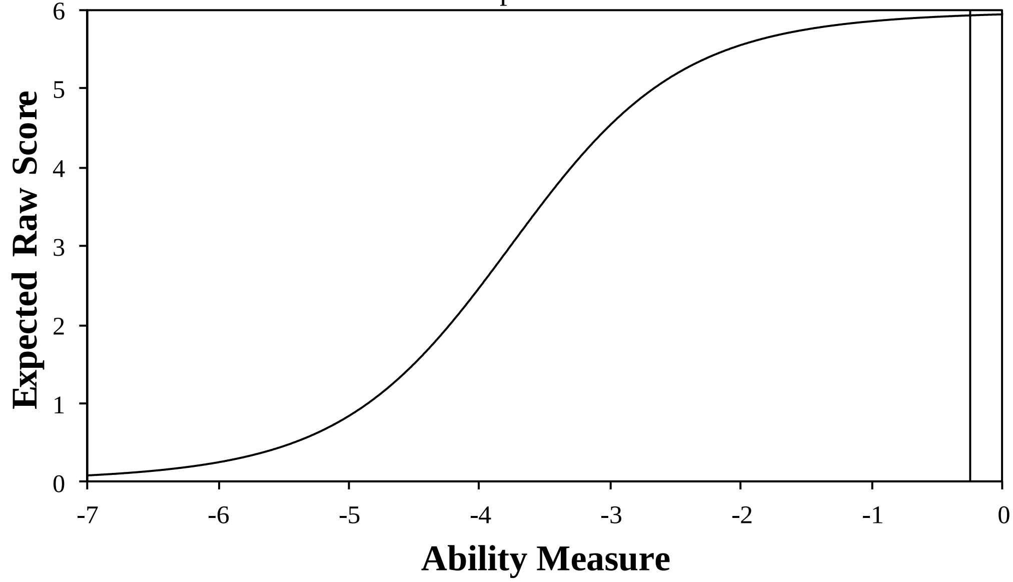


Figure 2.8.4.3.2

Test Characteristic Curve: Spek 4-5 A S403 Online

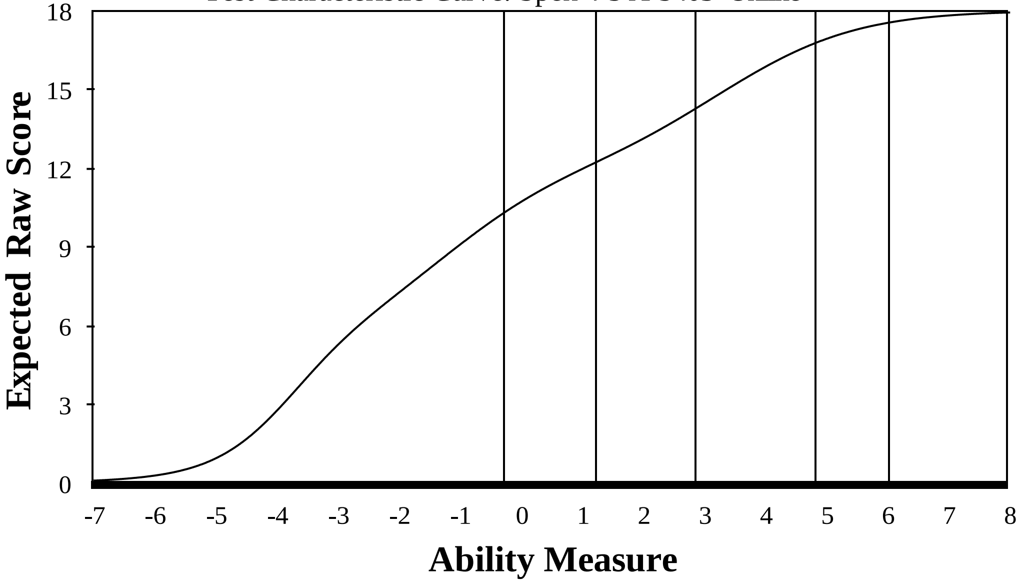


Figure 2.8.4.3.3

Test Characteristic Curve: Spek 4-5 B/C S403 Online

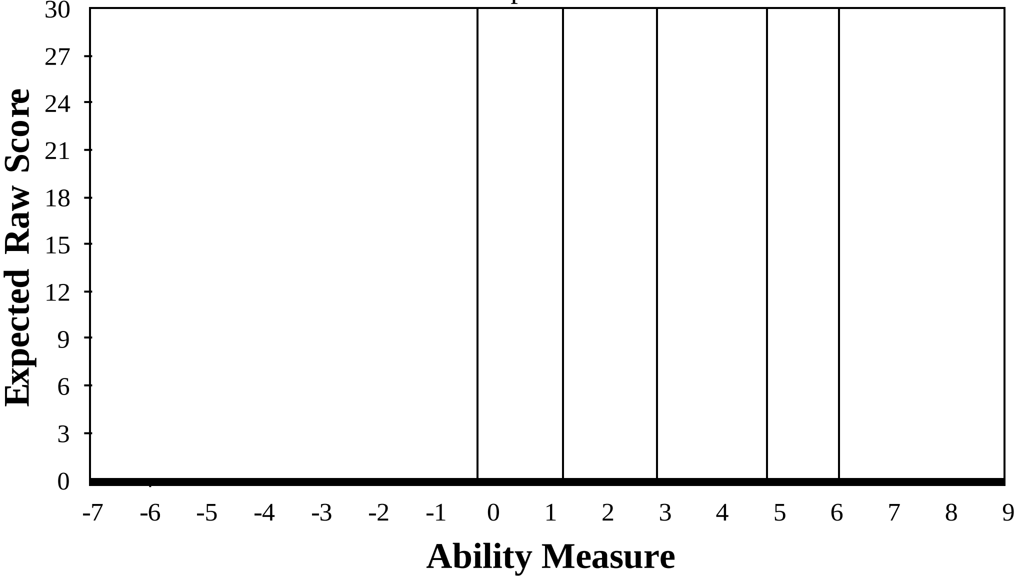
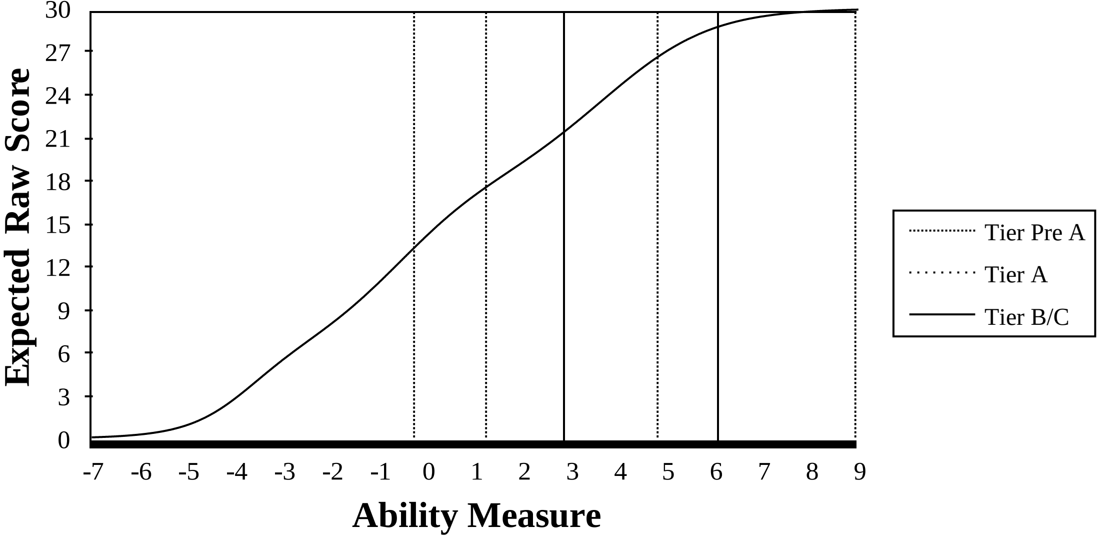


Figure 2.8.4.3.4

Test Characteristic Curve: Spek 4-5 S403 Online



2.8.4.4 *Grades 6–8*

Figure 2.8.4.4.1

Test Characteristic Curve: Spek 6-8 Pre-A S403 Online

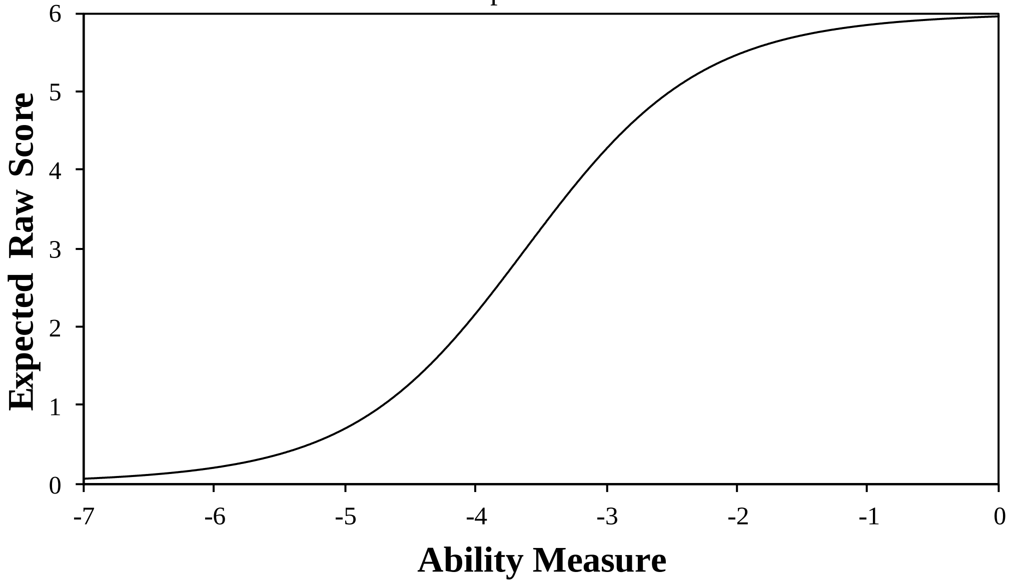


Figure 2.8.4.4.2

Test Characteristic Curve: Spek 6-8 A S403 Online

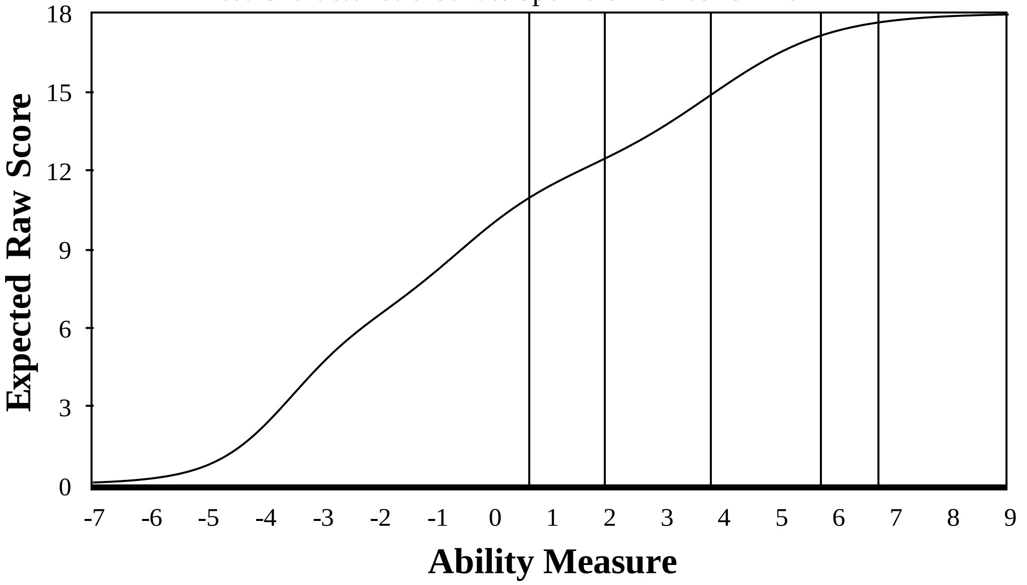


Figure 2.8.4.4.3

Test Characteristic Curve: Spek 6-8 B/C S403 Online

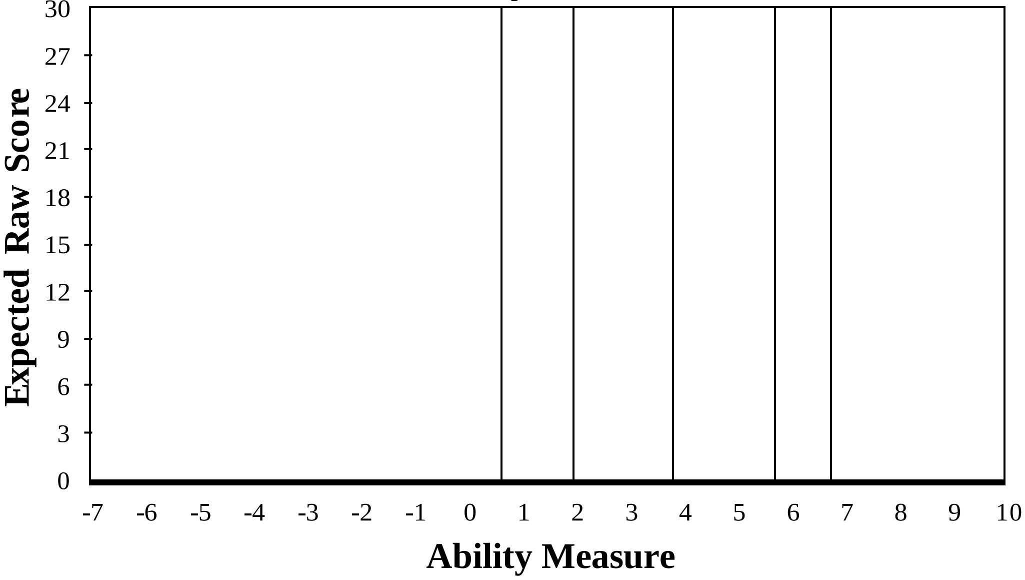
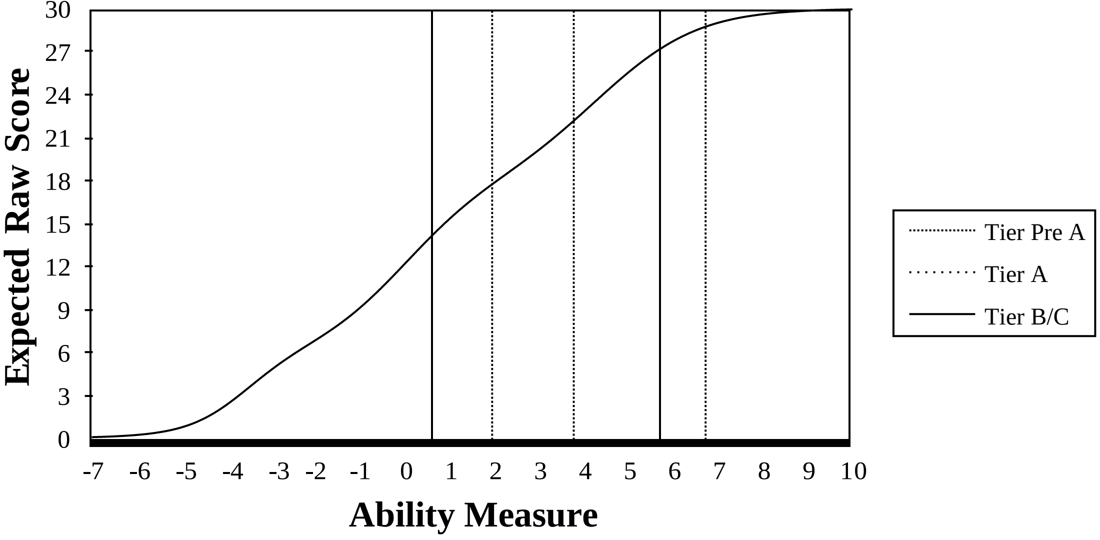


Figure 2.8.4.4.4

Test Characteristic Curve: Spek 6-8 S403 Online



2.8.4.5 *Grades 9–12*

Figure 2.8.4.5.1
Test Characteristic Curve: Spek 9-12 Pre-A S403 Online

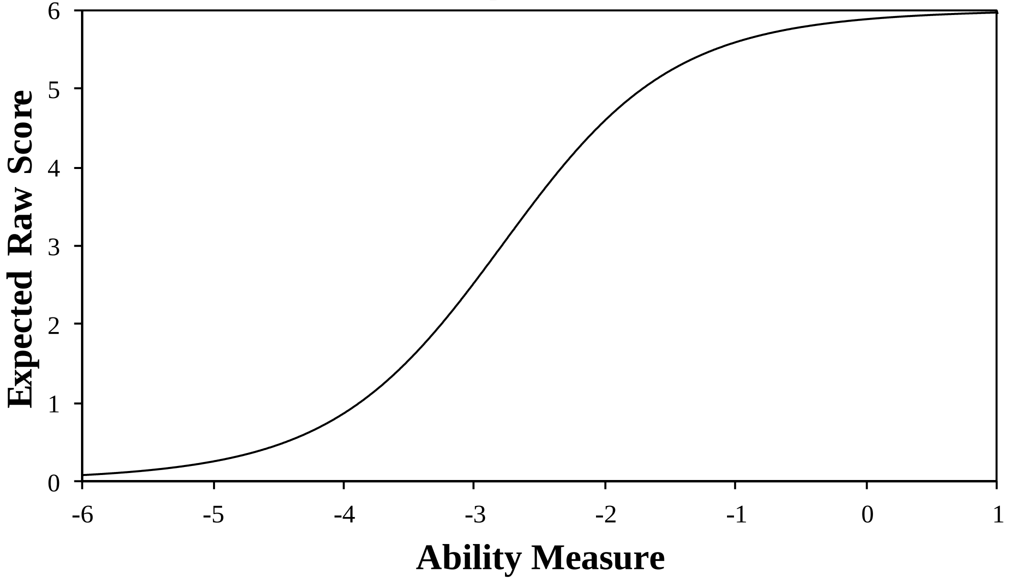


Figure 2.8.4.5.2
Test Characteristic Curve: Spek 9-12 A S403 Online

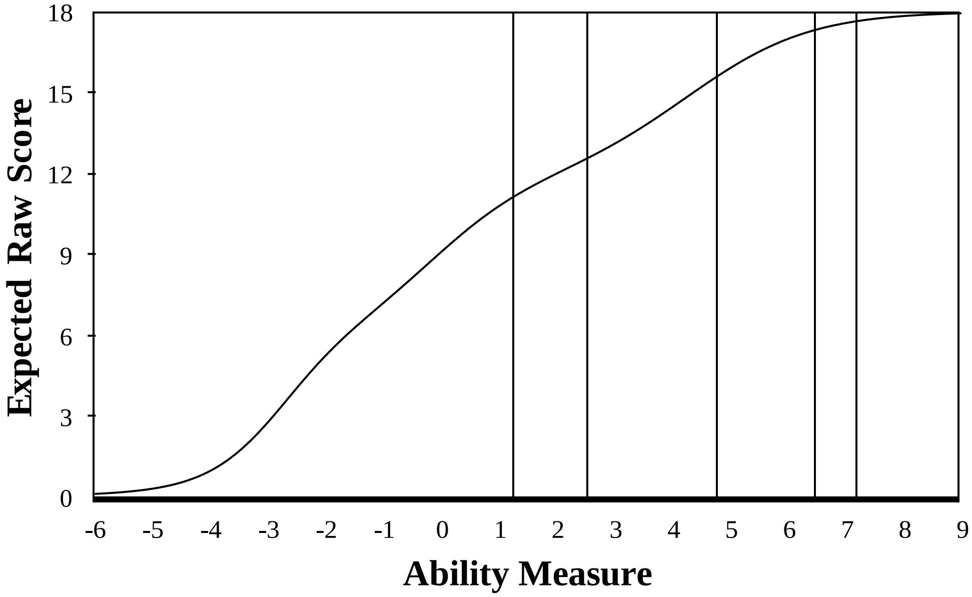


Figure 2.8.4.5.3
Test Characteristic Curve: Spek 9-12 B/C S403 Online

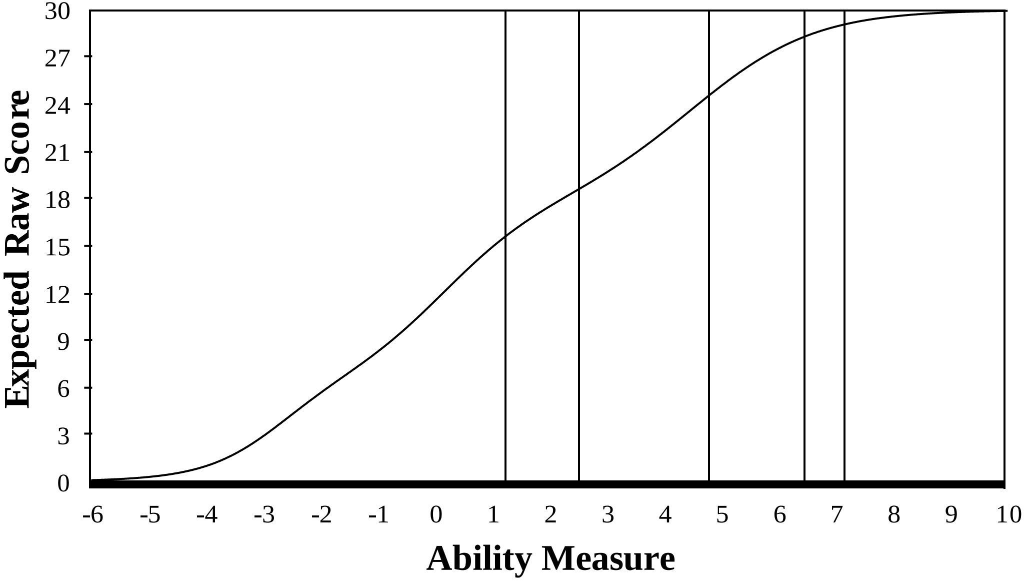
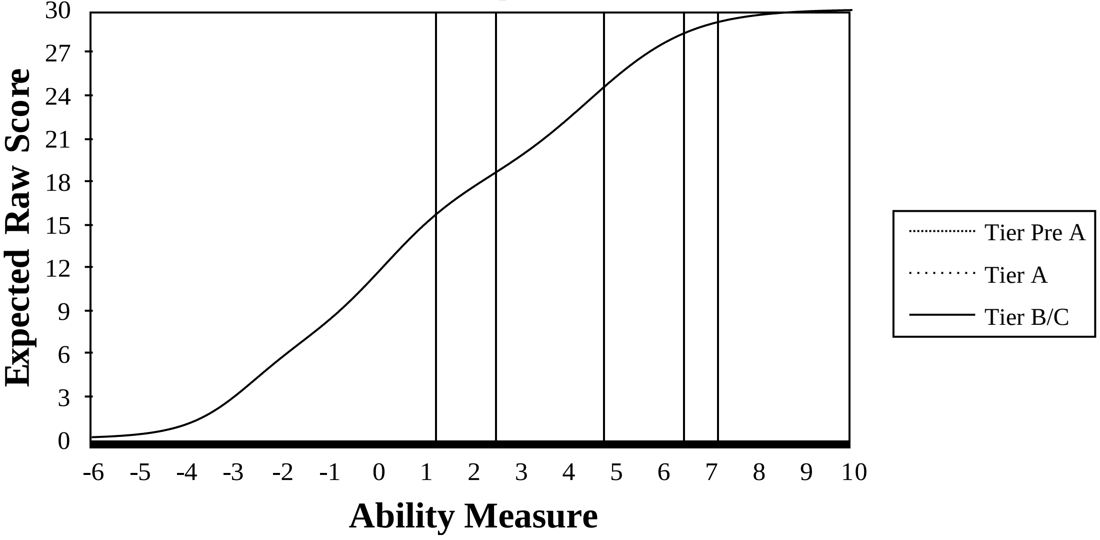


Figure 2.8.4.5.4
Test Characteristic Curve: Spek 9-12 S403 Online



2.9 Test Information Function

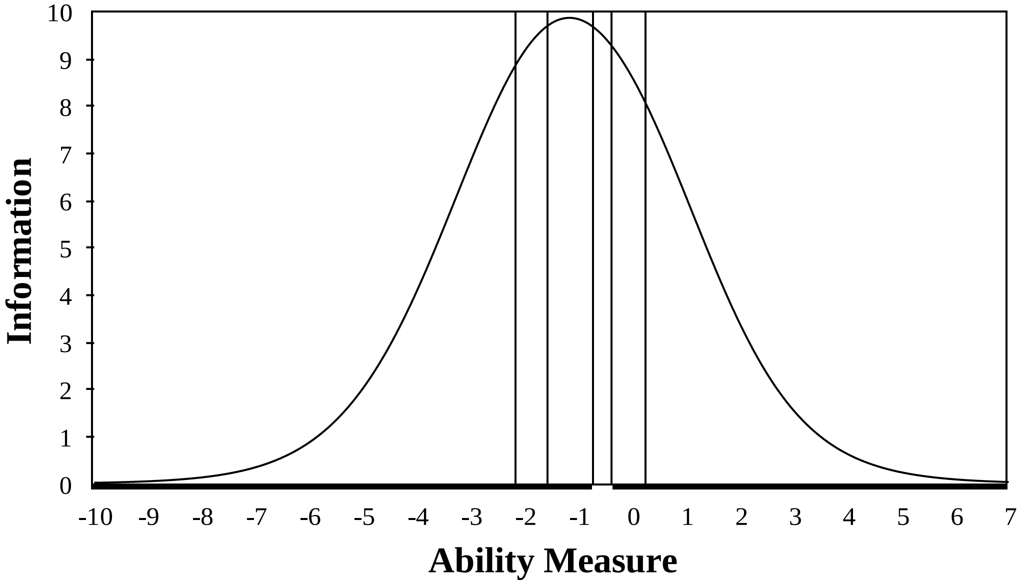
With the Rasch measurement model, as with any measurement model following item response theory, the relationship between the ability measure (in logits) and the accuracy of test scores can be modeled. It is recognized that tests measure most accurately when the abilities of the examinees and the difficulty 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 a test is too easy for an examinee (i.e., the examinee receives a perfect or near-perfect score), the examinee's ability cannot be accurately measured. The figures in this section show graphically how well the test is measuring across the ability measure spectrum. High test information values indicate more accuracy in measurement. Figures show 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 errors of measurement.

Five vertical lines in the figure indicate the five ACCESS cut scores for the highest grade in the grade-level cluster for the test form, dividing the figure into six sections for each of the WIDA proficiency levels (1–6) for the domain being tested. The ACCESS cut scores lines are presented along with the test information function to facilitate the interpretation of the test information curves. The test information curve and the corresponding ACCESS cut score lines are both expressed on the ACCESS logit scale. Note that for Speaking, in Tier Pre-A, all scores fall in the PL 1.0 range, so for some graphs there are no vertical lines expressing the cuts between proficiency levels.

2.9.1 Listening

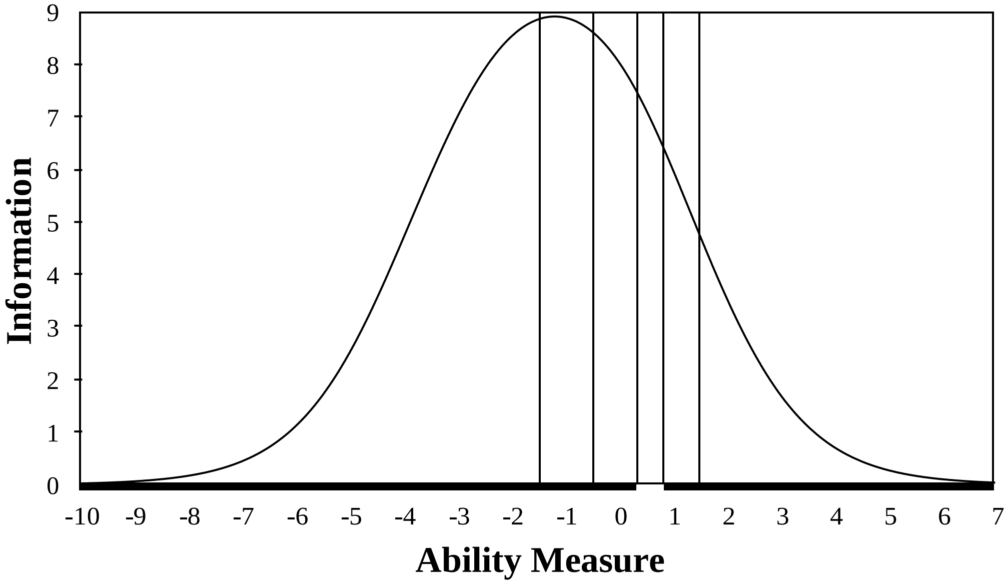
2.9.1.1 *Grade 1*

Figure 2.9.1.1
Test Information Function: List 1 S403 Online



2.9.1.2 *Grades 2–3*

Figure 2.9.1.2
Test Information Function: List 2-3 S403 Online



2.9.1.3 *Grades 4–5*

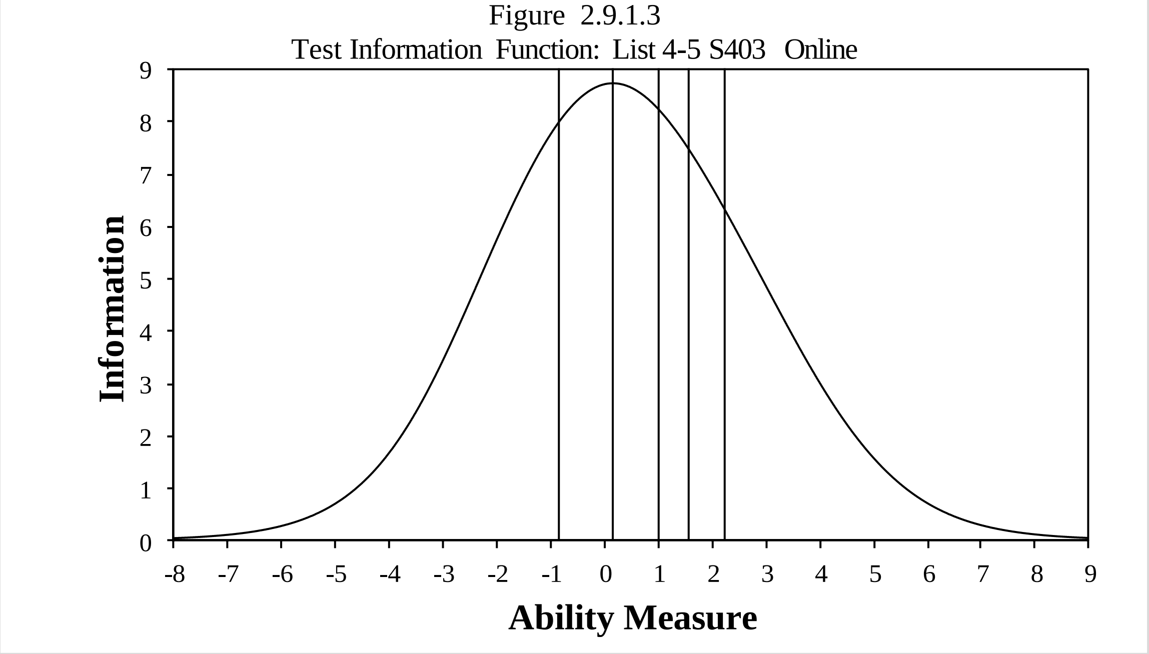


Figure 2.9.1.3
Test Information Function: List 4-5 S403 Online

2.9.1.4 *Grades 6–8*

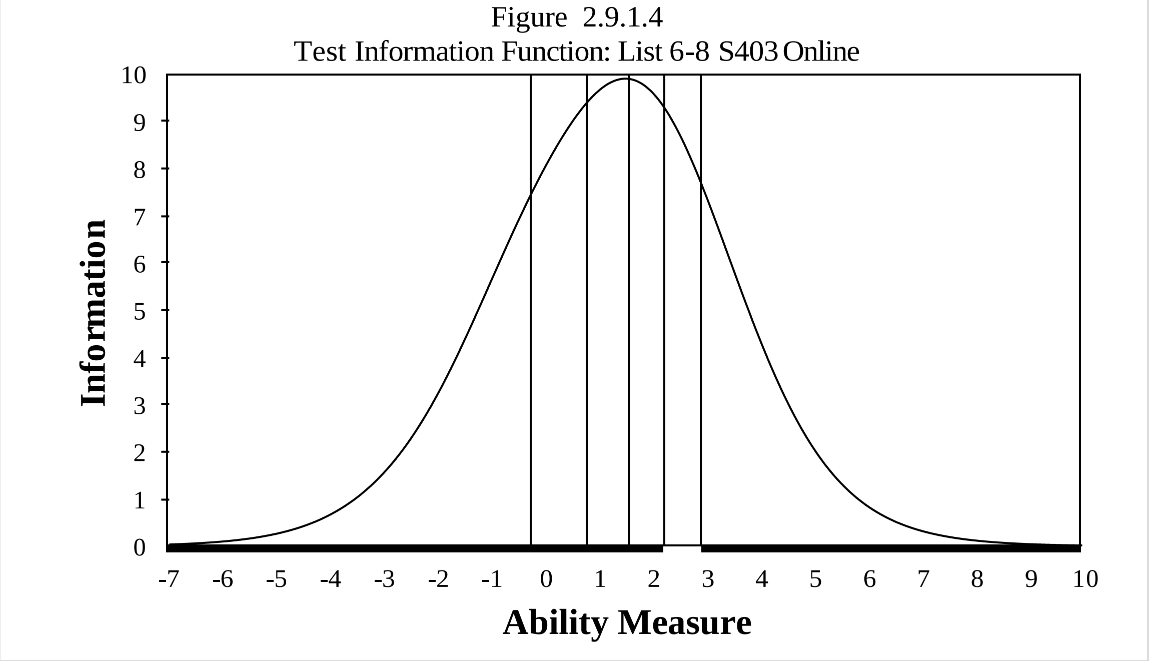
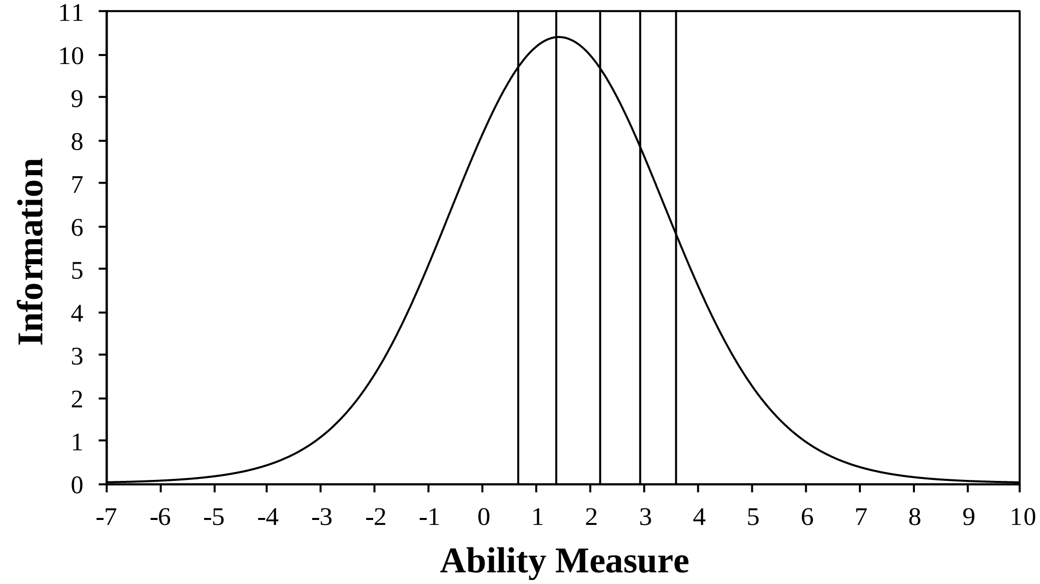


Figure 2.9.1.4
Test Information Function: List 6-8 S403 Online

2.9.1.5 *Grades 9–12*

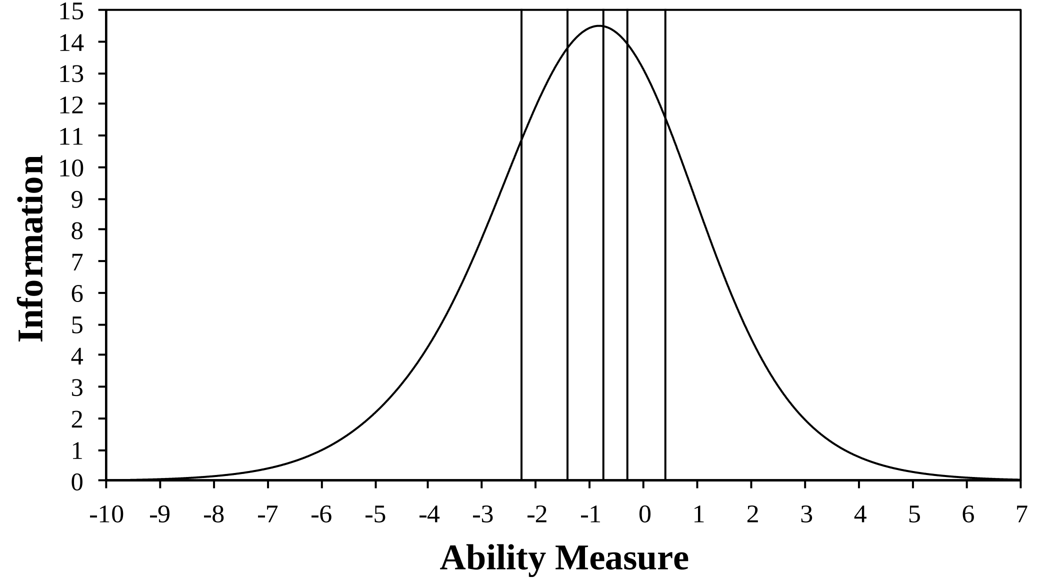
Figure 2.9.1.5
Test Information Function: List 9-12 S403 Online



2.9.2 Reading

2.9.2.1 *Grade 1*

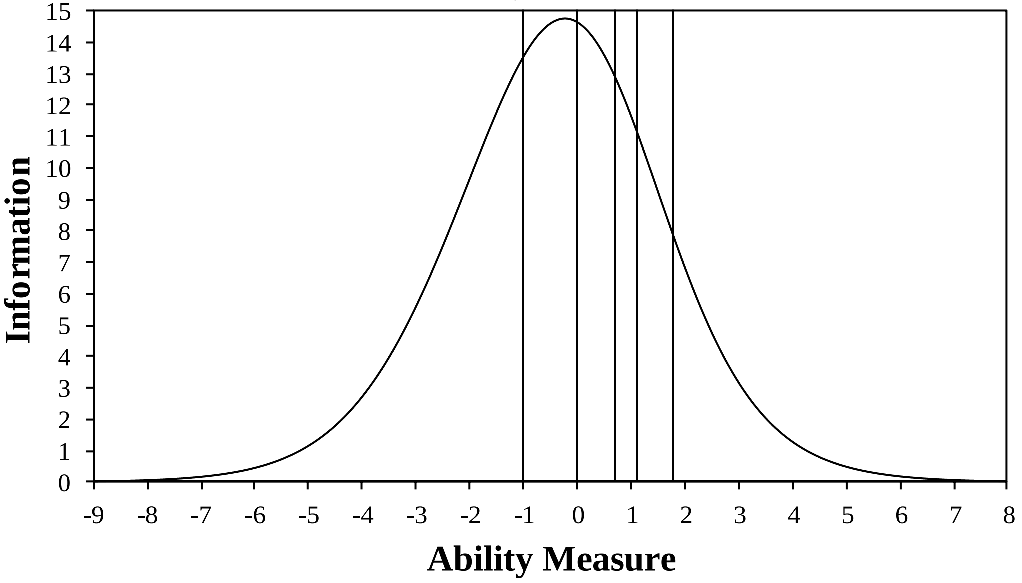
Figure 2.9.2.1
Test Information Function: Read 1 S403 Online



2.9.2.2 *Grades 2–3*

Figure 2.9.2.2

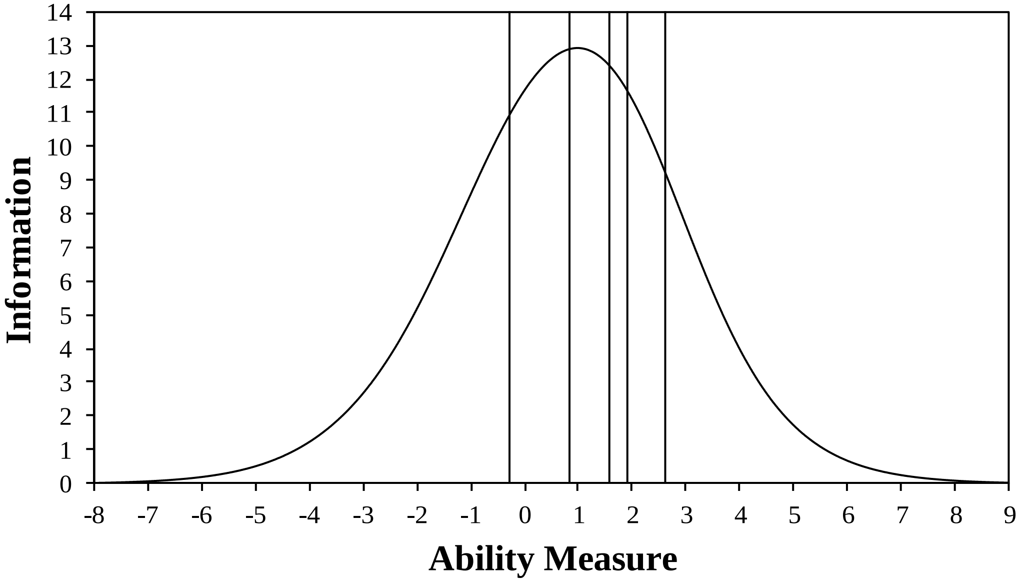
Test Information Function: Read 2-3 S403 Online



2.9.2.3 *Grades 4–5*

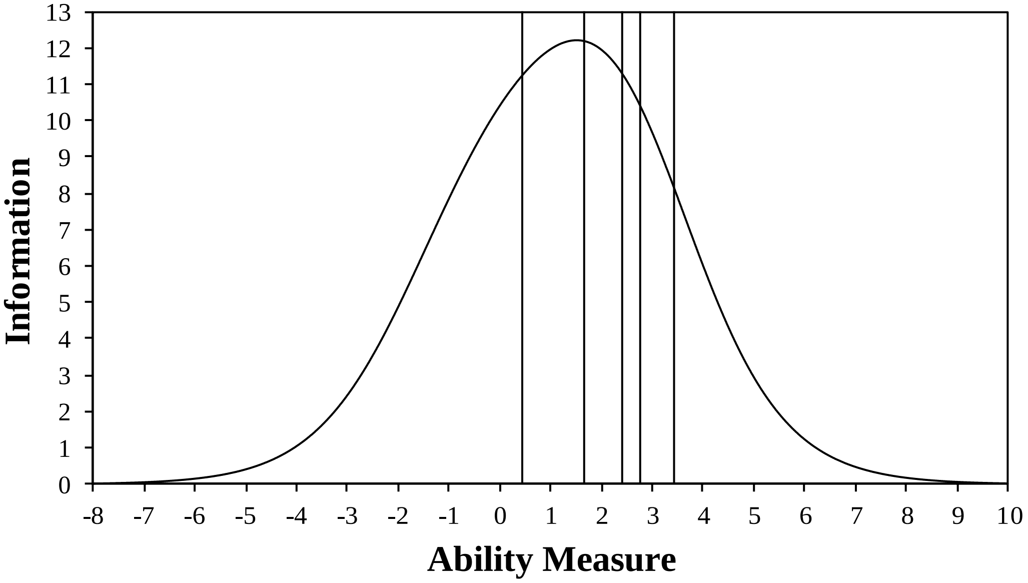
Figure 2.9.2.3

Test Information Function: Read 4-5 S403 Online



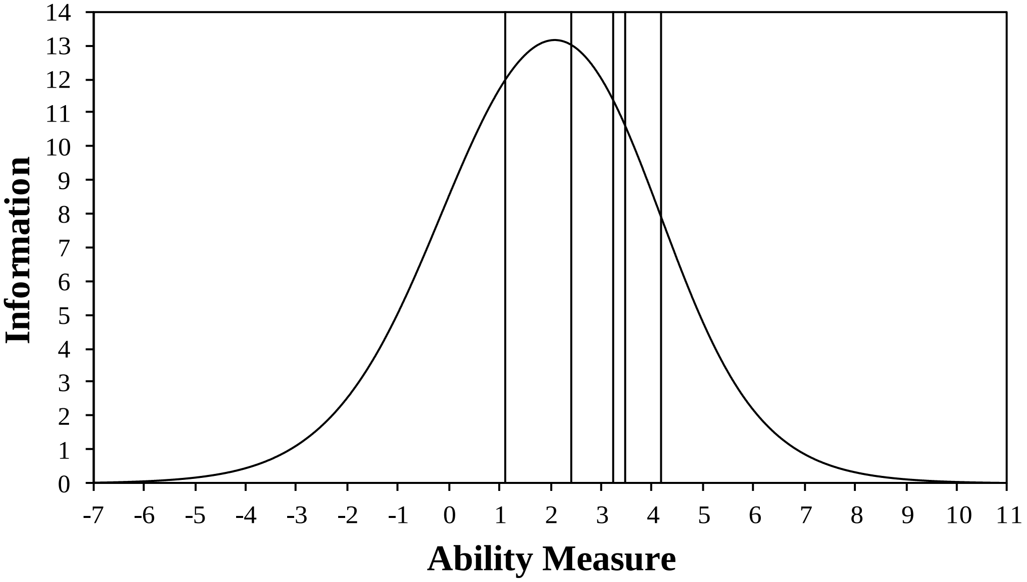
2.9.2.4 Grades 6–8

Figure 2.9.2.4
Test Information Function: Read 6-8 S403 Online



2.9.2.5 Grades 9–12

Figure 2.9.2.5
Test Information Function: Read 9-12 S403 Online



2.9.3 Writing

2.9.3.1 *Grade 1*

Figure 2.9.3.1.1

Test Information Function: Writ 1 A S403 Online

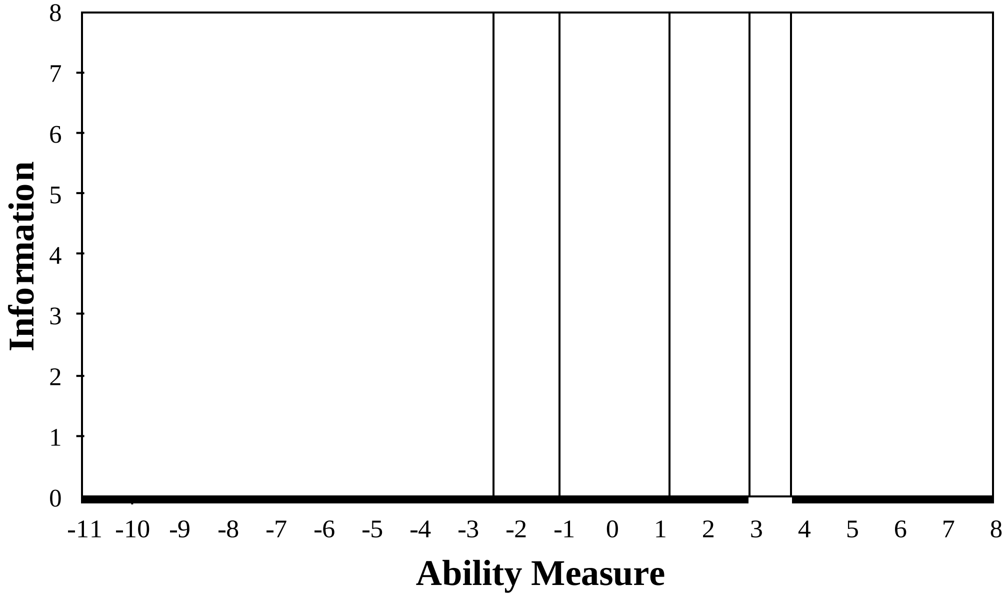


Figure 2.9.3.1.2

Test Information Function: Writ 1 B/C S403 Online

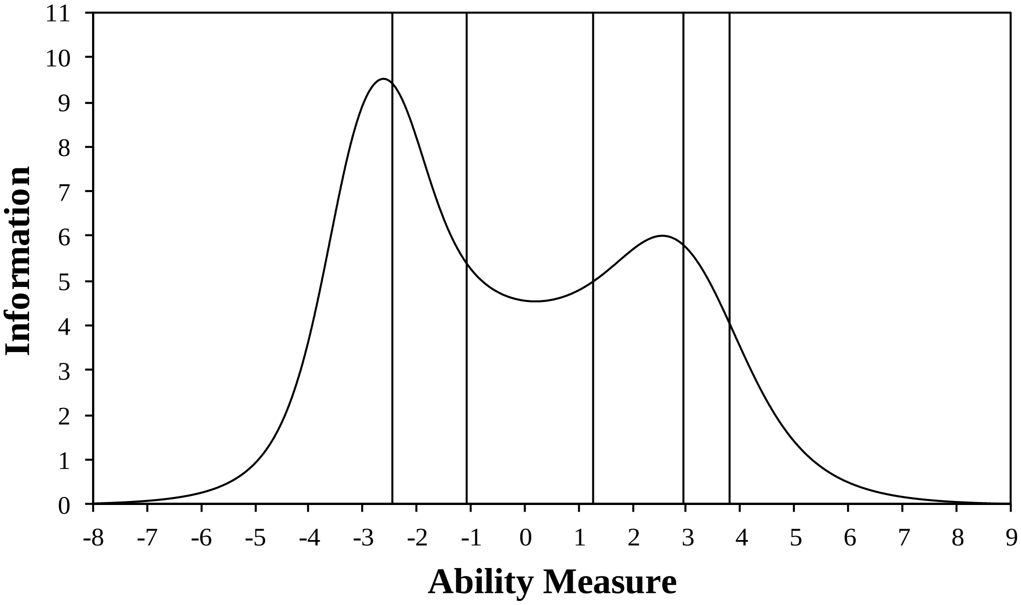
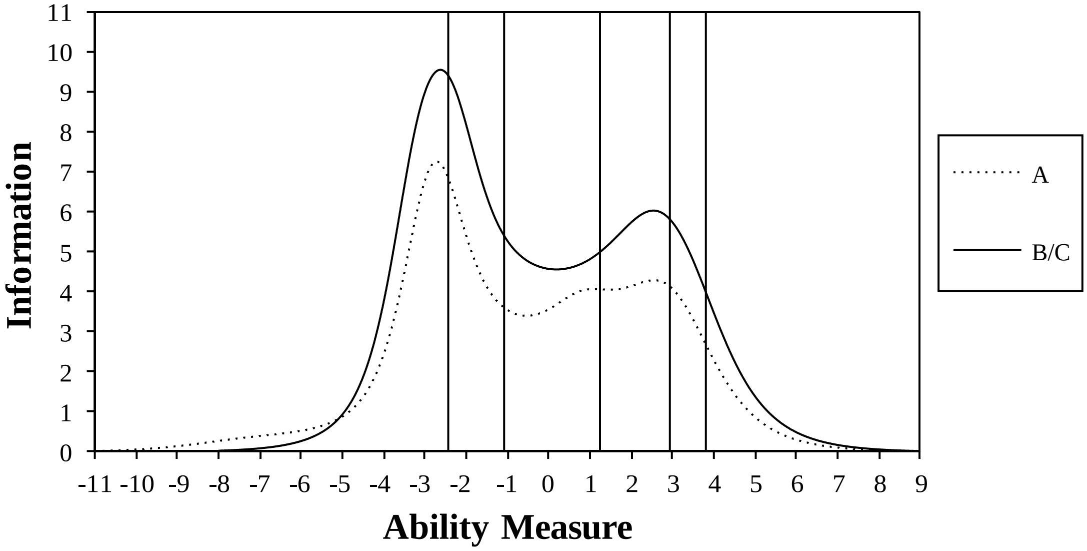


Figure 2.9.3.1.3
Test Information Function: Writ 1 S403 Online



2.9.3.2 *Grades 2–3*

Figure 2.9.3.2.1
Test Information Function: Writ 2-3 A S403 Online

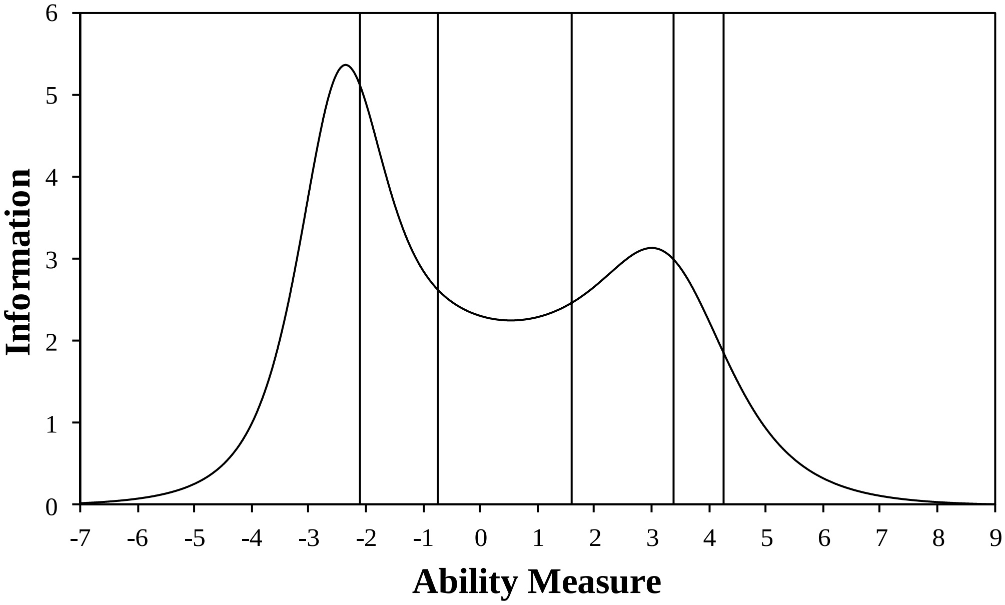


Figure 2.9.3.2.2
Test Information Function: Writ 2-3 B/C S403 Online

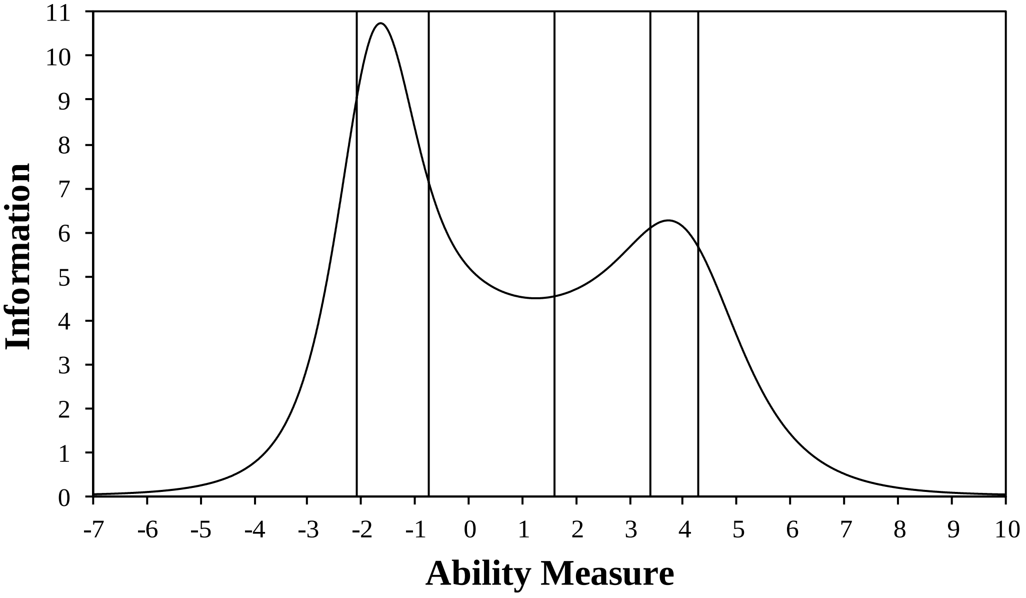
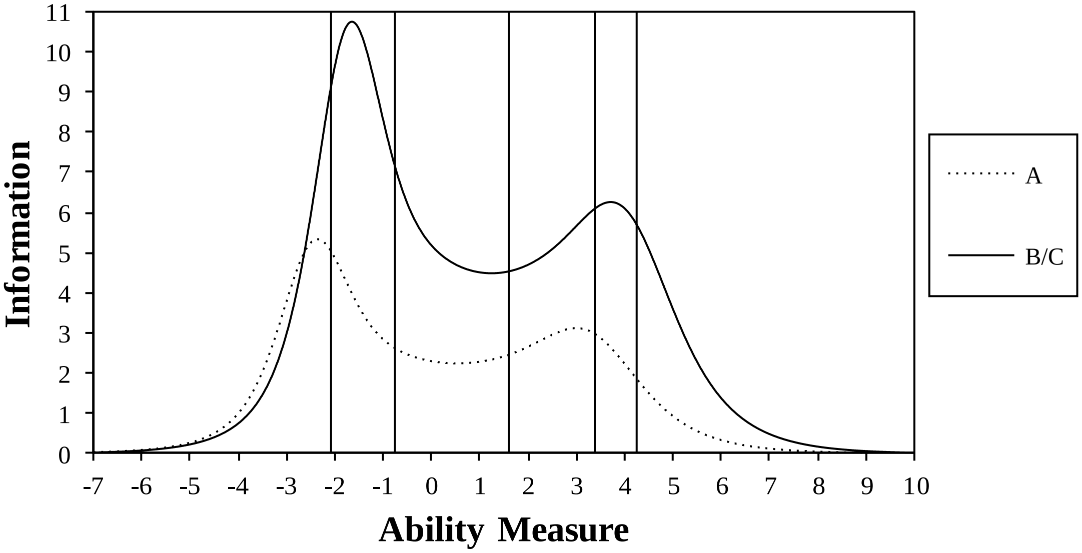


Figure 2.9.3.2.3
Test Information Function: Writ 2-3 S403 Online



2.9.3.3 Grades 4–5

Figure 2.9.3.3.1
Test Information Function: Writ 4-5 A S403 Online

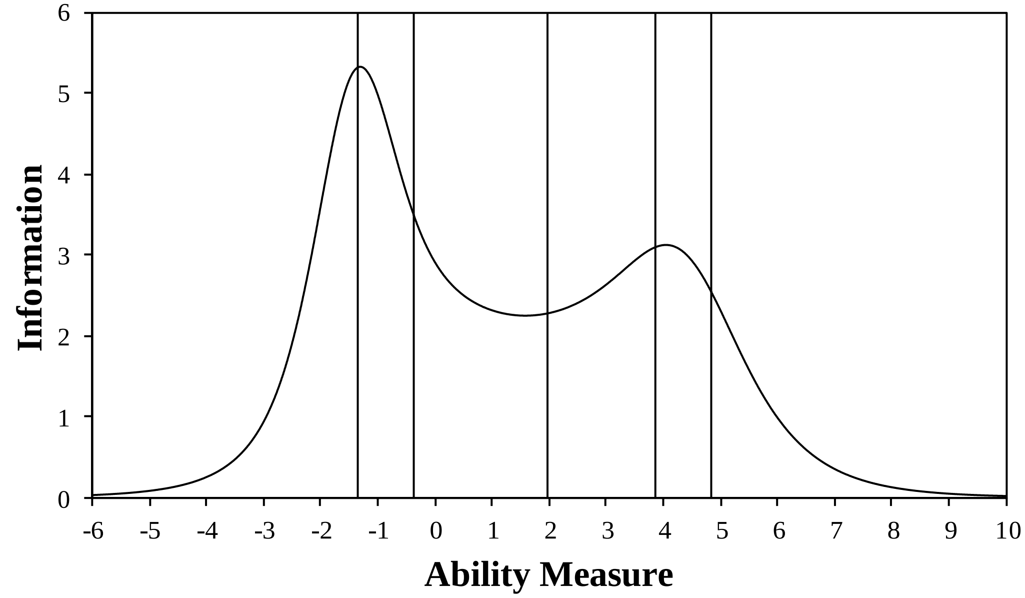


Figure 2.9.3.3.2
Test Information Function: Writ 4-5 B/C S403 Online

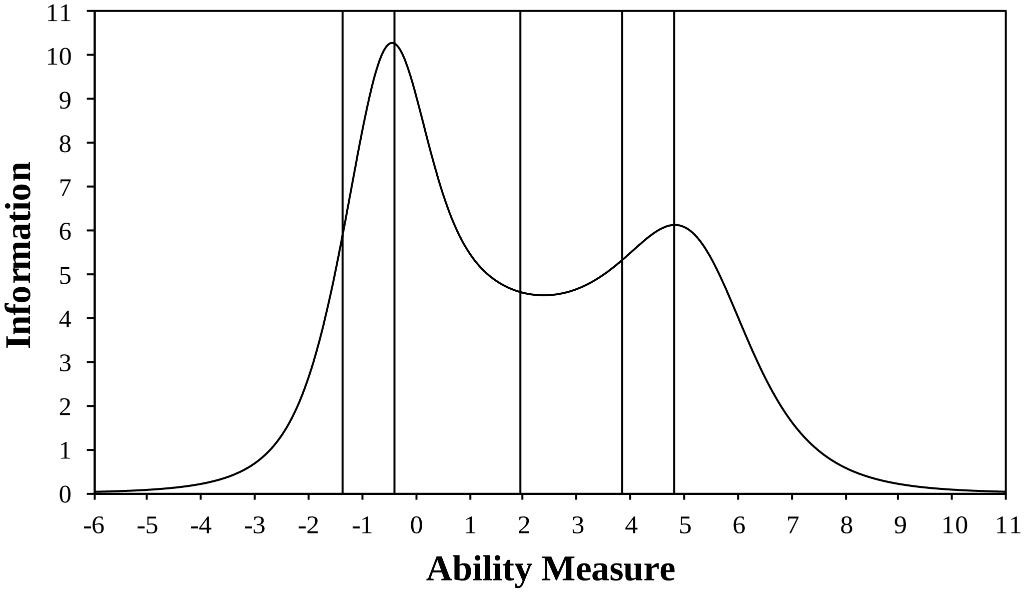
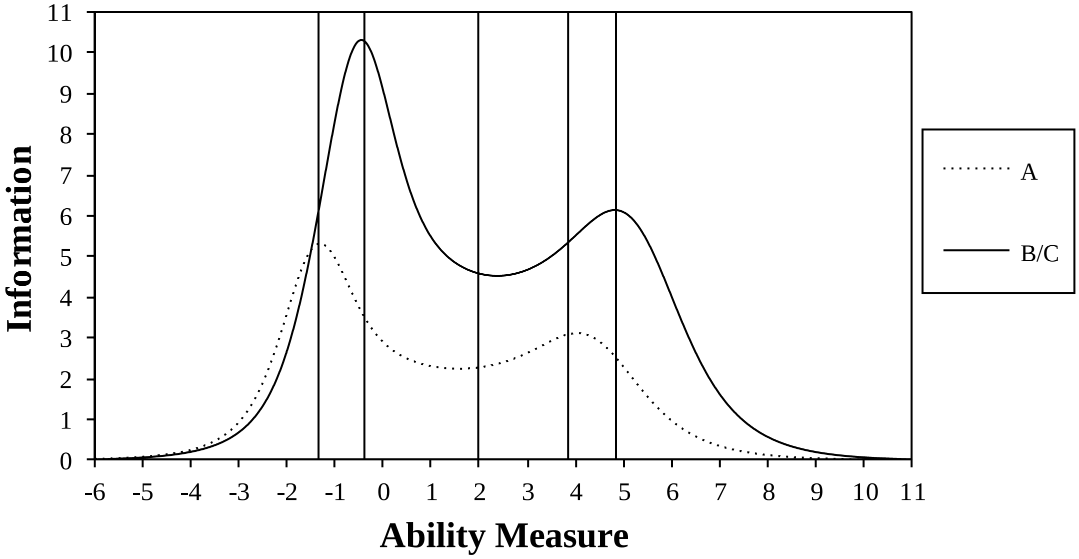


Figure 2.9.3.3.3

Test Information Function: Writ 4-5 S403 Online



2.9.3.4 Grades 6–8

Figure 2.9.3.4.1

Test Information Function: Writ 6-8 A S403 Online

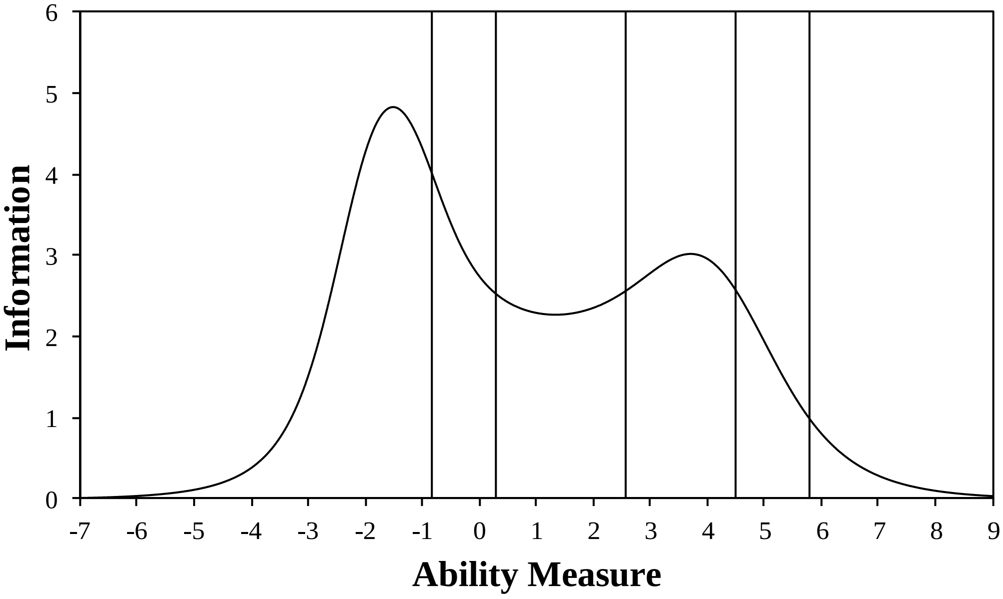


Figure 2.9.3.4.2
Test Information Function: Writ 6-8 B/C S403 Online

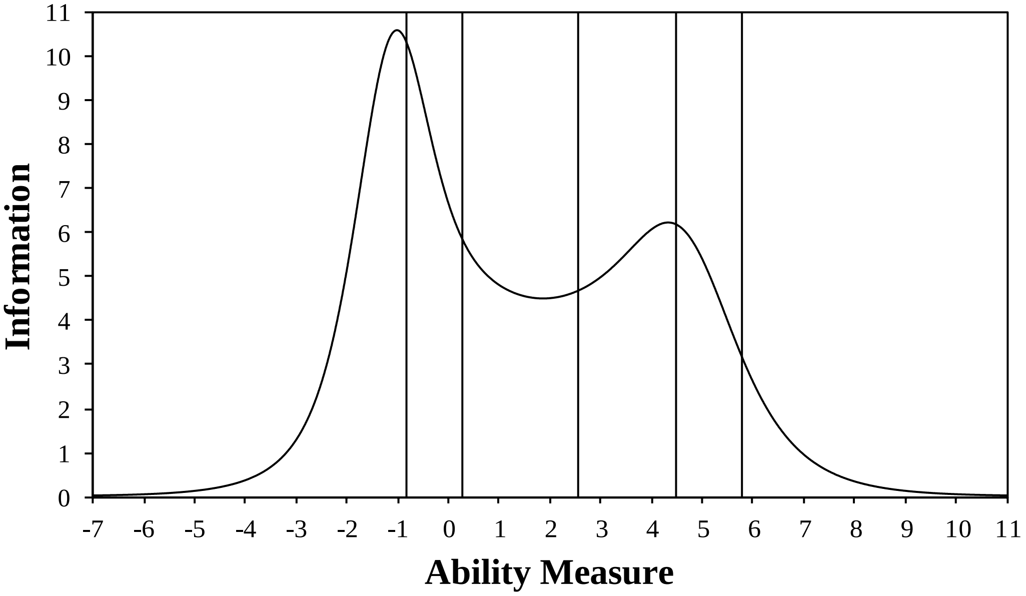
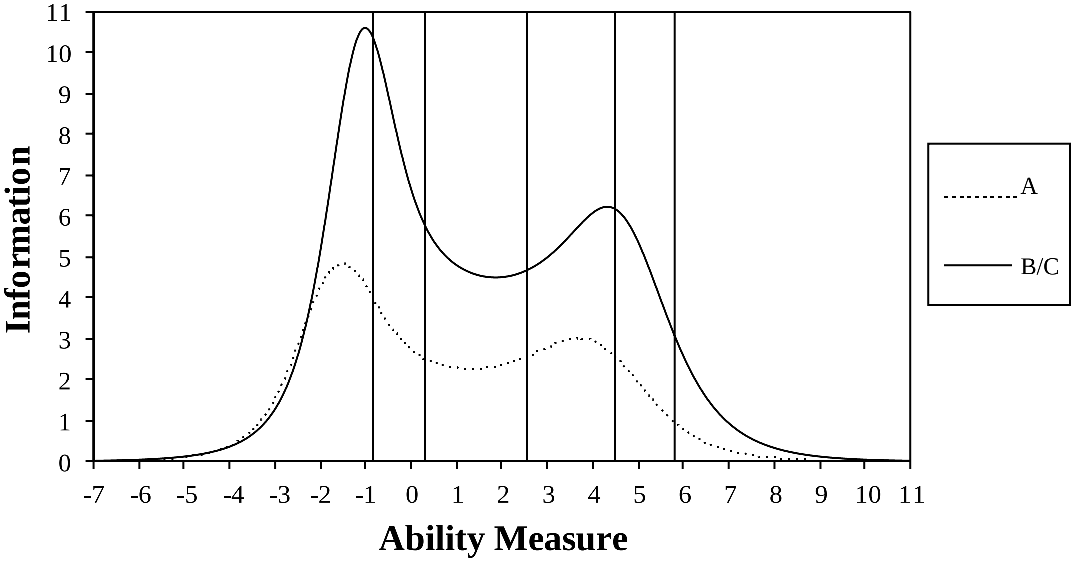


Figure 2.9.3.4.3
Test Information Function: Writ 6-8 S403 Online



2.9.3.5 Grades 9–12

Figure 2.9.3.5.1

Test Information Function: Writ 9-12 A S403 Online

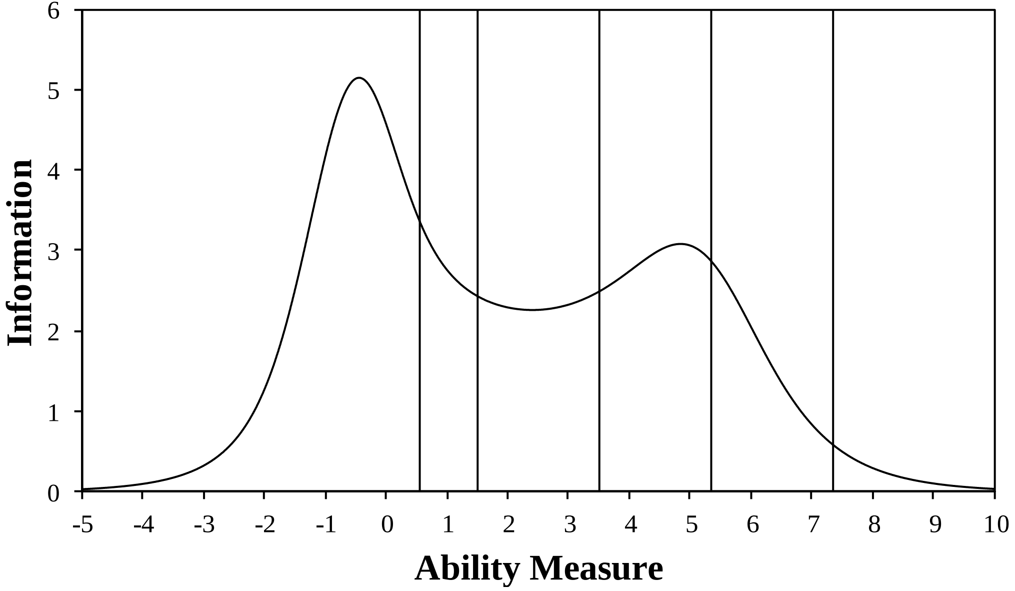


Figure 2.9.3.5.2

Test Information Function: Writ 9-12 B/C S403 Online

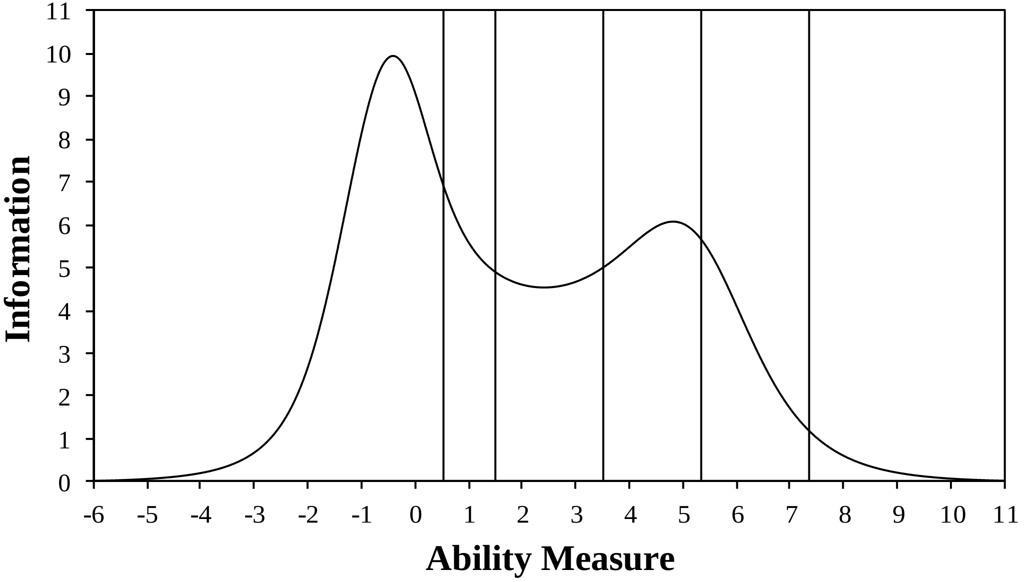
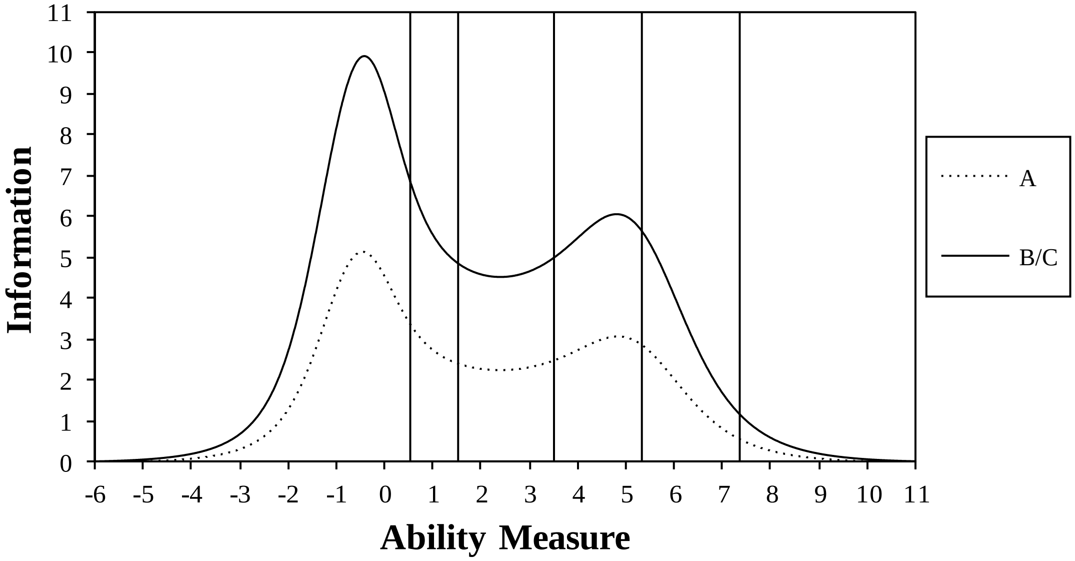


Figure 2.9.3.5.3
Test Information Function: Writ 9-12 S403 Online



2.9.4 Speaking

2.9.4.1 Grade 1

Figure 2.9.4.1.1

Test Information Function: Spek 1 Pre-A S403 Online

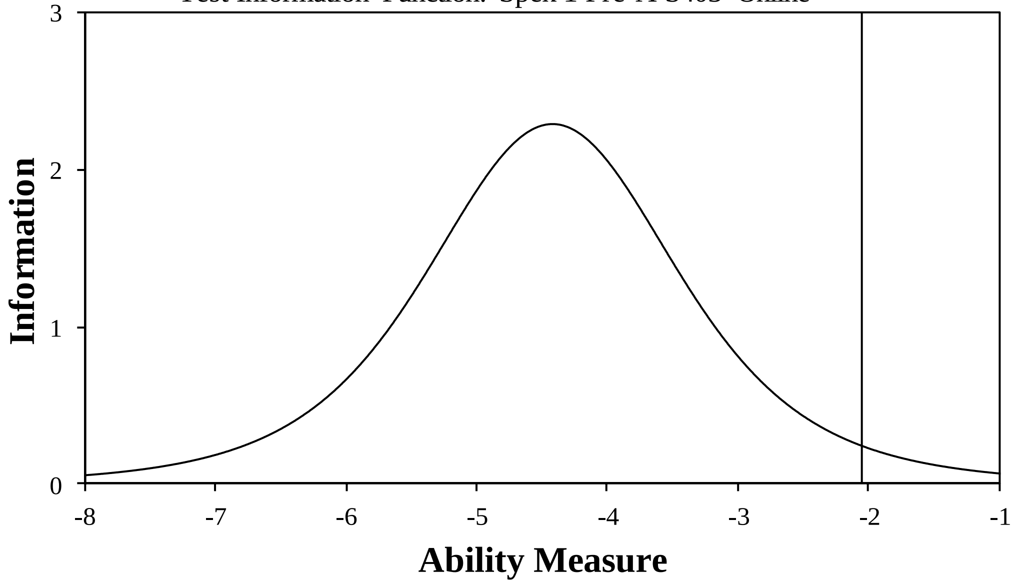


Figure 2.9.4.1.2

Test Information Function: Spek 1 A S403 Online

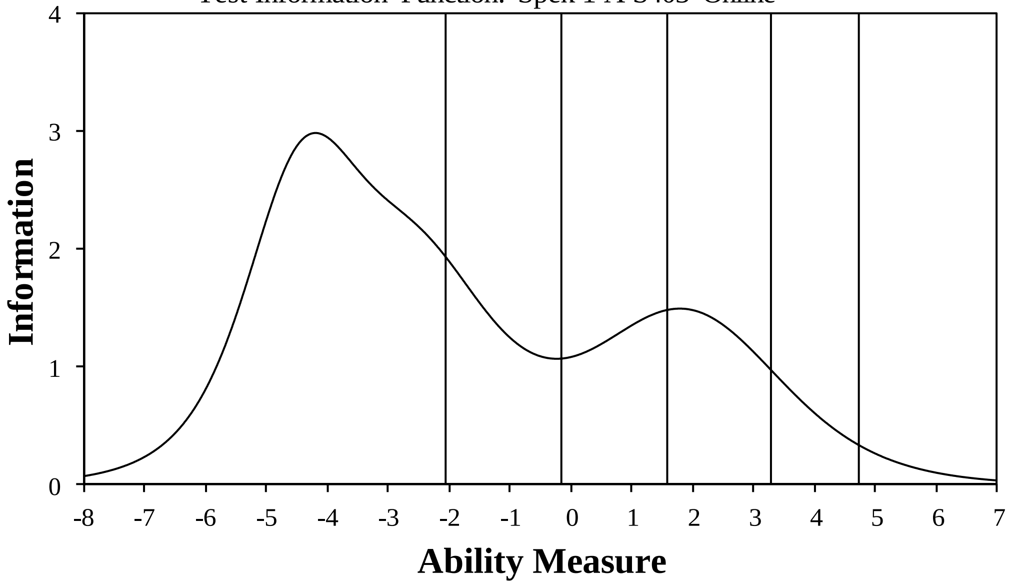


Figure 2.9.4.1.3
Test Information Function: Spek 1 B/C S403 Online

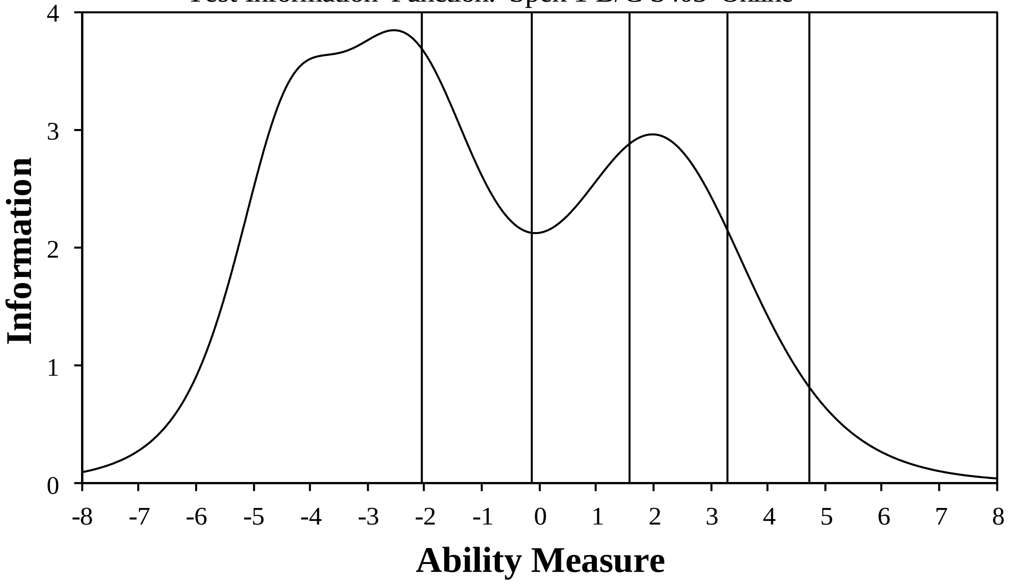
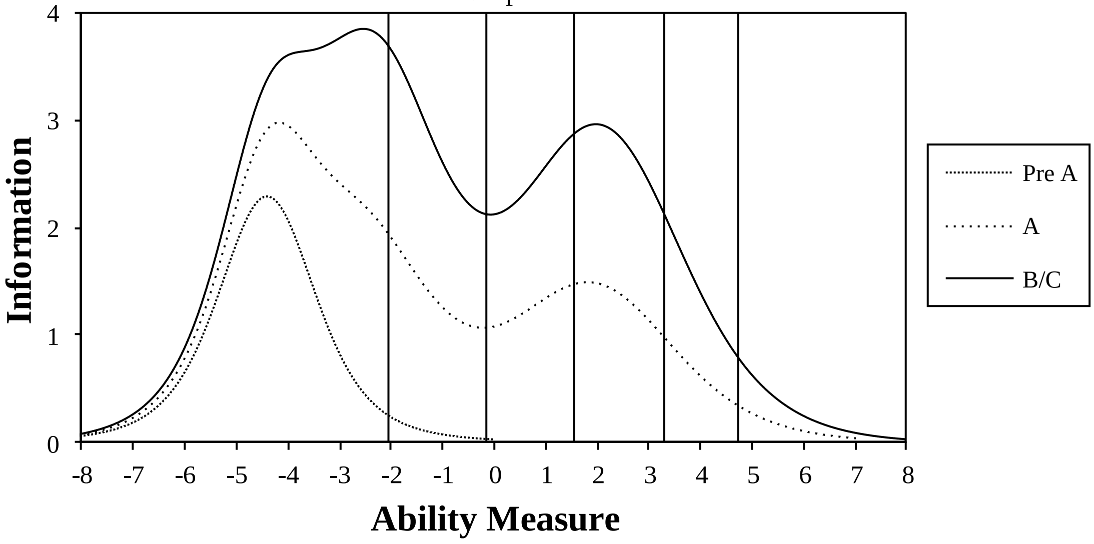


Figure 2.9.4.1.4
Test Information Function: Spek 1 S403 Online



2.9.4.2 Grades 2–3

Figure 2.9.4.2.1
Test Information Function: Spek 2-3 Pre-A S403 Online

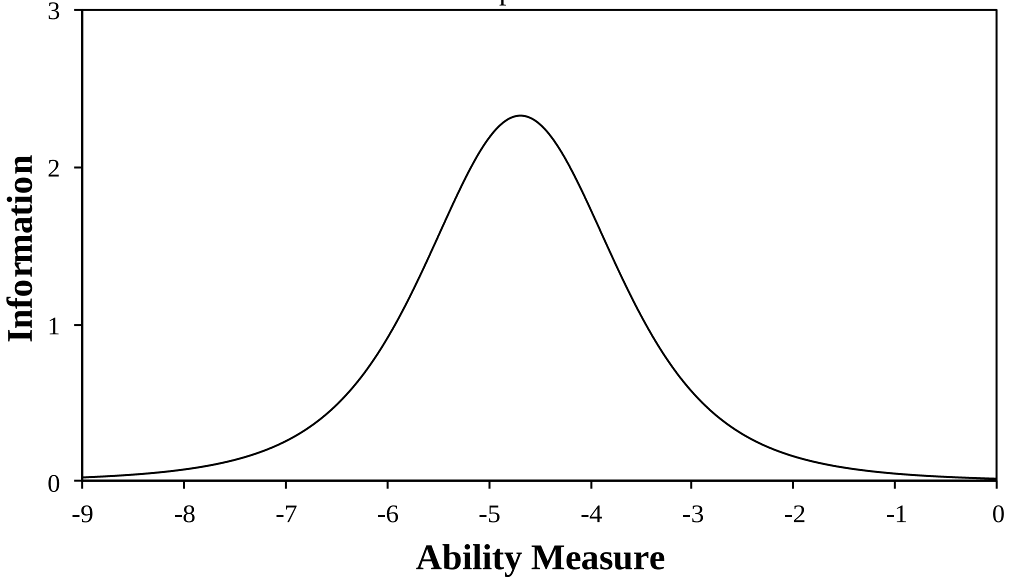


Figure 2.9.4.2.2
Test Information Function: Spek 2-3 A S403 Online

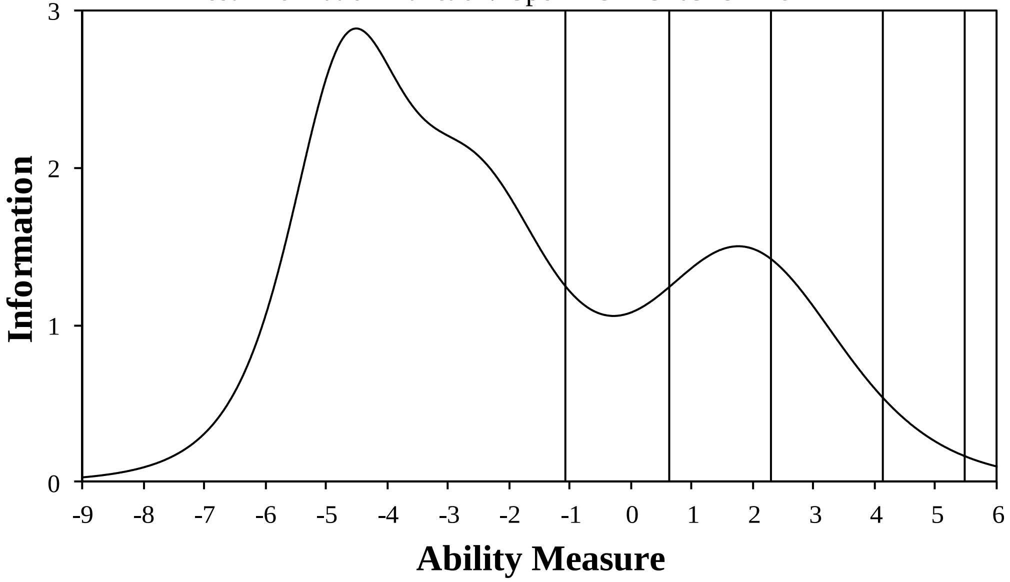


Figure 2.9.4.2.3

Test Information Function: Spek 2-3 B/C S403 Online

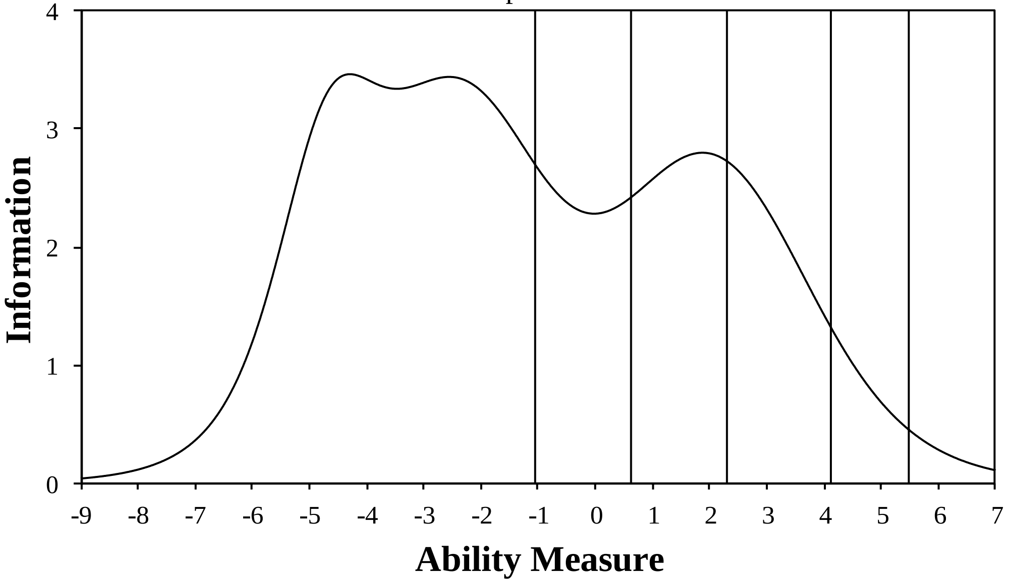
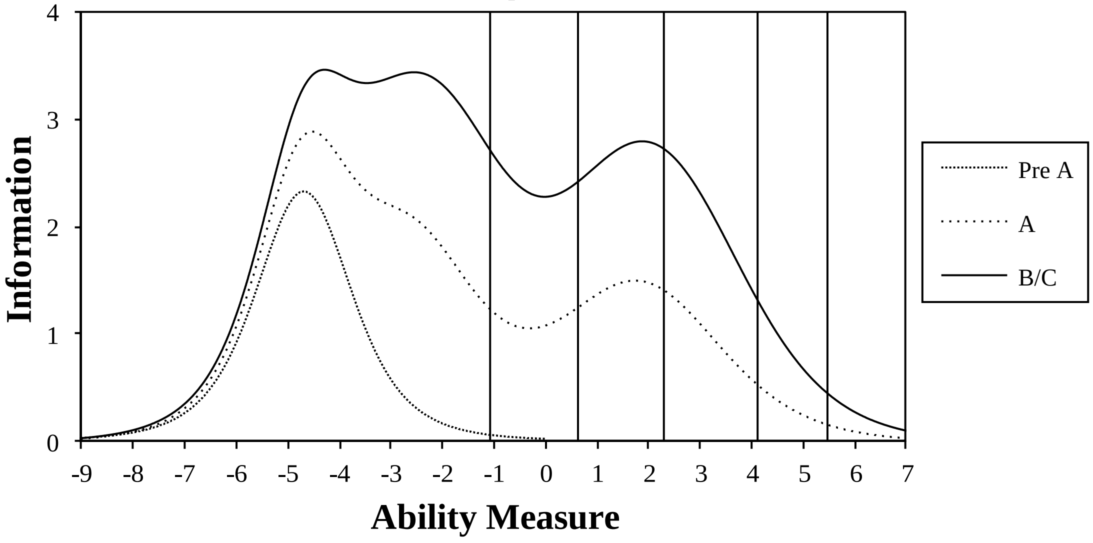


Figure 2.9.4.2.4

Test Information Function: Spek 2-3 S403 Online



2.9.4.3 *Grades 4–5*

Figure 2.9.4.3.1
Test Information Function: Spek 4-5 Pre-A S403 Online

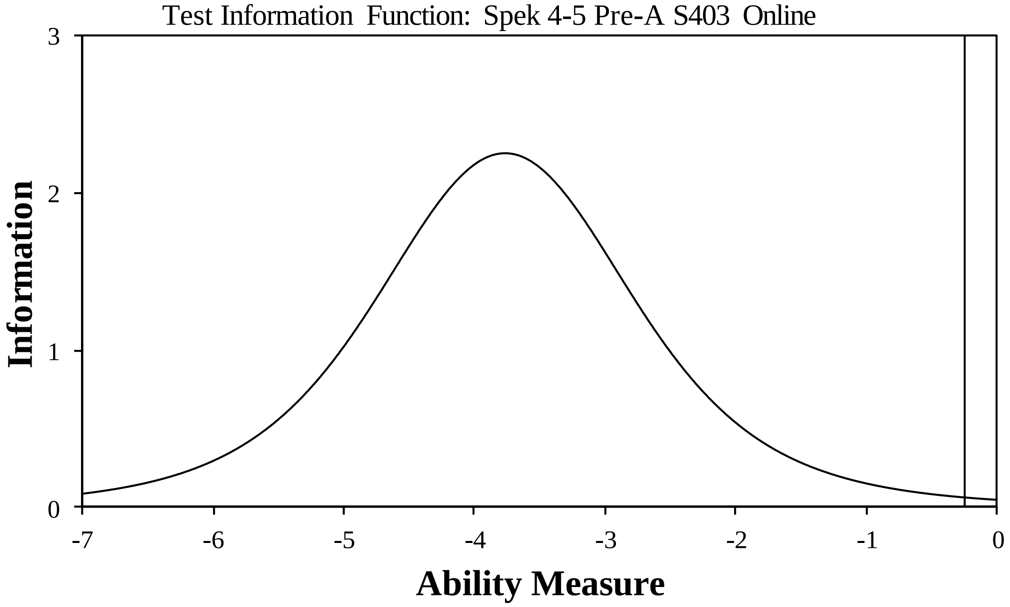


Figure 2.9.4.3.2
Test Information Function: Spek 4-5 A S403 Online

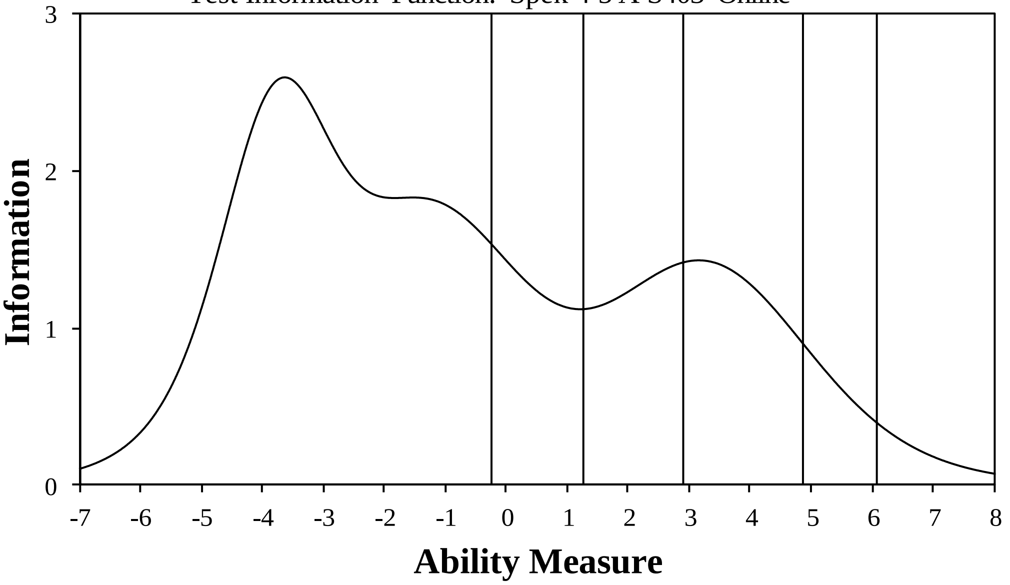


Figure 2.9.4.3.3
Test Information Function: Spek 4-5 B/C S403 Online

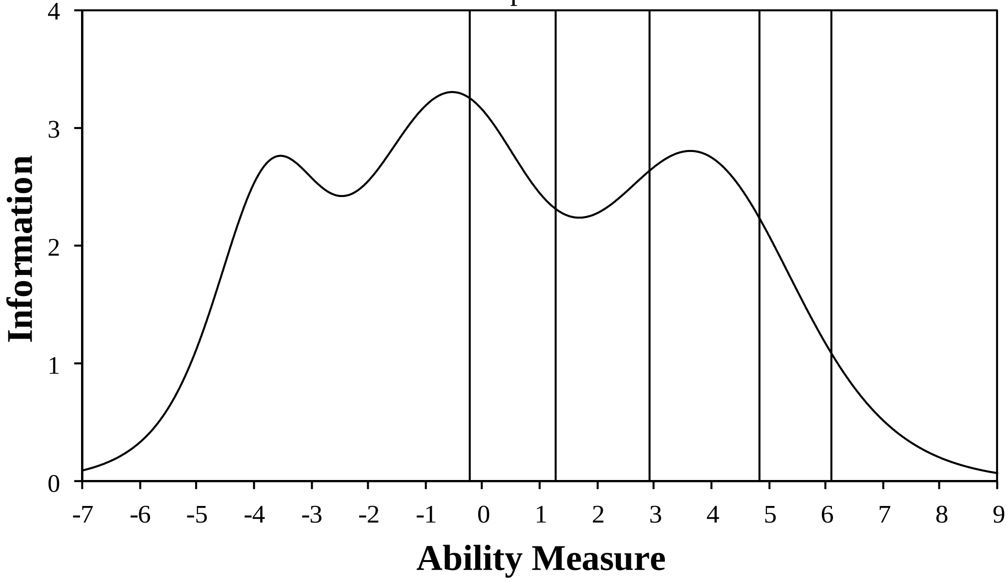
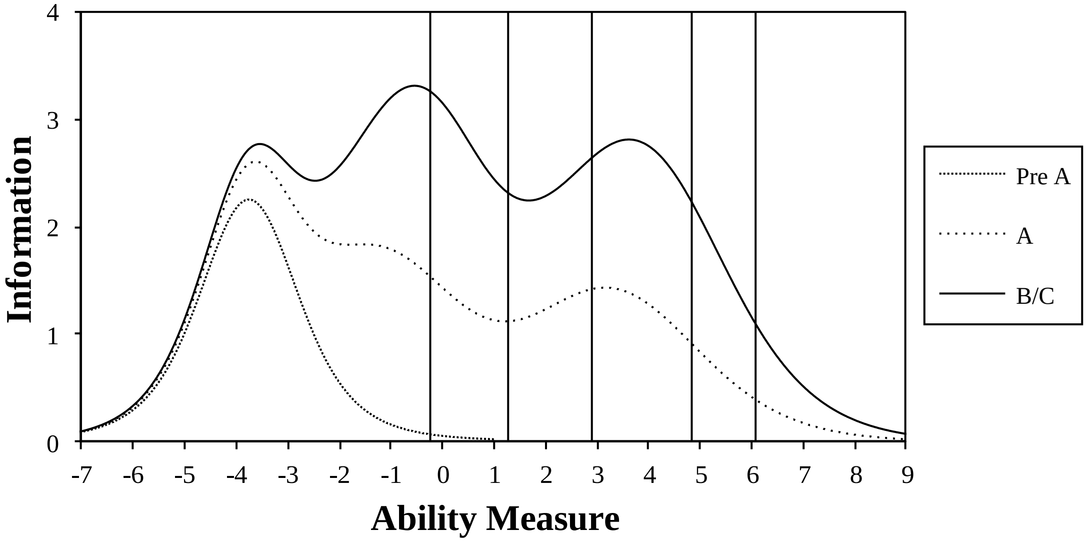


Figure 2.9.4.3.4
Test Information Function: Spek 4-5 S403 Online



2.9.4.4 *Grades 6–8*

Figure 2.9.4.4.1
Test Information Function: Spek 6-8 Pre-A S403 Online

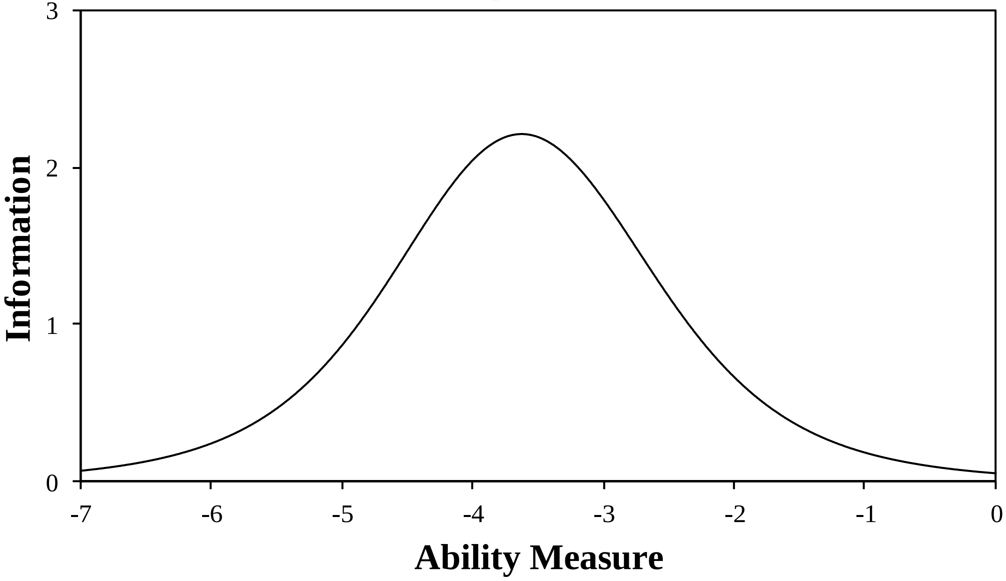


Figure 2.9.4.4.2
Test Information Function: Spek 6-8 A S403 Online

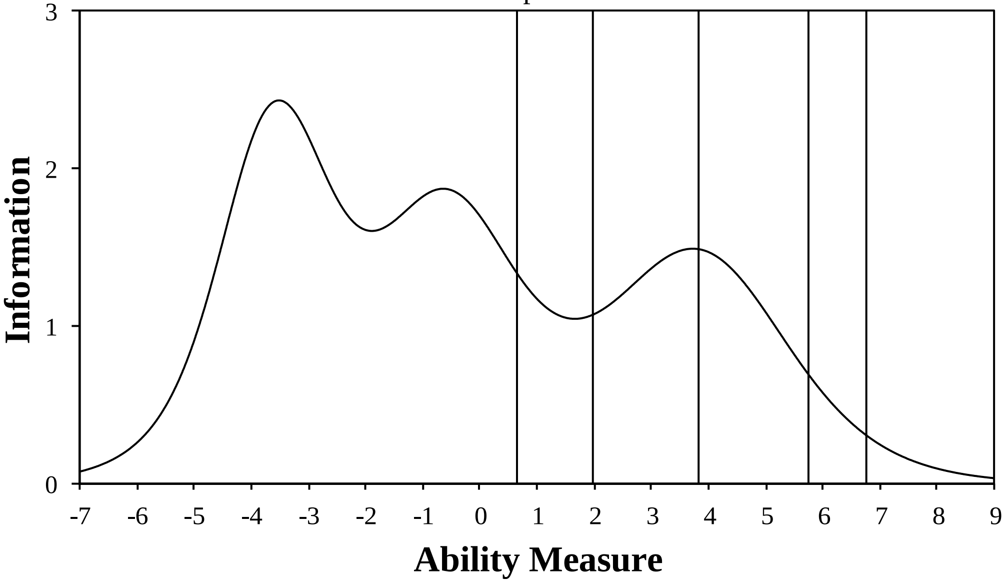


Figure 2.9.4.4.3
Test Information Function: Spek 6-8 B/C S403 Online

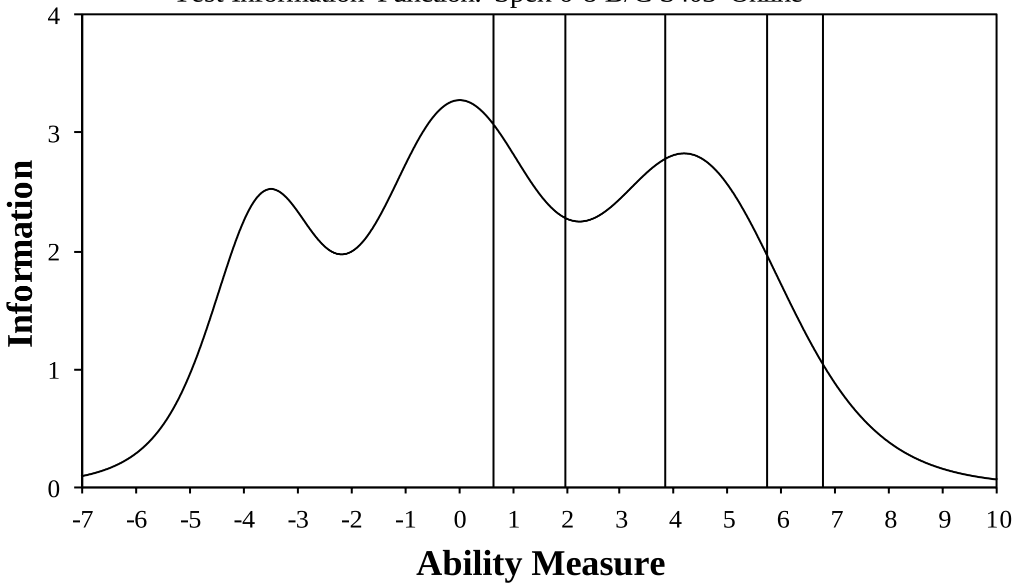
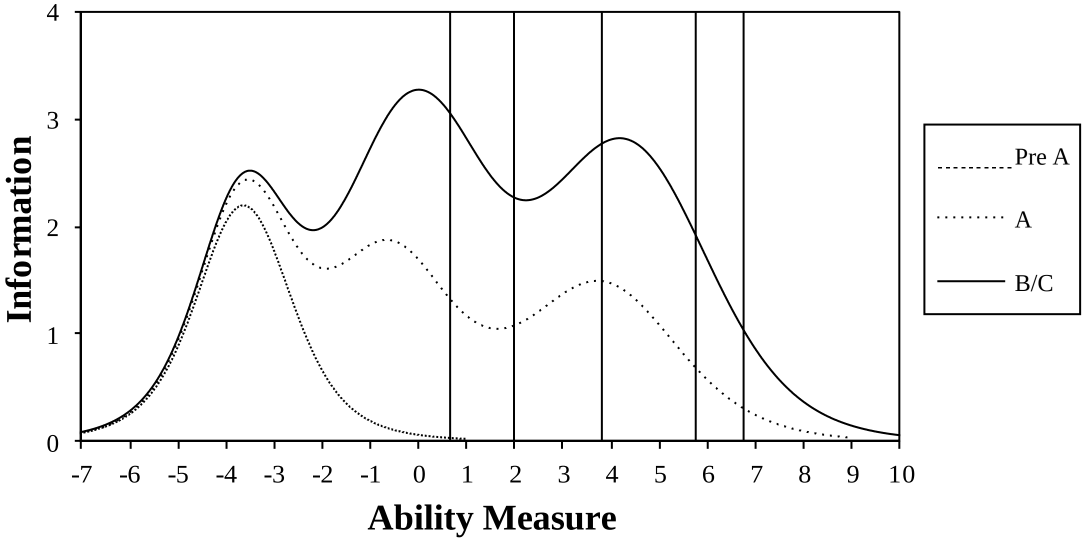


Figure 2.9.4.4.4
Test Information Function: Spek 6-8 S403 Online



2.9.4.5 *Grades 9–12*

Figure 2.9.4.5.1
Test Information Function: Spek 9-12 Pre-A S403 Online

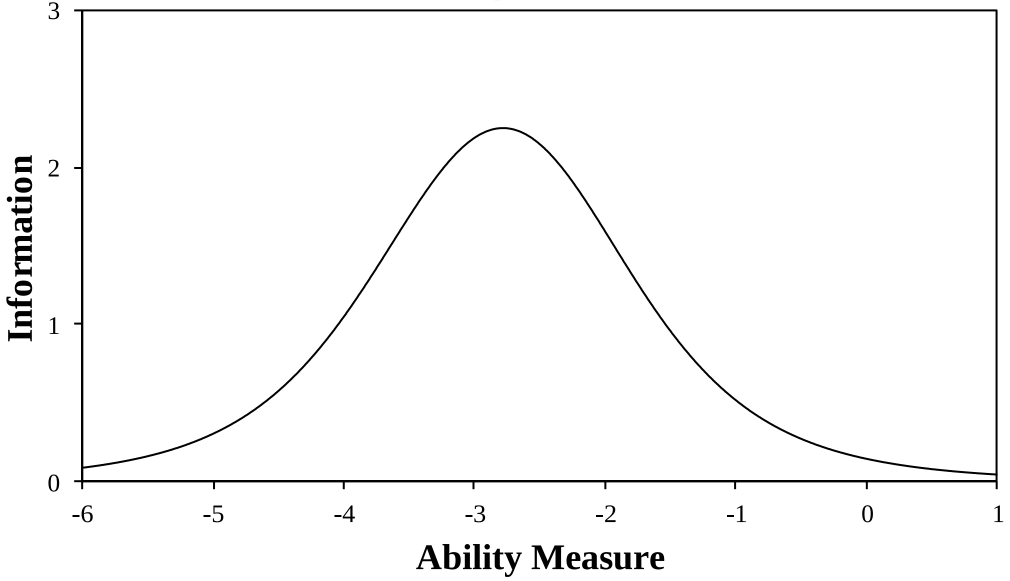


Figure 2.9.4.5.2
Test Information Function: Spek 9-12 A S403 Online

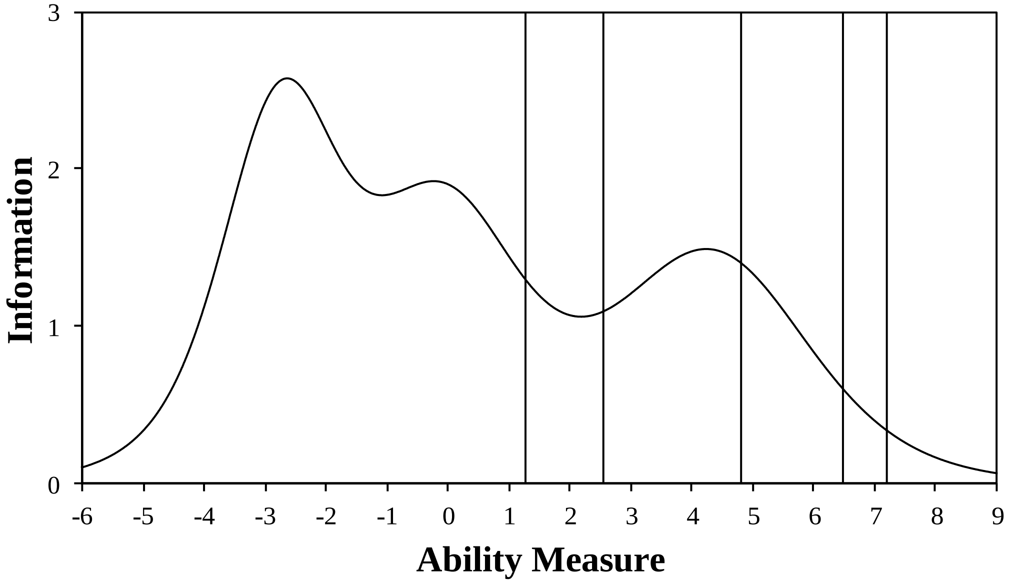


Figure 2.9.4.5.3
Test Information Function: Spek 9-12 B/C S403 Online

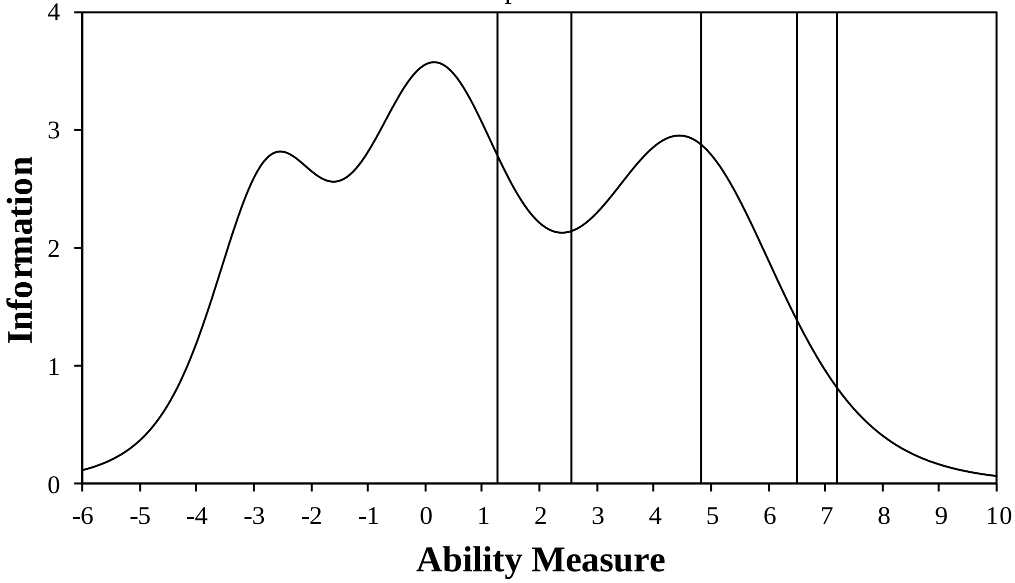
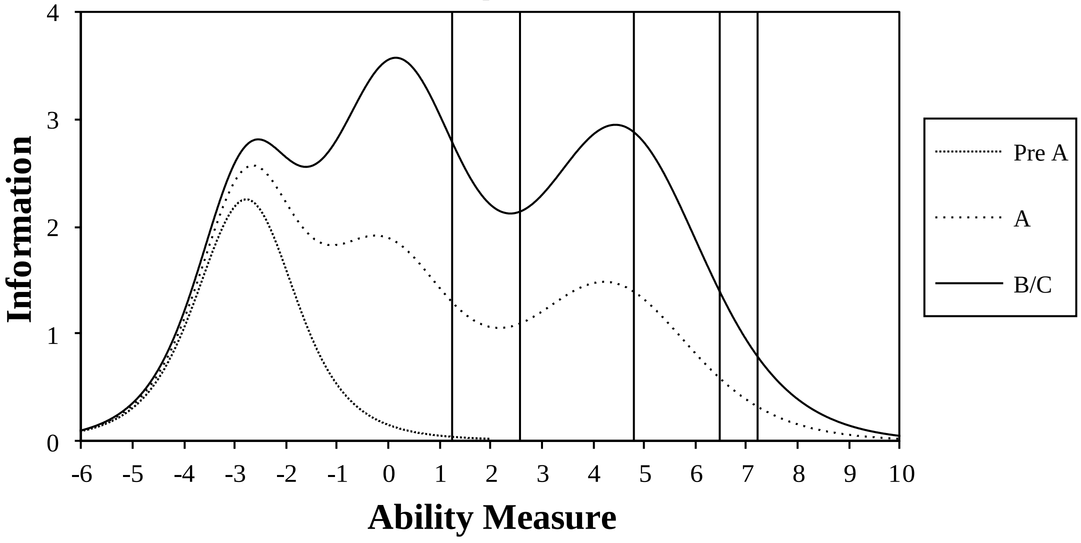


Figure 2.9.4.5.4
Test Information Function: Spek 9-12 S403 Online



2.10 Reliability

Listening and Reading

For tests administered using a multistage adaptive method, a reliability coefficient based on classical test theory such as Cronbach's coefficient alpha cannot be applied because not all students take the same set of items. Reliability for Listening and Reading was estimated using a method by Thissen (2000) by grade-level cluster:

$$\bar{\rho} = \frac{\sigma^2_{\theta} - average(CSEM^2_{observed})}{\sigma^2_{\theta}}$$

where

$\bar{\rho}$ is the average reliability
σ^2_{θ} is the variance of the distribution of student measure
$CSEM^2_{observed}$ is the squared observed conditional standard error of measurement for each student

This estimate is equivalent to the Rasch separation reliability coefficient (Linacre, 1999). Like Cronbach's alpha, the Rasch reliability coefficient is an estimate of the ratio of "true measure variance" to "observed measure variance." To obtain these values, item parameters and population student data were used as inputs in the Winsteps program. The Rasch separation reliability coefficient can be interpreted like Cronbach's coefficient alpha. It expresses how well the items on a test appear to measure the same construct.

Tables are provided below to express the Rasch separation reliability coefficient for ACCESS Online, for Listening and for Reading. For each domain, the first table provides the Rasch separation reliability coefficient for all test takers. Each row in the table represents a grade-level cluster, and values for the number of students, number of items, and reliability estimate are provided for each grade-level cluster. The second table for each domain provides the same information for the population of female students and the population of male students. The third table provides information by ethnicity, for Hispanic and non-Hispanic test takers, and the fourth table provides information for the population of students who have an individualized education plan (IEP).

Speaking and Writing

Cronbach's coefficient alpha is widely used as an estimate of reliability, particularly of the internal consistency of test items, and this statistic is appropriate for the fixed-form Speaking and Writing tests. Cronbach's alpha 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 examinees (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 sample of students tested as of the test. That is, the exact same test can produce widely disparate reliability indices based on the ability distribution of the group of examinees.

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

For the Writing test, a slight modification was made in the estimation of the Cronbach's alpha for tiered forms that have differential weighting across tasks. This modification is an attempt to take into account that some tasks are weighted more than others when deriving students' ability measure for these tiered forms. For Writing tasks with a weight greater than one, students' response to the tasks is replicated as a function of their weights. For example, the fourth task in Writing G1A is weighted three; therefore, students' response to this task is repeated three times when computing the Cronbach's alpha. This modification means that the number of pieces of information for Writing tasks that contribute to the estimation of the Cronbach's alpha for G1A is actually six, not four.

For Writing and Speaking, tables in this section also present the standard error of measurement (SEM) based on classical test theory. Unlike item response theory, 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 (SD) 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, with the assertion in the view of classical test theory 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. Statistically speaking, then, there is an expectation that an examinee's true score has a 68% probability of falling within the band extending from the observed score minus 1 SEM to the observed score plus 1 SEM.

Tables are provided below to express reliability information for Writing and Speaking. For each domain, the first table provides Cronbach's alpha and the SEM for all test takers. Each row in the table represents a test form, and values for the number of students, number of tasks, Cronbach's alpha, and SEM are provided for each test form. The second table for each domain provides the same information for the population of female students and the population of male students. The third table provides information by ethnicity, for Hispanic and non-Hispanic test takers, and the fourth table provides information for the population of students who have an IEP.

2.10.1 Listening

Table 2.10.1.1

Reliability: List S403 Online

Cluster	No. of Students	No. of Items	Rasch Reliability Estimate
1	181,107	54	0.85
2-3	372,664	54	0.84
4-5	322,964	54	0.81
6-8	277,787	54	0.85
9-12	301,765	54	0.86

Table 2.10.1.2

Reliability: List S403 Online by Gender

Cluster	No. of Items	Female		Male	
		No. of Students	Rasch Reliability Estimate	No. of Students	Rasch Reliability Estimate
1	54	84,584	0.84	95,256	0.86
2-3	54	173,798	0.83	196,719	0.84
4-5	54	146,422	0.80	174,527	0.81
6-8	54	120,267	0.85	155,344	0.86
9-12	54	130,894	0.86	167,760	0.86

Table 2.10.1.3

Reliability: List S403 Online by Ethnicity

Cluster	No. of Items	Hispanic		Other	
		No. of Students	Rasch Reliability Estimate	No. of Students	Rasch Reliability Estimate
1	54	115,237	0.85	60,698	0.85
2-3	54	246,502	0.84	116,392	0.83
4-5	54	221,969	0.80	89,207	0.80
6-8	54	186,367	0.85	78,708	0.85
9-12	54	196,149	0.86	91,582	0.86

Table 2.10.1.4

Reliability: List S403 Online by IEP status

Cluster	No. of Students	No. of Items	Rasch Reliability Estimate
1	14,826	54	0.89
2-3	35,854	54	0.87
4-5	40,540	54	0.81
6-8	46,517	54	0.83
9-12	39,698	54	0.83

2.10.2 Reading

Table 2.10.2.1

Reliability: Read S403 Online

Cluster	No. of Students	No. of Items	Rasch Reliability Estimate
1	183,421	72	0.89
2-3	373,371	72	0.88
4-5	318,641	69	0.89
6-8	277,957	72	0.91
9-12	297,819	72	0.91

Table 2.10.2.2

Reliability: Read S403 Online by Gender

Cluster	No. of Items	Female		Male	
		No. of Students	Rasch Reliability Estimate	No. of Students	Rasch Reliability Estimate
1	72	85,320	0.89	96,829	0.89
2-3	72	173,462	0.88	197,749	0.88
4-5	69	143,800	0.88	172,892	0.89
6-8	72	119,706	0.91	156,029	0.92
9-12	72	128,400	0.91	166,373	0.91

Table 2.10.2.3

Reliability: Read S403 Online by Ethnicity

Cluster	No. of Items	Hispanic		Other	
		No. of Students	Rasch Reliability Estimate	No. of Students	Rasch Reliability Estimate
1	72	116,808	0.87	61,402	0.90
2-3	72	247,022	0.88	116,661	0.88
4-5	69	218,993	0.88	88,005	0.89
6-8	72	186,901	0.91	78,222	0.92
9-12	72	194,471	0.91	89,384	0.91

Table 2.10.2.4

Reliability: Read S403 Online by IEP status

Cluster	No. of Students	No. of Items	Rasch Reliability Estimate
1	15,193	72	0.85
2-3	36,133	72	0.86
4-5	40,121	69	0.87
6-8	46,919	72	0.89
9-12	39,715	72	0.87

2.10.3 Writing

Table 2.10.3.1

Reliability: Writ S403 Online

Cluster	Tier	No. of Students	No. of Tasks	Cronbach's Alpha	SEM
1	A	154,550	4	0.865	1.804
	B/C	35,093	3	0.884	1.676
2-3	A	93,614	3	0.882	1.386
	B/C	296,663	3	0.916	1.622
4-5	A	44,209	3	0.909	1.305
	B/C	261,800	3	0.914	1.700
6-8	A	100,087	3	0.883	1.381
	B/C	168,607	3	0.901	1.651
9-12	A	108,872	3	0.889	1.557
	B/C	184,755	3	0.910	1.984

Table 2.10.3.2

Reliability: Writ S403 Online by Gender

			Female			Male		
Cluster	Tier	No. of Tasks	No. of Students	Cronbach's Alpha	SEM	No. of Students	Cronbach's Alpha	SEM
1	A	4	70,606	0.860	1.741	82,801	0.868	1.841
	B/C	3	17,888	0.877	1.662	17,033	0.888	1.690
2-3	A	3	39,244	0.880	1.368	53,589	0.881	1.398
	B/C	3	142,622	0.908	1.614	152,573	0.919	1.629
4-5	A	3	17,818	0.906	1.324	25,798	0.910	1.293
	B/C	3	120,260	0.905	1.678	140,215	0.919	1.717
6-8	A	3	39,185	0.878	1.409	59,839	0.882	1.362
	B/C	3	76,492	0.887	1.663	91,080	0.908	1.636
9-12	A	3	42,713	0.882	1.556	64,463	0.889	1.558
	B/C	3	83,610	0.902	1.942	99,810	0.914	2.015

Table 2.10.3.3

Reliability: Writ S403 Online by Ethnicity

Cluster	Tier	No. of Tasks	Hispanic			Other		
			No. of Students	Cronbach's Alpha	SEM	No. of Students	Cronbach's Alpha	SEM
1	A	4	103,777	0.866	1.829	46,409	0.859	1.719
	B/C	3	16,694	0.887	1.684	17,382	0.878	1.667
2-3	A	3	66,436	0.884	1.386	24,298	0.865	1.390
	B/C	3	191,471	0.916	1.621	97,936	0.914	1.627
4-5	A	3	29,792	0.909	1.307	11,393	0.883	1.330
	B/C	3	180,860	0.914	1.694	72,874	0.914	1.717
6-8	A	3	68,472	0.881	1.380	25,436	0.849	1.390
	B/C	3	112,276	0.902	1.610	50,132	0.897	1.732
9-12	A	3	73,714	0.888	1.558	27,984	0.860	1.568
	B/C	3	118,322	0.910	1.986	59,819	0.911	1.970

Table 2.10.3.4

Reliability: Writ S403 Online by IEP status

Cluster	Tier	No. of Students	No. of Tasks	Cronbach's Alpha	SEM
1	A	14,518	4	0.879	1.990
	B/C	1,148	3	0.926	1.795
2-3	A	18,131	3	0.877	1.418
	B/C	19,727	3	0.934	1.705
4-5	A	11,498	3	0.863	1.335
	B/C	26,859	3	0.929	1.794
6-8	A	23,824	3	0.831	1.372
	B/C	21,120	3	0.909	1.647
9-12	A	17,595	3	0.853	1.582
	B/C	21,674	3	0.912	2.012

2.10.4 Speaking

Table 2.10.4.1

Reliability: Spek S403 Online

Cluster	Tier	No. of Students	No. of Tasks	Cronbach's Alpha	SEM
1	Pre-A	5,158	3	0.848	0.838
	A	61,323	6	0.850	1.342
	B/C	111,043	6	0.822	1.584
2-3	Pre-A	16,359	3	0.868	0.699
	A	79,013	6	0.813	1.242
	B/C	270,495	6	0.780	1.534
4-5	Pre-A	6,011	3	0.828	0.883
	A	28,001	6	0.832	1.227
	B/C	279,646	6	0.815	1.466
6-8	Pre-A	11,494	3	0.835	0.756
	A	57,829	6	0.822	1.286
	B/C	205,516	6	0.819	1.477
9-12	Pre-A	23,102	3	0.864	0.602
	A	123,691	6	0.841	1.219
	B/C	152,516	6	0.844	1.390

Table 2.10.4.2

Reliability: Spek S403 Online by Gender

			Female			Male		
Cluster	Tier	No. of Tasks	No. of Students	Cronbach's Alpha	SEM	No. of Students	Cronbach's Alpha	SEM
1	Pre-A	3	1,958	0.863	0.786	3,118	0.836	0.868
	A	6	26,184	0.854	1.327	34,620	0.844	1.350
	B/C	6	54,800	0.823	1.579	55,610	0.818	1.586
2-3	Pre-A	3	6,684	0.874	0.688	9,410	0.864	0.700
	A	6	33,686	0.815	1.231	44,819	0.809	1.246
	B/C	6	130,200	0.778	1.539	138,957	0.776	1.527
4-5	Pre-A	3	2,489	0.847	0.843	3,326	0.815	0.907
	A	6	11,485	0.826	1.223	16,186	0.833	1.227
	B/C	6	128,031	0.812	1.464	150,187	0.816	1.467
6-8	Pre-A	3	4,772	0.837	0.739	6,412	0.833	0.761
	A	6	22,334	0.820	1.290	34,885	0.821	1.281
	B/C	6	91,268	0.823	1.493	113,004	0.817	1.460
9-12	Pre-A	3	8,932	0.839	0.599	13,686	0.875	0.598
	A	6	50,640	0.827	1.235	71,559	0.849	1.206
	B/C	6	69,582	0.844	1.402	81,807	0.844	1.376

Table 2.10.4.3

Reliability: Spek S403 Online by Ethnicity

Cluster	Tier	No. of Tasks	Hispanic			Other		
			No. of Students	Cronbach's Alpha	SEM	No. of Students	Cronbach's Alpha	SEM
1	Pre-A	3	3,461	0.846	0.852	1,448	0.847	0.789
	A	6	41,441	0.853	1.341	18,051	0.831	1.344
	B/C	6	67,860	0.822	1.574	40,169	0.818	1.597
2-3	Pre-A	3	11,341	0.871	0.707	4,201	0.848	0.641
	A	6	56,330	0.818	1.235	20,616	0.780	1.259
	B/C	6	174,272	0.784	1.522	89,533	0.769	1.556
4-5	Pre-A	3	3,781	0.819	0.899	1,216	0.839	0.767
	A	6	18,630	0.835	1.224	7,678	0.785	1.212
	B/C	6	193,139	0.815	1.453	77,670	0.810	1.496
6-8	Pre-A	3	7,510	0.824	0.768	2,329	0.862	0.586
	A	6	39,128	0.825	1.287	15,211	0.770	1.268
	B/C	6	137,924	0.821	1.461	59,981	0.808	1.509
9-12	Pre-A	3	15,676	0.865	0.616	5,449	0.858	0.491
	A	6	83,062	0.848	1.228	34,021	0.799	1.186
	B/C	6	96,026	0.852	1.388	51,027	0.818	1.392

Table 2.10.4.4

Reliability: Spek S403 Online by IEP status

Cluster	Tier	No. of Students	No. of Tasks	Cronbach's Alpha	SEM
1	Pre-A	984	3	0.813	0.835
	A	7,845	6	0.848	1.354
	B/C	5,780	6	0.814	1.612
2-3	Pre-A	3,728	3	0.847	0.606
	A	14,707	6	0.786	1.250
	B/C	17,029	6	0.789	1.536
4-5	Pre-A	617	3	0.810	0.769
	A	7,600	6	0.788	1.170
	B/C	31,143	6	0.822	1.457
6-8	Pre-A	1,519	3	0.856	0.611
	A	14,595	6	0.807	1.244
	B/C	30,284	6	0.819	1.460
9-12	Pre-A	2,728	3	0.900	0.584
	A	21,022	6	0.859	1.187
	B/C	16,025	6	0.854	1.371

2.11 Interrater Agreement

For the Writing and Speaking tests, tables provide information on interrater agreement for a sample of 20% of task raters. These tables show, for each of the tasks, the percentage of agreement between two raters. The first column shows the task and the second column shows the number of responses that were double scored. DRC selects a sample of 20% of all responses scored, chosen at random during the operational scoring process. The next columns shows the rates of agreement.

For Writing, with 0–6 as defined levels and the possibility of awarding a “plus” score between levels (e.g., 3, 3+, or 4 are all valid scores), scores that match or are contiguous (for example, if Rater 1 assigns a 3+ and Rater 2 assigns a score of 3, 3+, or 4) are categorized as agreement (%AG). Scores that are one whole score point apart (for example, if Rater 1 assigns a 3+ and Rater 2 assigns a score of 2+ or 4+) are categorized as adjacent (%AD). Otherwise, the raters are nonadjacent (%NA). Note that for Writing, interrater agreement is computed independently between ratings of keyboarded and handwritten responses.

For Speaking, when the two raters agree on the rating, an exact agreement is counted (%EX). If the two raters differ by one point, an adjacent agreement is counted (%AD). Otherwise, the raters are nonadjacent (%NA).

2.11.3 Writing

2.11.3.1 Grade 1

Table 2.11.3.1.1

Interrater Agreement: Writ 1 A S403 Online

Interrater Agreement	**Task**	**No. in Sample**	**% AG**	**% AD**	**% NA**
	1	68,258	100	0	0
	2	77,286	100	0	0
	3	83,970	98	2	0
	4	76,732	97	3	0

Table 2.11.3.1.2

Interrater Agreement: Writ 1 B/C S403 Online

Interrater Agreement	**Task**	**No. in Sample**	**% AG**	**% AD**	**% NA**
	1	14,338	97	3	0
	2	15,272	97	3	0
	3	14,668	98	2	0

2.11.3.2 Grades 2–3

Table 2.11.3.2.1

Interrater Agreement: Writ 2-3 A S403 Online

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	49,988	96	4	0
	2	49,716	97	3	0
	3	51,038	98	2	0

Table 2.11.3.2.2

Interrater Agreement: Writ 2-3 B/C S403 Online

Interrater Agreement	Task	No. in Sample	% AG	% AD	% NA
	1	121,760	97	3	0
	2	122,436	96	4	0
	3	122,716	94	6	0

2.11.3.3 Grades 4–5

Table 2.11.3.3.1

Interrater Agreement: Writ 4-5 A S403 Online

Interrater Agreement	Task	Mode of Response	No. in Sample	% AG	% AD	% NA
	1	HW	4,710	98	2	0
		KB	16,466	98	2	0
	2	HW	4,588	98	2	0
		KB	16,458	97	3	0
	3	HW	4,906	99	1	0
		KB	16,414	97	3	0

Table 2.11.3.3.2

Interrater Agreement: Writ 4-5 B/C S403 Online

Interrater Agreement	Task	Mode of Response	No. in Sample	% AG	% AD	% NA
	1	HW	18,664	97	3	0
		KB	104,616	97	3	0
	2	HW	18,474	98	2	0
		KB	104,838	97	3	0
	3	HW	19,190	97	3	0
		KB	105,028	98	2	0

2.11.3.4 *Grades 6–8*

Table 2.11.3.4.1

Interrater Agreement: Writ 6-8 A S403 Online

Interrater Agreement	Task	Mode of Response	No. in Sample	% AG	% AD	% NA
	1	HW	280	98	2	0
		KB	43,686	96	4	0
	2	HW	278	98	2	0
		KB	43,530	97	3	0
	3	HW	316	96	4	0
		KB	43,832	96	4	0

Table 2.11.3.4.2

Interrater Agreement: Writ 6-8 B/C S403 Online

Interrater Agreement	Task	Mode of Response	No. in Sample	% AG	% AD	% NA
	1	HW	310	99	1	0
		KB	75,854	98	2	0
	2	HW	294	95	5	0
		KB	77,040	97	3	0
	3	HW	324	94	6	0
		KB	76,644	97	3	0

2.11.3.5 *Grades 9–12*

Table 2.11.3.5.1

Interrater Agreement: Writ 9-12 A S403 Online

Interrater Agreement	Task	Mode of Response	No. in Sample	% AG	% AD	% NA
	1	HW	202	100	0	0
		KB	48,096	95	5	0
	2	HW	198	99	1	0
		KB	48,218	97	3	0
	3	HW	190	100	0	0
		KB	48,138	94	6	0

Table 2.11.3.5.2

Interrater Agreement: Writ 9-12 B/C S403 Online

Interrater Agreement	Task	Mode of Response	No. in Sample	% AG	% AD	% NA
	1	HW	158	100	0	0
		KB	84,810	97	3	0
	2	HW	154	99	1	0
		KB	85,376	97	3	0
	3	HW	136	97	3	0
		KB	87,820	95	5	0

2.11.4 Speaking

2.11.4.1 Grade 1

Table 2.11.4.1.1

Interrater Agreement: Spek 1 Pre-A S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	3,938	99	1	0
	2	3,894	98	2	0
	3	3,892	98	2	0

Table 2.11.4.1.2

Interrater Agreement: Spek 1 A S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	35,632	99	1	0
	2	35,634	92	8	0
	3	34,976	99	1	0
	4	34,994	87	13	0
	5	34,908	99	1	0
	6	34,908	90	10	0

Table 2.11.4.1.3

Interrater Agreement: Spek 1 B/C S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	55,196	89	11	0
	2	55,196	87	13	0
	3	57,562	85	15	0
	4	57,562	83	17	0
	5	56,560	89	11	0
	6	56,566	83	17	0

2.11.4.2 Grades 2–3

Table 2.11.4.2.1

Interrater Agreement: Spek 2-3 Pre-A S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	11,974	98	2	0
	2	11,270	99	1	0
	3	11,530	99	1	0

Table 2.11.4.2.2

Interrater Agreement: Spek 2-3 A S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	41,724	99	1	0
	2	41,724	85	15	0
	3	41,438	100	0	0
	4	41,438	89	11	0
	5	41,752	100	0	0
	6	41,752	87	12	0

Table 2.11.4.2.3

Interrater Agreement: Spek 2-3 B/C S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	133,362	79	20	0
	2	133,362	79	20	1
	3	129,550	85	15	0
	4	129,548	79	20	0
	5	130,100	88	12	0
	6	130,100	83	17	0

2.11.4.3 Grades 4–5

Table 2.11.4.3.1

Interrater Agreement: Spek 4-5 Pre-A S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	4,584	97	3	0
	2	5,136	99	1	0
	3	4,806	97	3	0

Table 2.11.4.3.2

Interrater Agreement: Spek 4-5 A S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	14,922	98	2	0
	2	14,922	83	17	0
	3	15,736	99	1	0
	4	15,736	90	10	0
	5	15,526	98	2	0
	6	15,526	90	10	0

Table 2.11.4.3.3

Interrater Agreement: Spek 4-5 B/C S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	136,240	79	21	0
	2	136,236	79	21	0
	3	136,386	81	18	0
	4	136,394	78	21	0
	5	133,242	79	21	0
	6	133,244	75	24	0

2.11.4.4 Grades 6–8

Table 2.11.4.4.1

Interrater Agreement: Spek 6-8 Pre-A S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	7,480	97	3	0
	2	7,724	98	2	0
	3	7,800	98	2	0

Table 2.11.4.4.2

Interrater Agreement: Spek 6-8 A S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	30,038	99	1	0
	2	30,042	86	14	0
	3	30,282	99	1	0
	4	30,282	87	13	0
	5	31,726	99	1	0
	6	31,726	89	11	0

Table 2.11.4.4.3

Interrater Agreement: Spek 6-8 B/C S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	104,696	84	16	0
	2	104,702	79	20	0
	3	104,522	83	17	0
	4	104,520	80	19	0
	5	103,326	84	16	0
	6	103,328	81	18	0

2.11.4.5 *Grades 9–12*

Table 2.11.4.5.1

Interrater Agreement: Spek 9-12 Pre-A S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	14,002	99	1	0
	2	13,804	99	1	0
	3	13,774	98	2	0

Table 2.11.4.5.2

Interrater Agreement: Spek 9-12 A S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	67,328	99	1	0
	2	67,328	88	12	0
	3	67,372	99	1	0
	4	67,372	82	17	1
	5	66,788	99	1	0
	6	66,788	86	13	0

Table 2.11.4.5.3

Interrater Agreement: Spek 9-12 B/C S403 Online

Interrater Agreement	Task	No. in Sample	% EX	% AD	% NA
	1	79,680	81	19	0
	2	79,680	83	17	0
	3	82,276	78	22	1
	4	82,260	77	23	1
	5	82,912	84	16	0
	6	82,912	77	22	0

2.12 Conditional Standard Errors of Measurement at Cut Score

The tables in this section present information on the conditional standard errors of measurement (CSEM) at the most important points at which decisions are made about students based on performance on ACCESS—the cut points between language proficiency levels. Because the cut points depend on the grade level, information is provided for each grade level within a grade-level cluster.

Since the Listening and Reading tests are multistage adaptive tests, the CSEM will vary for the same scale score since students are routed to take different items; therefore, it is not possible to present a single value for the CSEM of the scale score which corresponds to each cut score. In the tables for Listening and Reading, the leftmost column shows the proficiency level cut (e.g., 1/2, which is the cut between PL 1 and PL 2). The second column shows the grade level. The third column shows the cut score in the scale score metric (e.g., 305). The next columns present the mean, standard deviation, minimum, and maximum of the CSEM of all students at the cut scores. Note that there are some rare cases where there are no observed scale scores corresponding to the cut score values; therefore, these descriptive statistics cannot be provided.

For Writing and Speaking, the values are presented by tier. From these tables, it is possible to identify how well the different Writing and Speaking tiers are targeted for making decisions about students at the various proficiency level cuts. For example, Tier A is intended for students at the lowest end of the language proficiency continuum. Optimally, Tier A forms should have the lowest CSEM of any tier at the 1/2 proficiency level cut and a relatively low CSEM at the 2/3 proficiency level cut. At the other end of the continuum, Tier B/C forms should optimally have the lowest CSEM at the 5/6 proficiency level cut and a relatively low CSEM at the 4/5 proficiency level cut. These tables provide comparable information on how well the two tier forms are targeted to provide the most accurate measure in order to place their intended examinees into the language proficiency levels that they target. In the tables for Writing and Speaking, the leftmost column shows the proficiency level cut (e.g., 1/2, which is the cut between PL 1 and PL 2). The second column shows the grade level. The third column shows the cut score in the scale score metric (e.g., 305). In the last column(s), the corresponding CSEM is given for each cut score in the scale score metric for Writing and Speaking.

2.12.1 Listening

2.12.1.1 *Grade 1*

Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: List 1 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	No. of Students	Min.	Max.	Mean	Std. Dev.
1/2	1	236	N/A	N/A	N/A	N/A	N/A
2/3	1	259	43	16.84	17.86	17.15	0.47
3/4	1	291	55	17.35	18.88	18.47	0.53
4/5	1	303	385	17.35	19.90	17.64	0.79
5/6	1	327	740	17.86	21.94	18.06	0.42

2.12.1.2 *Grades 2–3*

Table 2.12.1.2

Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: List 2-3 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	No. of Students	Min.	Max.	Mean	Std. Dev.
1/2	2	245	N/A	N/A	N/A	N/A	N/A
	3	262	39	19.90	20.92	20.13	0.39
2/3	2	283	156	17.86	19.90	18.79	1.01
	3	300	379	17.86	19.90	18.12	0.60
3/4	2	314	263	18.88	21.43	18.90	0.18
	3	331	9	19.39	21.43	20.52	1.08
4/5	2	330	972	18.88	19.90	19.31	0.23
	3	349	333	20.41	20.41	20.41	0.00
5/6	2	354	15	21.43	26.02	23.88	2.37
	3	374	119	25.51	25.51	25.51	0.00

2.12.1.3 *Grades 4–5*

Table 2.12.1.3

Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: List 4-5 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	No. of Students	Min.	Max.	Mean	Std. Dev.
1/2	4	275	95	20.92	20.92	20.92	0.00
	5	285	N/A	N/A	N/A	N/A	N/A
2/3	4	313	558	17.35	18.37	17.36	0.09
	5	323	68	17.35	17.35	17.35	0.00
3/4	4	343	78	18.37	19.39	19.35	0.20
	5	354	1	19.90	19.90	19.90	0.00
4/5	4	363	309	18.37	19.90	18.72	0.63
	5	375	1,595	19.39	19.90	19.39	0.04
5/6	4	388	N/A	N/A	N/A	N/A	N/A
	5	401	58	20.41	20.41	20.41	0.00

2.12.1.4 *Grades 6–8*

Table 2.12.1.4

Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: List 6-8 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	No. of Students	Min.	Max.	Mean	Std. Dev.
1/2	6	294	N/A	N/A	N/A	N/A	N/A
	7	302	4	20.41	20.41	20.41	0.00
	8	308	163	19.90	19.90	19.90	0.00
2/3	6	332	N/A	N/A	N/A	N/A	N/A
	7	340	23	17.35	17.35	17.35	0.00
	8	347	347	16.33	17.86	16.48	0.46
3/4	6	363	N/A	N/A	N/A	N/A	N/A
	7	370	3	16.33	16.84	16.50	0.29
	8	377	235	16.84	17.35	16.97	0.23
4/5	6	385	892	16.33	17.35	17.27	0.20
	7	394	87	16.84	17.35	16.96	0.22
	8	402	132	16.84	18.37	17.79	0.40
5/6	6	411	424	16.84	17.86	17.07	0.30
	7	420	4,552	16.84	17.86	16.84	0.08
	8	427	N/A	N/A	N/A	N/A	N/A

2.12.1.5 Grades 9–12

Table 2.12.1.5

Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: List 9-12 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	No. of Students	Min.	Max.	Mean	Std. Dev.
1/2	9	314	N/A	N/A	N/A	N/A	N/A
	10	325	N/A	N/A	N/A	N/A	N/A
	11	335	N/A	N/A	N/A	N/A	N/A
	12	342	1,550	16.84	19.39	19.38	0.11
2/3	9	353	52	16.84	16.84	16.84	0.00
	10	358	37	16.84	16.84	16.84	0.00
	11	364	2,097	16.33	16.84	16.34	0.09
	12	368	4	16.33	16.84	16.71	0.26
3/4	9	383	158	16.33	17.35	16.68	0.24
	10	389	23	16.84	16.84	16.84	0.00
	11	394	709	16.33	16.84	16.74	0.20
	12	398	35	16.84	17.35	17.08	0.26
4/5	9	409	2,793	16.84	17.35	17.34	0.05
	10	415	273	17.35	17.86	17.57	0.25
	11	420	568	17.86	19.39	18.37	0.11
	12	426	1,095	17.86	18.88	18.10	0.37
5/6	9	434	26	19.39	19.39	19.39	0.00
	10	441	N/A	N/A	N/A	N/A	N/A
	11	447	122	20.41	20.41	20.41	0.00
	12	452	10	21.94	21.94	21.94	0.00

2.12.2 Reading

2.12.2.1 Grade 1

Table 2.12.2.1

Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: Read 1 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	No. of Students	Min.	Max.	Mean	Std. Dev.
1/2	1	264	2,134	10.71	12.24	12.05	0.40
2/3	1	286	651	9.69	10.71	9.82	0.25
3/4	1	304	1,032	9.69	10.20	9.89	0.25
4/5	1	315	57	9.69	10.20	9.93	0.26
5/6	1	334	3,169	10.20	11.22	10.20	0.02

2.12.2.2 Grades 2–3

Table 2.12.2.2

Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: Read 2-3 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	No. of Students	Min.	Max.	Mean	Std. Dev.
1/2	2	283	258	10.71	12.24	11.09	0.26
	3	297	789	10.20	10.71	10.67	0.14
2/3	2	307	801	9.69	10.71	10.12	0.24
	3	323	1,790	9.69	10.20	9.97	0.25
3/4	2	326	593	9.69	10.20	9.88	0.25
	3	342	769	9.69	10.71	10.28	0.38
4/5	2	337	6,581	9.69	10.20	10.20	0.02
	3	352	363	10.71	11.73	10.78	0.21
5/6	2	355	29	11.22	11.73	11.68	0.16
	3	370	67	12.76	12.76	12.76	0.00

2.12.2.3 Grades 4–5

Table 2.12.2.3

Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: Read 4-5 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	No. of Students	Min.	Max.	Mean	Std. Dev.
1/2	4	307	388	11.22	13.27	11.81	0.24
	5	316	965	10.71	12.24	12.05	0.41
2/3	4	335	981	10.20	11.22	10.70	0.13
	5	345	1,353	10.20	10.71	10.50	0.25
3/4	4	354	412	10.20	10.71	10.30	0.20
	5	364	408	10.20	10.71	10.47	0.26
4/5	4	364	318	10.20	10.71	10.25	0.14
	5	373	5,269	10.20	11.22	10.24	0.12
5/6	4	382	7,268	10.20	10.71	10.21	0.02
	5	391	5,491	10.71	11.73	10.72	0.03

2.12.2.4 *Grades 6–8*

Table 2.12.2.4

Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: Read 6-8 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	No. of Students	Min.	Max.	Mean	Std. Dev.
1/2	6	323	582	11.22	11.73	11.24	0.08
	7	329	75	11.73	12.24	11.90	0.24
	8	335	171	11.73	12.24	11.88	0.23
2/3	6	353	1,149	10.71	11.73	10.75	0.15
	7	360	3,361	10.20	12.24	10.25	0.22
	8	366	451	10.20	12.24	10.34	0.43
3/4	6	373	574	10.20	11.22	10.56	0.29
	7	380	309	10.20	11.73	10.73	0.23
	8	386	966	10.20	13.78	10.41	0.37
4/5	6	382	139	10.20	11.22	10.78	0.24
	7	389	85	10.20	11.73	11.04	0.48
	8	395	319	10.71	12.24	11.21	0.47
5/6	6	399	174	10.71	10.71	10.71	0.00
	7	406	1,574	10.71	10.71	10.71	0.00
	8	412	25	11.73	12.76	11.78	0.20

2.12.2.5 *Grades 9–12*

Table 2.12.2.5

Descriptive Statistics of Conditional Standard Error of Measurement at Cut Scores: Read 9-12 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	No. of Students	Min.	Max.	Mean	Std. Dev.
1/2	9	340	268	11.22	12.24	11.26	0.18
	10	344	295	11.22	11.73	11.30	0.19
	11	348	371	11.22	12.24	11.27	0.15
	12	352	249	11.22	11.73	11.26	0.13
2/3	9	372	3,150	10.20	11.22	10.25	0.15
	10	377	2,366	10.20	11.22	10.24	0.16
	11	382	730	10.20	11.22	10.26	0.21
	12	386	625	10.20	11.73	10.34	0.33
3/4	9	392	1,220	10.20	10.71	10.23	0.11
	10	397	535	10.20	11.22	10.37	0.26
	11	402	137	10.20	11.73	10.59	0.35
	12	407	89	10.20	12.24	10.99	0.41
4/5	9	401	1,997	10.20	11.73	10.25	0.15
	10	406	218	10.20	12.24	10.43	0.33
	11	410	92	10.20	11.73	10.85	0.34
	12	414	214	10.71	12.24	10.85	0.37
5/6	9	418	2,030	10.71	11.22	10.72	0.04
	10	423	42	11.73	12.24	12.22	0.11
	11	427	2,252	11.22	13.27	11.23	0.08
	12	432	1,370	11.73	13.27	11.74	0.05

2.12.3 Writing

2.12.3.1 *Grade 1*

Table 2.12.3.1

Conditional Standard Error of Measurement at Cut Scores: Writ 1 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	**Tier B/C**
1/2	1	238	10.31	8.86
2/3	1	275	14.23	11.55
3/4	1	337	13.43	12.08
4/5	1	382	13.16	11.28
5/6	1	405	16.65	13.43

2.12.3.2 *Grades 2–3*

Table 2.12.3.2

Conditional Standard Error of Measurement at Cut Scores: Writ 2-3 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	**Tier B/C**
1/2	2	242	11.55	9.64
	3	247	12.08	8.86
2/3	2	279	16.38	9.67
	3	283	16.65	10.20
3/4	2	341	17.45	12.62
	3	346	17.18	12.62
4/5	2	388	15.31	11.01
	3	394	15.60	10.74
5/6	2	411	18.26	11.01
	3	418	20.14	11.33

2.12.3.3 Grades 4–5

Table 2.12.3.3

Conditional Standard Error of Measurement at Cut Scores: Writ 4-5 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	4	266	11.55	11.01
	5	267	11.55	10.71
2/3	4	288	13.69	8.32
	5	293	14.50	8.32
3/4	4	351	17.99	12.46
	5	356	17.72	12.62
4/5	4	401	15.57	11.81
	5	407	15.31	11.55
5/6	4	425	15.84	11.01
	5	433	16.92	10.74

2.12.3.4 Grades 6–8

Table 2.12.3.4

Conditional Standard Error of Measurement at Cut Scores: Writ 6-8 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	6	268	12.35	8.59
	7	273	12.62	8.30
	8	281	13.43	8.32
2/3	6	298	15.57	9.91
	7	305	16.38	10.74
	8	311	16.92	11.14
3/4	6	361	17.18	12.62
	7	367	16.92	12.51
	8	372	16.65	12.35
4/5	6	413	15.84	10.82
	7	419	16.38	10.74
	8	424	16.92	10.74
5/6	6	441	20.41	12.08
	7	450	23.63	13.43
	8	459	27.66	15.31

2.12.3.5 *Grades 9–12*

Table 2.12.3.5

Conditional Standard Error of Measurement at Cut Scores: Writ 9-12 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	9	289	11.81	8.59
	10	298	12.08	8.65
	11	308	13.16	9.40
	12	318	14.77	10.47
2/3	9	319	14.77	10.47
	10	326	15.84	11.04
	11	335	16.65	11.76
	12	344	17.18	12.08
3/4	9	378	17.72	12.62
	10	385	17.45	12.35
	11	391	17.45	12.35
	12	398	16.94	11.98
4/5	9	430	15.31	11.01
	10	436	15.33	10.93
	11	441	15.57	11.01
	12	447	15.84	11.28
5/6	9	469	20.14	14.23
	10	479	23.90	16.65
	11	490	29.27	20.41
	12	501	35.71	24.97

2.12.4 Speaking

2.12.4.1 Grade 1

Table 2.12.4.1

Conditional Standard Error of Measurement at Cut Scores: Spek 1 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	**Tier B/C**
1/2	1	205	21.06	15.21
2/3	1	261	28.37	20.18
3/4	1	311	23.98	17.26
4/5	1	361	29.83	20.18
5/6	1	403	50.31	32.76

Note: Tier Pre-A is not presented as it is not possible for Tier Pre-A students to receive a proficiency level higher than 2.

2.12.4.2 Grades 2–3

Table 2.12.4.2

Conditional Standard Error of Measurement at Cut Scores: Spek 2-3 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	**Tier B/C**
1/2	2	220	23.98	16.67
	3	234	26.32	17.84
2/3	2	273	27.49	19.30
	3	283	26.32	18.72
3/4	2	322	23.98	17.55
	3	332	24.57	17.84
4/5	2	374	34.80	22.81
	3	386	40.36	25.74
5/6	2	415	62.01	37.44
	3	425	72.83	43.58

Note: Tier Pre-A is not presented as it is not possible for Tier Pre-A students to receive a proficiency level higher than 2.

2.12.4.3 *Grades 4–5*

Table 2.12.4.3

Conditional Standard Error of Measurement at Cut Scores: Spek 4-5 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	4	246	22.52	16.09
	5	258	23.69	16.09
2/3	4	293	27.49	18.72
	5	302	27.49	19.30
3/4	4	342	24.86	18.43
	5	350	24.57	17.84
4/5	4	397	28.37	18.72
	5	407	31.00	19.60
5/6	4	435	42.41	25.45
	5	443	47.09	28.08

Note: Tier Pre-A is not presented as it is not possible for Tier Pre-A students to receive a proficiency level higher than 2.

2.12.4.4 *Grades 6–8*

Table 2.12.4.4

Conditional Standard Error of Measurement at Cut Scores: Spek 6-8 S403 Online

Proficiency Level Cut Point	Grade	Cut Score	SEM	
			Tier A	Tier B/C
1/2	6	268	22.81	16.09
	7	277	24.28	16.38
	8	284	25.45	16.67
2/3	6	310	28.66	18.72
	7	317	28.37	19.01
	8	323	28.08	19.30
3/4	6	360	24.28	18.43
	7	369	23.98	17.84
	8	377	23.98	17.55
4/5	6	417	29.54	19.01
	7	425	32.46	19.89
	8	433	35.10	21.06
5/6	6	451	44.75	25.15
	7	457	48.84	26.91
	8	463	53.52	29.25

Note: Tier Pre-A is not presented as it is not possible for Tier Pre-A students to receive a proficiency level higher than 2.

2.12.4.5 *Grades 9–12*

Table 2.12.4.5

Conditional Standard Error of Measurement at Cut Scores: Spek 9-12 S403 Online

Proficiency Level Cut Point	**Grade**	**Cut Score**	**SEM**	
			Tier A	**Tier B/C**
1/2	9	290	23.98	16.38
	10	295	24.57	16.96
	11	299	25.45	17.26
	12	302	25.74	17.55
2/3	9	328	28.37	19.89
	10	333	28.37	20.18
	11	337	28.08	20.18
	12	340	27.79	19.89
3/4	9	385	23.98	17.26
	10	393	23.98	16.96
	11	400	24.28	16.96
	12	406	24.86	17.26
4/5	9	440	32.46	21.35
	10	446	34.51	22.81
	11	451	36.85	23.98
	12	455	38.61	25.15
5/6	9	468	46.21	29.54
	10	471	48.26	30.71
	11	474	50.31	32.17
	12	476	51.77	33.05

Note: Tier Pre-A is not presented as it is not possible for Tier Pre-A students to receive a proficiency level higher than 2.

2.13 Accuracy and Consistency

One of the main purposes of the WIDA ACCESS program is to identify the English language proficiency level of students with respect to the WIDA ELD Standards. Because of the emphasis on the classification of student performance, a psychometric property of interest is how accurately and consistently ACCESS can classify students into WIDA proficiency categories. The analyses utilize the methods outlined by 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).

Classification accuracy is defined conceptually as the extent to which the proficiency classifications of students based on the assessment would agree with those made on the basis of their true scores. True scores are, of course, unknown. A common practice is to estimate the true scores based on observed scores and the parameters of the model used in estimating the true scores. Classification consistency is defined conceptually as the extent to which the proficiency classifications of students agree given two independent administrations of the same or two parallel test forms. However, it is impractical to obtain repeated administrations of the same or parallel test forms because of cost, testing burden, and effects of student memory and practice. A common approach is to estimate classification consistency from a single administration of a test.

Tables in this section present three pieces of information related to the accuracy and consistency of placement into the WIDA language proficiency levels for each domain. A separate table is provided for each grade in a grade-level cluster.

The first section provides overall indices related to the accuracy and consistency of classification, as well as Cohen’s kappa. Overall accuracy indicates the percentage of all students who would be classified into the same language proficiency level by both the administered test and the true score distribution. For example, an overall accuracy of 0.774 means that an estimated 77% of students would be classified into the correct performance level according to observed and true scores. Overall consistency indicates the percentage of all students who would be classified into the same language proficiency level by both the administered test and by a parallel test. For example, an overall classification consistency of 0.664 means that an estimated 66% of students would be classified into the same performance level if two parallel forms were administered. Cohen’s kappa assesses the proportion of consistency classifications after removing the proportion of consistent classifications that would be expected by chance. Due to this correction, Cohen’s kappa values tend to be lower than overall consistency.

The second section of the table shows accuracy and consistency information conditional on proficiency level. The conditional accuracy indices examine the percentage of students classified by both tests into a proficiency level, divided by all students classified into that proficiency level according to the true score distribution. For example, an accuracy of 0.774 at PL 3 means that an estimated 77% of students are classified into PL 3 when their true scores suggest that they would be in PL 3. The conditional consistency indices examine the percentage of students classified by both tests into a proficiency level, divided by all students classified into that proficiency level according to a parallel test. For example, a classification consistency of 0.664 at PL 4 means that an estimated 66% of students would be classified into PL 4 when a parallel form suggests that these students would be classified in PL 4.

The third section of the table provides indices of classification accuracy, including the false-positives and the false-negatives, 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 (e.g., determining which students have reached PL 5). For example, to evaluate the degree of confidence that one can have in a decision made based on the domain scores as to whether or not students are being accurately (consistently) classified into PL 5 ("Bridging"), one can look at the accuracy (consistency) index provided in the table for the cut score 4/5. Note that the accuracy and consistency are generally higher at the cut points than over the proficiency levels.

In terms of classifications around a single cut point, students can be misclassified in one of two ways. Students who are below the proficiency level cut score (based on their "true score") but are classified based on the observed score as being above the cut score are considered to be false-positives. Students who are above the proficiency level cut score (based on their "true score") but are classified as being below a cut score based on the observed score are considered to be false-negatives. All other students are considered to be accurately placed either above or below the cut score. The accuracy indices at the cut score examine the percentage of students who are accurately placed above and below the cut score.

The consistency indices at the cut examine the percentage of students classified consistently above and below the cut score.

In several cases no test takers were placed into the proficiency level, and accuracy and consistency of classification conditional on that level cannot be computed. Also, when the number of test takers placed into the proficiency level is smaller than 200, BB-CLASS produces estimates that are out of bounds (for example, a negative estimated percentage of students placed in the proficiency level). In these cases, 'N/A' has been placed in the table. In addition, there are a few cases where, due to the small percentage of test takers placed into the proficiency level and the range of observed scale scores, accuracy of classification conditional on that level cannot be estimated by BB-CLASS. In such cases, a hyphen (-) has been placed in the table.

The approach taken by Livingston and Lewis (1995) and implemented here uses information about the reliability of the test, the cut scores, and the observed distribution of scores. Then, using a four-parameter beta distribution, the distribution of the true scores and of scores on a parallel form is modeled. The Livingston and Lewis procedure requires that the reliability estimate of the test form be provided in estimating the classification consistency and accuracy statistics. For Listening and Reading, the Rasch reliability estimates by grade-level clusters were used in the procedure. Since the Writing and Speaking tests were tiered, it was necessary to produce a single reliability estimate across tiers for the Livingston and Lewis procedure. This is a weighted reliability estimate across tiers (see Section 2.10).

2.13.1 Listening

Table 2.13.1.1

Accuracy and Consistency of Classification Indices: List (Grade 1) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.704	0.640		0.429	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.857		0.732	
	2	0.313		0.223	
	3	0.423		0.306	
	4	0.215		0.152	
	5	0.391		0.294	
	6	0.892		0.858	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.962	0.010	0.027	0.949
	2/3	0.953	0.017	0.030	0.931
	3/4	0.925	0.036	0.040	0.891
	4/5	0.908	0.042	0.050	0.870
	5/6	0.881	0.057	0.062	0.834

Table 2.13.1.2

Accuracy and Consistency of Classification Indices: List (Grade 2) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.576	0.491		0.357	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.846		0.688	
	2	0.383		0.280	
	3	0.553		0.431	
	4	0.182		0.138	
	5	0.428		0.331	
	6	0.818		0.730	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.951	0.012	0.036	0.932
	2/3	0.915	0.045	0.039	0.876
	3/4	0.872	0.050	0.078	0.825
	4/5	0.870	0.070	0.059	0.818
	5/6	0.876	0.065	0.059	0.826

Table 2.13.1.3

Accuracy and Consistency of Classification Indices: List (Grade 3) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.611	0.529		0.372	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.820		0.666	
	2	0.422		0.314	
	3	0.517		0.394	
	4	0.180		0.131	
	5	0.400		0.300	
	6	0.855		0.788	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.958	0.013	0.029	0.941
	2/3	0.926	0.034	0.040	0.895
	3/4	0.896	0.036	0.068	0.855
	4/5	0.890	0.059	0.051	0.840
	5/6	0.862	0.078	0.060	0.810

Table 2.13.1.4

Accuracy and Consistency of Classification Indices: List (Grade 4) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.808	0.758		0.419	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.917		0.771	
	2	0.279		0.189	
	3	0.522		0.352	
	4	0.254		0.175	
	5	0.410		0.296	
	6	0.926		0.911	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.988	0.001	0.010	0.986
	2/3	0.984	0.007	0.008	0.974
	3/4	0.959	0.013	0.028	0.942
	4/5	0.942	0.025	0.033	0.914
	5/6	0.902	0.041	0.057	0.862

Table 2.13.1.5

Accuracy and Consistency of Classification Indices: List (Grade 5) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.798	0.745		0.417	
Conditional on Level	Level	Accuracy		Consistency	
	1	0.902		0.766	
	2	0.306		0.214	
	3	0.450		0.299	
	4	0.250		0.170	
	5	0.403		0.289	
	6	0.921		0.906	
Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.985	0.002	0.013	0.981
	2/3	0.980	0.008	0.012	0.969
	3/4	0.959	0.015	0.026	0.939
	4/5	0.940	0.028	0.032	0.911
	5/6	0.897	0.044	0.060	0.855

Table 2.13.1.6

Accuracy and Consistency of Classification Indices: List (Grade 6) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.641	0.546		0.395	
Conditional on Level	Level	Accuracy		Consistency	
	1	0.880		0.717	
	2	0.450		0.312	
	3	0.635		0.497	
	4	0.438		0.339	
	5	0.488		0.383	
	6	0.853		0.781	
Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.985	0.002	0.013	0.981
	2/3	0.964	0.018	0.018	0.944
	3/4	0.906	0.036	0.058	0.872
	4/5	0.886	0.049	0.065	0.839
	5/6	0.876	0.069	0.055	0.826

Table 2.13.1.7

Accuracy and Consistency of Classification Indices: List (Grade 7) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.611	0.518		0.378	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.861		0.666	
	2	0.399		0.277	
	3	0.600		0.476	
	4	0.437		0.344	
	5	0.406		0.310	
	6	0.874		0.795	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.976	0.004	0.020	0.968
	2/3	0.947	0.028	0.025	0.918
	3/4	0.883	0.045	0.072	0.844
	4/5	0.882	0.051	0.068	0.832
	5/6	0.888	0.070	0.042	0.843

Table 2.13.1.8

Accuracy and Consistency of Classification Indices: List (Grade 8) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.605	0.517		0.376	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.832		0.648	
	2	0.429		0.308	
	3	0.550		0.432	
	4	0.405		0.314	
	5	0.367		0.277	
	6	0.875		0.799	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.970	0.007	0.023	0.958
	2/3	0.935	0.032	0.033	0.904
	3/4	0.886	0.044	0.070	0.847
	4/5	0.884	0.052	0.064	0.834
	5/6	0.884	0.071	0.044	0.837

Table 2.13.1.9

Accuracy and Consistency of Classification Indices: List (Grade 9) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.583	0.478		0.362	
Conditional on Level	Level	Accuracy		Consistency	
	1	0.779		0.578	
	2	0.506		0.396	
	3	0.580		0.473	
	4	0.490		0.385	
	5	0.443		0.332	
	6	0.855		0.746	
Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.950	0.011	0.039	0.932
	2/3	0.903	0.053	0.044	0.861
	3/4	0.874	0.062	0.064	0.828
	4/5	0.902	0.052	0.047	0.858
	5/6	0.929	0.047	0.025	0.899

Table 2.13.1.10

Accuracy and Consistency of Classification Indices: List (Grade 10) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.576	0.474		0.363	
Conditional on Level	Level	Accuracy		Consistency	
	1	0.799		0.636	
	2	0.433		0.332	
	3	0.557		0.449	
	4	0.485		0.380	
	5	0.433		0.324	
	6	0.859		0.750	
Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.934	0.018	0.048	0.909
	2/3	0.901	0.053	0.046	0.859
	3/4	0.877	0.063	0.061	0.832
	4/5	0.903	0.050	0.047	0.861
	5/6	0.930	0.046	0.024	0.900

Table 2.13.1.11

Accuracy and Consistency of Classification Indices: List (Grade 11) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.569	0.467		0.356	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.841		0.693	
	2	0.331		0.243	
	3	0.593		0.480	
	4	0.435		0.342	
	5	0.471		0.355	
	6	0.827		0.705	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.932	0.018	0.049	0.905
	2/3	0.902	0.059	0.039	0.858
	3/4	0.870	0.060	0.070	0.824
	4/5	0.896	0.055	0.048	0.852
	5/6	0.930	0.043	0.027	0.899

Table 2.13.1.12

Accuracy and Consistency of Classification Indices: List (Grade 12) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.571	0.469		0.359	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.790		0.649	
	2	0.403		0.304	
	3	0.537		0.431	
	4	0.489		0.383	
	5	0.444		0.333	
	6	0.842		0.719	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.926	0.028	0.046	0.894
	2/3	0.894	0.055	0.051	0.850
	3/4	0.878	0.061	0.061	0.832
	4/5	0.905	0.052	0.044	0.863
	5/6	0.932	0.045	0.023	0.904

2.13.2 Reading

Table 2.13.2.1

Accuracy and Consistency of Classification Indices: Read (Grade 1) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.614	0.508		0.396	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.751		0.639	
	2	0.634		0.526	
	3	0.519		0.414	
	4	0.415		0.308	
	5	0.513		0.385	
	6	0.855		0.742	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.910	0.047	0.043	0.873
	2/3	0.885	0.058	0.057	0.840
	3/4	0.902	0.055	0.043	0.865
	4/5	0.929	0.036	0.036	0.897
	5/6	0.961	0.025	0.014	0.944

Table 2.13.2.2

Accuracy and Consistency of Classification Indices: Read (Grade 2) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.609	0.500		0.388	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.793		0.653	
	2	0.638		0.528	
	3	0.579		0.467	
	4	0.334		0.251	
	5	0.619		0.494	
	6	0.775		0.632	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.941	0.022	0.037	0.915
	2/3	0.894	0.051	0.055	0.852
	3/4	0.894	0.054	0.052	0.850
	4/5	0.904	0.059	0.037	0.867
	5/6	0.947	0.030	0.022	0.923

Table 2.13.2.3

Accuracy and Consistency of Classification Indices: Read (Grade 3) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.601	0.497		0.389	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.770		0.645	
	2	0.627		0.518	
	3	0.499		0.389	
	4	0.287		0.213	
	5	0.603		0.481	
	6	0.792		0.667	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.927	0.032	0.041	0.897
	2/3	0.894	0.051	0.056	0.851
	3/4	0.898	0.056	0.046	0.855
	4/5	0.901	0.061	0.038	0.863
	5/6	0.939	0.033	0.028	0.912

Table 2.13.2.4

Accuracy and Consistency of Classification Indices: Read (Grade 4) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.622	0.521		0.415	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.824		0.687	
	2	0.665		0.547	
	3	0.537		0.423	
	4	0.344		0.260	
	5	0.575		0.466	
	6	0.828		0.726	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.963	0.012	0.025	0.948
	2/3	0.920	0.035	0.045	0.887
	3/4	0.896	0.054	0.050	0.854
	4/5	0.892	0.059	0.049	0.850
	5/6	0.917	0.048	0.035	0.883

Table 2.13.2.5

Accuracy and Consistency of Classification Indices: Read (Grade 5) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.619	0.520		0.415	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.830		0.710	
	2	0.647		0.532	
	3	0.539		0.425	
	4	0.248		0.184	
	5	0.552		0.440	
	6	0.822		0.719	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.952	0.018	0.031	0.932
	2/3	0.909	0.043	0.048	0.872
	3/4	0.896	0.050	0.053	0.853
	4/5	0.897	0.060	0.044	0.855
	5/6	0.917	0.048	0.034	0.884

Table 2.13.2.6

Accuracy and Consistency of Classification Indices: Read (Grade 6) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.678	0.584		0.465	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.907		0.854	
	2	0.625		0.517	
	3	0.655		0.548	
	4	0.303		0.225	
	5	0.514		0.389	
	6	0.674		0.452	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.929	0.028	0.043	0.901
	2/3	0.907	0.057	0.036	0.870
	3/4	0.913	0.043	0.044	0.879
	4/5	0.934	0.035	0.031	0.905
	5/6	0.968	0.023	0.009	0.954

Table 2.13.2.7

Accuracy and Consistency of Classification Indices: Read (Grade 7) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.678	0.587		0.470	
Conditional on Level	Level	Accuracy		Consistency	
	1	0.906		0.853	
	2	0.594		0.486	
	3	0.647		0.540	
	4	0.287		0.211	
	5	0.457		0.334	
	6	0.797		0.634	
Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.926	0.030	0.045	0.895
	2/3	0.905	0.061	0.034	0.869
	3/4	0.916	0.041	0.043	0.884
	4/5	0.938	0.032	0.030	0.909
	5/6	0.963	0.025	0.012	0.948

Table 2.13.2.8

Accuracy and Consistency of Classification Indices: Read (Grade 8) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.670	0.578		0.461	
Conditional on Level	Level	Accuracy		Consistency	
	1	0.892		0.836	
	2	0.623		0.513	
	3	0.582		0.467	
	4	0.251		0.186	
	5	0.561		0.432	
	6	0.715		0.552	
Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.923	0.034	0.043	0.892
	2/3	0.908	0.054	0.038	0.872
	3/4	0.917	0.043	0.040	0.883
	4/5	0.927	0.044	0.029	0.898
	5/6	0.959	0.022	0.018	0.940

Table 2.13.2.9

Accuracy and Consistency of Classification Indices: Read (Grade 9) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.675	0.581		0.479	
Conditional on Level	Level	Accuracy		Consistency	
	1	0.872		0.802	
	2	0.661		0.559	
	3	0.604		0.487	
	4	0.267		0.196	
	5	0.540		0.417	
	6	0.833		0.726	
Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.926	0.030	0.045	0.896
	2/3	0.913	0.050	0.037	0.878
	3/4	0.922	0.041	0.037	0.890
	4/5	0.931	0.041	0.028	0.903
	5/6	0.952	0.029	0.019	0.933

Table 2.13.2.10

Accuracy and Consistency of Classification Indices: Read (Grade 10) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.657	0.561		0.459	
Conditional on Level	Level	Accuracy		Consistency	
	1	0.872		0.794	
	2	0.665		0.563	
	3	0.589		0.476	
	4	0.289		0.218	
	5	0.573		0.449	
	6	0.786		0.661	
Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.933	0.025	0.042	0.907
	2/3	0.911	0.051	0.038	0.875
	3/4	0.910	0.049	0.041	0.875
	4/5	0.919	0.048	0.034	0.887
	5/6	0.951	0.026	0.023	0.929

Table 2.13.2.11

Accuracy and Consistency of Classification Indices: Read (Grade 11) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.661	0.567		0.467	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.859		0.773	
	2	0.698		0.599	
	3	0.546		0.436	
	4	0.304		0.226	
	5	0.502		0.389	
	6	0.843		0.737	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.934	0.025	0.040	0.908
	2/3	0.913	0.045	0.042	0.877
	3/4	0.912	0.050	0.038	0.878
	4/5	0.921	0.042	0.037	0.889
	5/6	0.945	0.034	0.021	0.922

Table 2.13.2.12

Accuracy and Consistency of Classification Indices: Read (Grade 12) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.671	0.576		0.472	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.862		0.781	
	2	0.703		0.607	
	3	0.566		0.459	
	4	0.321		0.240	
	5	0.526		0.409	
	6	0.839		0.717	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.930	0.027	0.042	0.903
	2/3	0.912	0.047	0.041	0.876
	3/4	0.913	0.053	0.034	0.882
	4/5	0.925	0.039	0.036	0.895
	5/6	0.957	0.027	0.016	0.938

2.13.3 Writing

Table 2.13.3.1

Accuracy and Consistency of Classification Indices: Writ (Grade 1) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.749	0.665		0.460	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.782		0.676	
	2	0.864		0.793	
	3	0.548		0.468	
	4	-		-	
	5	N/A		N/A	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.924	0.045	0.031	0.888
	2/3	0.831	0.037	0.132	0.781
	3/4	0.992	0.008	0.000	0.994
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Table 2.13.3.2

Accuracy and Consistency of Classification Indices: Writ (Grade 2) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.816	0.754		0.518	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.828		0.727	
	2	0.780		0.672	
	3	0.824		0.821	
	4	-		-	
	5	N/A		N/A	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.976	0.013	0.012	0.964
	2/3	0.937	0.025	0.038	0.911
	3/4	0.904	0.096	0.000	0.878
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Table 2.13.3.3

Accuracy and Consistency of Classification Indices: Writ (Grade 3) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.790	0.714		0.507	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.799		0.707	
	2	0.750		0.631	
	3	0.855		0.795	
	4	0.659		0.565	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.981	0.010	0.009	0.974
	2/3	0.950	0.022	0.028	0.931
	3/4	0.861	0.058	0.081	0.812
	4/5	0.994	0.006	0.000	0.995
	5/6	N/A	N/A	N/A	N/A

Table 2.13.3.4

Accuracy and Consistency of Classification Indices: Writ (Grade 4) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.775	0.698		0.494	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.818		0.720	
	2	0.614		0.471	
	3	0.858		0.733	
	4	0.725		0.687	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.986	0.009	0.006	0.978
	2/3	0.969	0.014	0.017	0.957
	3/4	0.835	0.038	0.127	0.772
	4/5	0.985	0.015	0.000	0.985
	5/6	N/A	N/A	N/A	N/A

Table 2.13.3.5

Accuracy and Consistency of Classification Indices: Writ (Grade 5) S403 Online

<table>
<tr><th rowspan="2">Overall Indices</th><th>Accuracy</th><th colspan="2">Consistency</th><th colspan="2">Kappa (k)</th></tr>
<tr><td>0.799</td><td colspan="2">0.729</td><td colspan="2">0.552</td></tr>
<tr><th rowspan="7">Conditional on Level</th><th>Level</th><th colspan="2">Accuracy</th><th colspan="2">Consistency</th></tr>
<tr><td>1</td><td colspan="2">0.783</td><td colspan="2">0.674</td></tr>
<tr><td>2</td><td colspan="2">0.686</td><td colspan="2">0.549</td></tr>
<tr><td>3</td><td colspan="2">0.851</td><td colspan="2">0.743</td></tr>
<tr><td>4</td><td colspan="2">0.779</td><td colspan="2">0.747</td></tr>
<tr><td>5</td><td colspan="2">-</td><td colspan="2">-</td></tr>
<tr><td>6</td><td colspan="2">N/A</td><td colspan="2">N/A</td></tr>
<tr><th rowspan="7">Indices at Proficiency Level Cut Points</th><th rowspan="2">Proficiency Level Cut Point</th><th colspan="3">Accuracy</th><th rowspan="2">Consistency</th></tr>
<tr><th>Accuracy</th><th>False Positives</th><th>False Negatives</th></tr>
<tr><td>1/2</td><td>0.986</td><td>0.009</td><td>0.005</td><td>0.978</td></tr>
<tr><td>2/3</td><td>0.966</td><td>0.014</td><td>0.020</td><td>0.953</td></tr>
<tr><td>3/4</td><td>0.871</td><td>0.034</td><td>0.095</td><td>0.819</td></tr>
<tr><td>4/5</td><td>0.976</td><td>0.024</td><td>0.000</td><td>0.975</td></tr>
<tr><td>5/6</td><td>N/A</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
</table>

Table 2.13.3.6

Accuracy and Consistency of Classification Indices: Writ (Grade 6) S403 Online

<table>
<tr><th rowspan="2">Overall Indices</th><th>Accuracy</th><th colspan="2">Consistency</th><th colspan="2">Kappa (k)</th></tr>
<tr><td>0.774</td><td colspan="2">0.690</td><td colspan="2">0.510</td></tr>
<tr><th rowspan="7">Conditional on Level</th><th>Level</th><th colspan="2">Accuracy</th><th colspan="2">Consistency</th></tr>
<tr><td>1</td><td colspan="2">0.830</td><td colspan="2">0.723</td></tr>
<tr><td>2</td><td colspan="2">0.704</td><td colspan="2">0.582</td></tr>
<tr><td>3</td><td colspan="2">0.800</td><td colspan="2">0.740</td></tr>
<tr><td>4</td><td colspan="2">0.729</td><td colspan="2">0.620</td></tr>
<tr><td>5</td><td colspan="2">N/A</td><td colspan="2">N/A</td></tr>
<tr><td>6</td><td colspan="2">N/A</td><td colspan="2">N/A</td></tr>
<tr><th rowspan="7">Indices at Proficiency Level Cut Points</th><th rowspan="2">Proficiency Level Cut Point</th><th colspan="3">Accuracy</th><th rowspan="2">Consistency</th></tr>
<tr><th>Accuracy</th><th>False Positives</th><th>False Negatives</th></tr>
<tr><td>1/2</td><td>0.967</td><td>0.015</td><td>0.018</td><td>0.952</td></tr>
<tr><td>2/3</td><td>0.934</td><td>0.025</td><td>0.041</td><td>0.909</td></tr>
<tr><td>3/4</td><td>0.872</td><td>0.070</td><td>0.058</td><td>0.826</td></tr>
<tr><td>4/5</td><td>N/A</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
<tr><td>5/6</td><td>N/A</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
</table>

Table 2.13.3.7

Accuracy and Consistency of Classification Indices: Writ (Grade 7) S403 Online

<table>
<tr><th rowspan="2">Overall Indices</th><th>Accuracy</th><th colspan="2">Consistency</th><th colspan="2">Kappa (k)</th></tr>
<tr><td>0.769</td><td colspan="2">0.683</td><td colspan="2">0.533</td></tr>
<tr><th rowspan="7">Conditional on Level</th><th>Level</th><th colspan="2">Accuracy</th><th colspan="2">Consistency</th></tr>
<tr><td>1</td><td colspan="2">0.830</td><td colspan="2">0.725</td></tr>
<tr><td>2</td><td colspan="2">0.715</td><td colspan="2">0.601</td></tr>
<tr><td>3</td><td colspan="2">0.763</td><td colspan="2">0.695</td></tr>
<tr><td>4</td><td colspan="2">0.793</td><td colspan="2">0.697</td></tr>
<tr><td>5</td><td colspan="2">N/A</td><td colspan="2">N/A</td></tr>
<tr><td>6</td><td colspan="2">N/A</td><td colspan="2">N/A</td></tr>
<tr><th rowspan="7">Indices at Proficiency Level Cut Points</th><th rowspan="2">Proficiency Level Cut Point</th><th colspan="3">Accuracy</th><th rowspan="2">Consistency</th></tr>
<tr><th>Accuracy</th><th>False Positives</th><th>False Negatives</th></tr>
<tr><td>1/2</td><td>0.959</td><td>0.019</td><td>0.023</td><td>0.940</td></tr>
<tr><td>2/3</td><td>0.926</td><td>0.026</td><td>0.048</td><td>0.897</td></tr>
<tr><td>3/4</td><td>0.884</td><td>0.069</td><td>0.046</td><td>0.841</td></tr>
<tr><td>4/5</td><td>N/A</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
<tr><td>5/6</td><td>N/A</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
</table>

Table 2.13.3.8

Accuracy and Consistency of Classification Indices: Writ (Grade 8) S403 Online

<table>
<tr><th rowspan="2">Overall Indices</th><th>Accuracy</th><th colspan="2">Consistency</th><th colspan="2">Kappa (k)</th></tr>
<tr><td>0.768</td><td colspan="2">0.681</td><td colspan="2">0.526</td></tr>
<tr><th rowspan="7">Conditional on Level</th><th>Level</th><th colspan="2">Accuracy</th><th colspan="2">Consistency</th></tr>
<tr><td>1</td><td colspan="2">0.867</td><td colspan="2">0.780</td></tr>
<tr><td>2</td><td colspan="2">0.638</td><td colspan="2">0.513</td></tr>
<tr><td>3</td><td colspan="2">0.792</td><td colspan="2">0.723</td></tr>
<tr><td>4</td><td colspan="2">0.743</td><td colspan="2">0.643</td></tr>
<tr><td>5</td><td colspan="2">-</td><td colspan="2">-</td></tr>
<tr><td>6</td><td colspan="2">N/A</td><td colspan="2">N/A</td></tr>
<tr><th rowspan="7">Indices at Proficiency Level Cut Points</th><th rowspan="2">Proficiency Level Cut Point</th><th colspan="3">Accuracy</th><th rowspan="2">Consistency</th></tr>
<tr><th>Accuracy</th><th>False Positives</th><th>False Negatives</th></tr>
<tr><td>1/2</td><td>0.958</td><td>0.017</td><td>0.025</td><td>0.941</td></tr>
<tr><td>2/3</td><td>0.930</td><td>0.029</td><td>0.041</td><td>0.901</td></tr>
<tr><td>3/4</td><td>0.882</td><td>0.064</td><td>0.055</td><td>0.837</td></tr>
<tr><td>4/5</td><td>0.997</td><td>0.003</td><td>0.000</td><td>0.996</td></tr>
<tr><td>5/6</td><td>N/A</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
</table>

Table 2.13.3.9

Accuracy and Consistency of Classification Indices: Writ (Grade 9) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.783	0.698		0.554	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.842		0.740	
	2	0.677		0.551	
	3	0.814		0.737	
	4	0.769		0.703	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.967	0.015	0.018	0.953
	2/3	0.937	0.027	0.036	0.911
	3/4	0.889	0.047	0.064	0.844
	4/5	0.989	0.011	0.000	0.986
	5/6	N/A	N/A	N/A	N/A

Table 2.13.3.10

Accuracy and Consistency of Classification Indices: Writ (Grade 10) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.799	0.721		0.572	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.844		0.737	
	2	0.680		0.554	
	3	0.829		0.766	
	4	0.790		0.723	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.972	0.012	0.016	0.960
	2/3	0.943	0.024	0.033	0.920
	3/4	0.895	0.050	0.055	0.852
	4/5	0.989	0.011	0.000	0.986
	5/6	N/A	N/A	N/A	N/A

Table 2.13.3.11

Accuracy and Consistency of Classification Indices: Writ (Grade 11) S403 Online

<table>
<tr><td rowspan="2">Overall Indices</td><td>Accuracy</td><td colspan="2">Consistency</td><td colspan="2">Kappa (k)</td></tr>
<tr><td>0.799</td><td colspan="2">0.720</td><td colspan="2">0.575</td></tr>
<tr><td rowspan="7">Conditional on Level</td><td>Level</td><td colspan="2">Accuracy</td><td colspan="2">Consistency</td></tr>
<tr><td>1</td><td colspan="2">0.860</td><td colspan="2">0.761</td></tr>
<tr><td>2</td><td colspan="2">0.623</td><td colspan="2">0.495</td></tr>
<tr><td>3</td><td colspan="2">0.826</td><td colspan="2">0.764</td></tr>
<tr><td>4</td><td colspan="2">0.809</td><td colspan="2">0.737</td></tr>
<tr><td>5</td><td colspan="2">-</td><td colspan="2">-</td></tr>
<tr><td>6</td><td colspan="2">N/A</td><td colspan="2">N/A</td></tr>
<tr><td rowspan="7">Indices at Proficiency Level Cut Points</td><td rowspan="2">Proficiency Level Cut Point</td><td colspan="3">Accuracy</td><td rowspan="2">Consistency</td></tr>
<tr><td>Accuracy</td><td>False Positives</td><td>False Negatives</td></tr>
<tr><td>1/2</td><td>0.968</td><td>0.012</td><td>0.020</td><td>0.954</td></tr>
<tr><td>2/3</td><td>0.941</td><td>0.026</td><td>0.033</td><td>0.916</td></tr>
<tr><td>3/4</td><td>0.896</td><td>0.051</td><td>0.053</td><td>0.854</td></tr>
<tr><td>4/5</td><td>0.995</td><td>0.005</td><td>0.000</td><td>0.993</td></tr>
<tr><td>5/6</td><td>N/A</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
</table>

Table 2.13.3.12

Accuracy and Consistency of Classification Indices: Writ (Grade 12) S403 Online

<table>
<tr><td rowspan="2">Overall Indices</td><td>Accuracy</td><td colspan="2">Consistency</td><td colspan="2">Kappa (k)</td></tr>
<tr><td>0.776</td><td colspan="2">0.691</td><td colspan="2">0.542</td></tr>
<tr><td rowspan="7">Conditional on Level</td><td>Level</td><td colspan="2">Accuracy</td><td colspan="2">Consistency</td></tr>
<tr><td>1</td><td colspan="2">0.856</td><td colspan="2">0.766</td></tr>
<tr><td>2</td><td colspan="2">0.642</td><td colspan="2">0.518</td></tr>
<tr><td>3</td><td colspan="2">0.787</td><td colspan="2">0.715</td></tr>
<tr><td>4</td><td colspan="2">0.789</td><td colspan="2">0.702</td></tr>
<tr><td>5</td><td colspan="2">N/A</td><td colspan="2">N/A</td></tr>
<tr><td>6</td><td colspan="2">N/A</td><td colspan="2">N/A</td></tr>
<tr><td rowspan="7">Indices at Proficiency Level Cut Points</td><td rowspan="2">Proficiency Level Cut Point</td><td colspan="3">Accuracy</td><td rowspan="2">Consistency</td></tr>
<tr><td>Accuracy</td><td>False Positives</td><td>False Negatives</td></tr>
<tr><td>1/2</td><td>0.960</td><td>0.018</td><td>0.023</td><td>0.942</td></tr>
<tr><td>2/3</td><td>0.933</td><td>0.027</td><td>0.040</td><td>0.906</td></tr>
<tr><td>3/4</td><td>0.883</td><td>0.063</td><td>0.054</td><td>0.837</td></tr>
<tr><td>4/5</td><td>N/A</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
<tr><td>5/6</td><td>N/A</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
</table>

2.13.4 Speaking

Table 2.13.4.1

Accuracy and Consistency of Classification Indices: Spek (Grade 1) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.681	0.576		0.387	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.781		0.659	
	2	0.793		0.674	
	3	0.578		0.527	
	4	-		-	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.938	0.034	0.028	0.907
	2/3	0.843	0.041	0.116	0.787
	3/4	0.900	0.100	0.000	0.861
	4/5	0.997	0.003	0.000	0.997
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.2

Accuracy and Consistency of Classification Indices: Spek (Grade 2) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.669	0.564		0.342	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.799		0.673	
	2	0.759		0.606	
	3	0.577		0.530	
	4	-		-	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.941	0.030	0.029	0.911
	2/3	0.785	0.050	0.165	0.716
	3/4	0.941	0.059	0.000	0.909
	4/5	0.998	0.002	0.000	0.998
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.3

Accuracy and Consistency of Classification Indices: Spek (Grade 3) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.643	0.549		0.332	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.828		0.716	
	2	0.746		0.578	
	3	0.540		0.502	
	4	-		-	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.940	0.029	0.031	0.911
	2/3	0.762	0.038	0.199	0.698
	3/4	0.939	0.061	0.000	0.906
	4/5	0.998	0.002	0.000	0.998
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.4

Accuracy and Consistency of Classification Indices: Spek (Grade 4) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.638	0.541		0.321	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.765		0.623	
	2	0.708		0.557	
	3	0.642		0.573	
	4	0.539		0.444	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.967	0.017	0.016	0.949
	2/3	0.897	0.037	0.066	0.855
	3/4	0.784	0.118	0.098	0.730
	4/5	0.986	0.014	0.000	0.985
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.5

Accuracy and Consistency of Classification Indices: Spek (Grade 5) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.626	0.528		0.319	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.769		0.640	
	2	0.711		0.558	
	3	0.579		0.555	
	4	-		-	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.952	0.026	0.022	0.928
	2/3	0.869	0.037	0.094	0.821
	3/4	0.804	0.196	0.000	0.757
	4/5	0.991	0.009	0.000	0.991
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.6

Accuracy and Consistency of Classification Indices: Spek (Grade 6) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.646	0.546		0.334	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.811		0.701	
	2	0.708		0.528	
	3	0.596		0.569	
	4	-		-	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.950	0.026	0.024	0.926
	2/3	0.857	0.032	0.110	0.802
	3/4	0.838	0.162	0.000	0.787
	4/5	0.998	0.002	0.000	0.998
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.7

Accuracy and Consistency of Classification Indices: Spek (Grade 7) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.633	0.532		0.329	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.807		0.700	
	2	0.688		0.508	
	3	0.567		0.534	
	4	-		-	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.935	0.035	0.030	0.904
	2/3	0.818	0.034	0.147	0.758
	3/4	0.876	0.124	0.000	0.824
	4/5	0.996	0.004	0.000	0.996
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.8

Accuracy and Consistency of Classification Indices: Spek (Grade 8) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.680	0.566		0.369	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.883		0.796	
	2	0.522		0.350	
	3	0.643		0.612	
	4	-		-	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.931	0.028	0.042	0.900
	2/3	0.842	0.042	0.116	0.772
	3/4	0.898	0.102	0.000	0.841
	4/5	0.997	0.003	0.000	0.997
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.9

Accuracy and Consistency of Classification Indices: Spek (Grade 9) S403 Online

Overall Indices	Accuracy	Consistency	Kappa (k)
	0.688	0.602	0.418

Conditional on Level	Level	Accuracy	Consistency
	1	0.872	0.790
	2	0.645	0.498
	3	0.589	0.540
	4	-	-
	5	N/A	N/A
	6	N/A	N/A

Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.915	0.039	0.046	0.878
	2/3	0.799	0.044	0.157	0.738
	3/4	0.967	0.033	0.000	0.954
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.10

Accuracy and Consistency of Classification Indices: Spek (Grade 10) S403 Online

Overall Indices	Accuracy	Consistency	Kappa (k)
	0.688	0.602	0.414

Conditional on Level	Level	Accuracy	Consistency
	1	0.848	0.758
	2	0.680	0.537
	3	0.594	0.544
	4	-	-
	5	N/A	N/A
	6	N/A	N/A

Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.914	0.043	0.043	0.874
	2/3	0.790	0.043	0.167	0.732
	3/4	0.979	0.021	0.000	0.968
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.11

Accuracy and Consistency of Classification Indices: Spek (Grade 11) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.721	0.630		0.447	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.885		0.805	
	2	0.591		0.441	
	3	0.679		0.630	
	4	-		-	
	5	N/A		N/A	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.918	0.035	0.048	0.881
	2/3	0.811	0.049	0.140	0.744
	3/4	0.986	0.014	0.000	0.977
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

Table 2.13.4.12

Accuracy and Consistency of Classification Indices: Spek (Grade 12) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.714	0.621		0.431	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.900		0.814	
	2	0.537		0.391	
	3	0.673		0.629	
	4	-		-	
	5	N/A		N/A	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.920	0.030	0.050	0.882
	2/3	0.799	0.053	0.149	0.726
	3/4	0.985	0.015	0.000	0.979
	4/5	N/A	N/A	N/A	N/A
	5/6	N/A	N/A	N/A	N/A

3 Analyses of Composite Scores

Four composite scores are calculated for ACCESS Online: Oral Language, Literacy, Comprehension, and Overall. Composite scores are calculated as weighted averages of domain scale scores, as follows:

- Oral Language: 50% Listening + 50% Speaking
- Literacy: 50% Reading + 50% Writing
- Comprehension: 30% Listening + 70% Reading
- Overall Composite: 15% Listening + 15% Speaking + 35% Reading + 35% Writing

This weighting resulted from a policy decision by the WIDA Board before the first operational administration of ACCESS, based on the view that literacy skills are paramount in developing academic language proficiency.

3.1 Scale Score Distribution for Composites

Figures and tables in this section provide scale score distributions for each of the composites, for each grade-level cluster.

For each cluster, the figure shows the distribution of the scale scores for the composite. Scale scores are plotted on the horizontal axis, grouped into units of five scale score points (e.g., 100-104, 105-109, 110-114, etc.). The number of students with scale scores falling into each range is plotted on the vertical axis.

Each table shows, by grade and by total for the grade-level cluster:

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

3.1.1 Oral

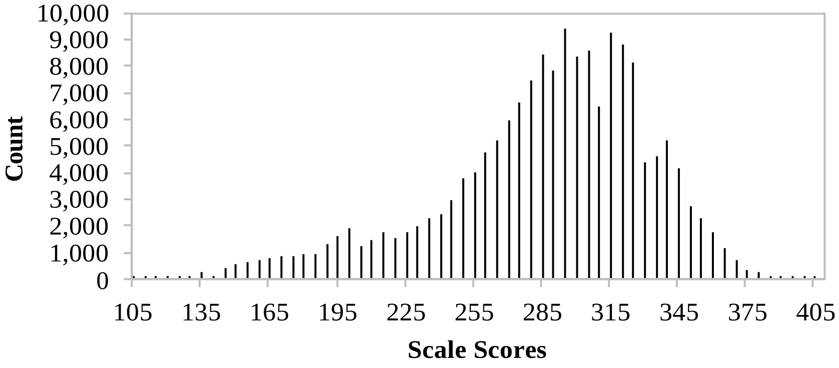
3.1.1.1 *Grade 1*

Table 3.1.1.1

Scale Score Descriptive Statistics: Oral 1 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	169,891	105	416	288.50	45.91
Total	169,891	105	416	288.50	45.91

Figure 3.1.1.1
Scale Scores: Oral 1 S403 Online



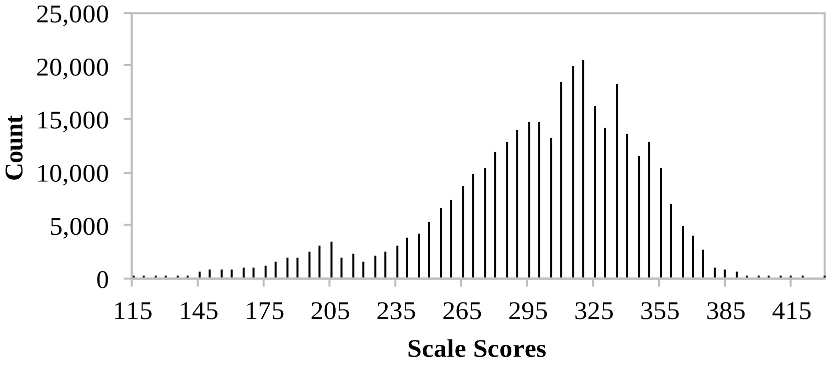
3.1.1.2 Grades 2–3

Table 3.1.1.2

Scale Score Descriptive Statistics: Oral 2-3 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	174,336	115	431	292.72	45.02
3	176,031	115	431	310.63	45.79
Total	350,367	115	431	301.72	46.29

Figure 3.1.1.2
Scale Scores: Oral 2-3 S403 Online



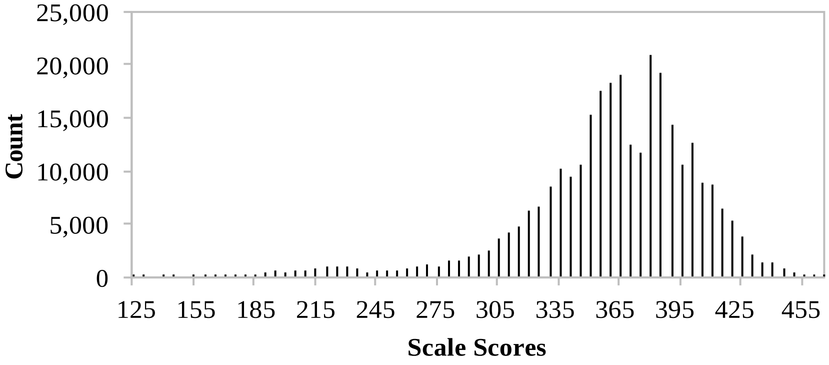
3.1.1.3 Grades 4–5

Table 3.1.1.3

Scale Score Descriptive Statistics: Oral 4-5 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	165,948	125	485	360.04	41.79
5	134,068	142	485	365.62	44.95
Total	300,016	125	485	362.53	43.32

Figure 3.1.1.3
Scale Scores: Oral 4-5 S403 Online



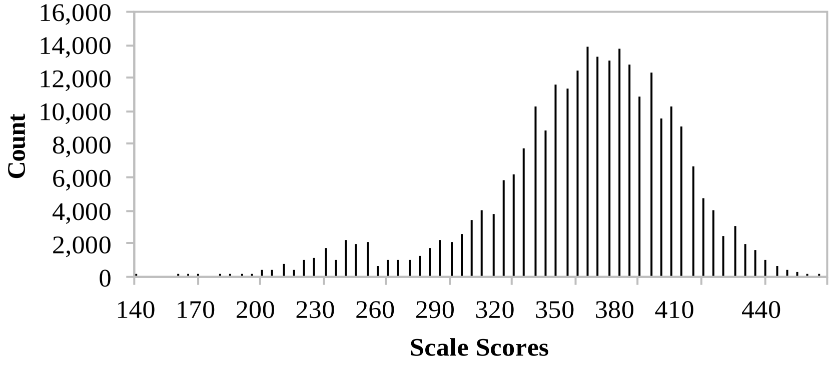
3.1.1.4 *Grades 6–8*

Table 3.1.1.4

Scale Score Descriptive Statistics: Oral 6-8 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	94,973	140	488	352.79	40.60
7	84,402	140	495	354.97	45.78
8	78,961	160	488	358.16	50.41
Total	258,336	140	495	355.14	45.52

Figure 3.1.1.4
Scale Scores: Oral 6-8 S403 Online



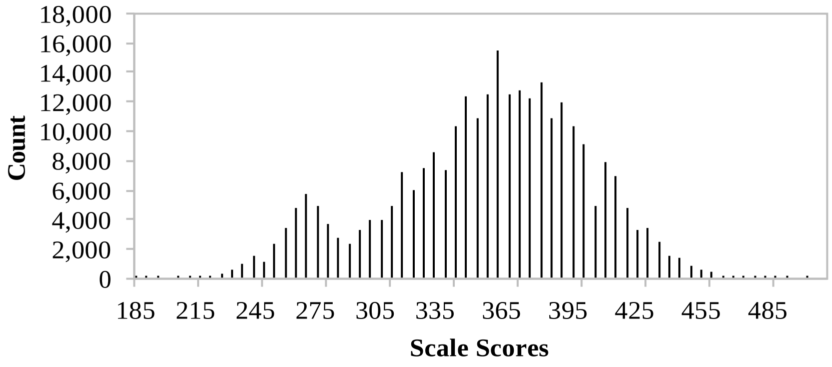
3.1.1.5 *Grades 9–12*

Table 3.1.1.5

Scale Score Descriptive Statistics: Oral 9-12 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	91,005	194	487	346.02	45.47
10	74,941	187	486	351.49	46.16
11	66,181	194	492	355.79	47.02
12	49,142	194	502	355.86	46.98
Total	281,269	187	502	351.49	46.47

Figure 3.1.1.5
Scale Scores: Oral 9-12 S403 Online



3.1.2 Literacy

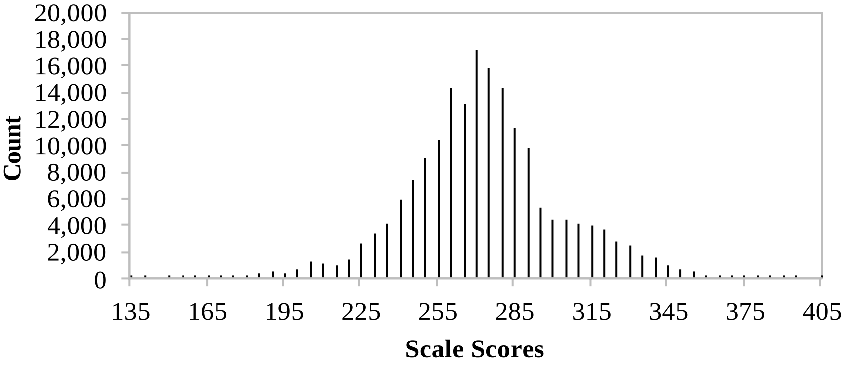
3.1.2.1 *Grade 1*

Table 3.1.2.1

Scale Score Descriptive Statistics: Litr 1 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	183,361	137	408	274.84	29.40
Total	183,361	137	408	274.84	29.40

Figure 3.1.2.1
Scale Scores: Litr 1 S403 Online



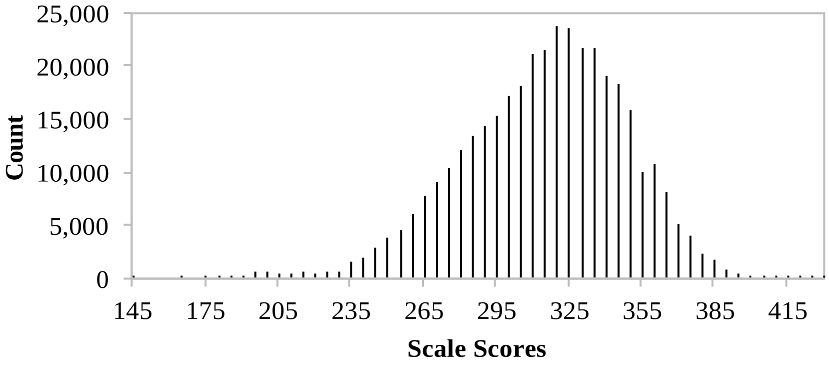
3.1.2.2 *Grades 2–3*

Table 3.1.2.2

Scale Score Descriptive Statistics: Litr 2-3 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	186,324	166	412	307.72	30.85
3	186,906	146	431	325.67	33.74
Total	373,230	146	431	316.71	33.55

Figure 3.1.2.2
Scale Scores: Litr 2-3 S403 Online



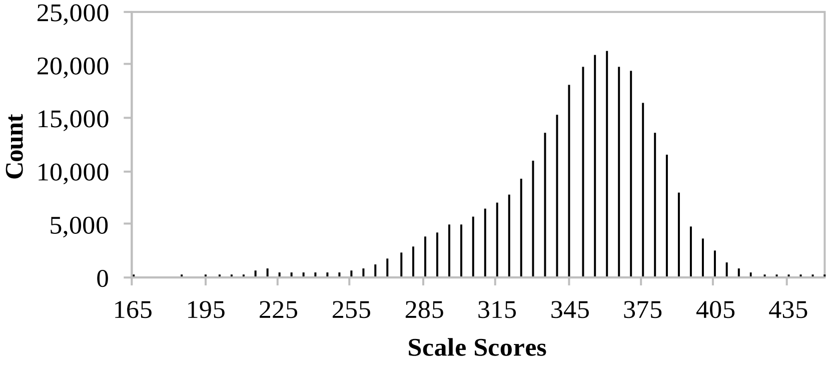
3.1.2.3 *Grades 4–5*

Table 3.1.2.3

Scale Score Descriptive Statistics: Litr 4-5 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	160,386	165	441	345.67	33.02
5	130,620	186	453	352.65	35.21
Total	291,006	165	453	348.80	34.19

Figure 3.1.2.3
Scale Scores: Litr 4-5 S403 Online



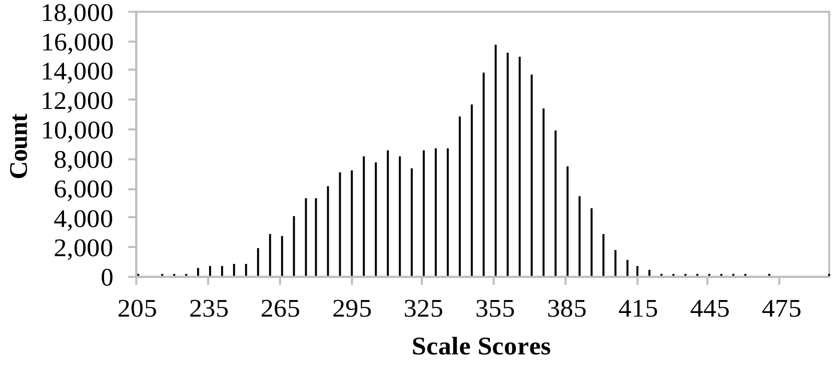
3.1.2.4 *Grades 6–8*

Table 3.1.2.4

Scale Score Descriptive Statistics: Litr 6-8 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	93,378	205	471	334.34	35.72
7	83,957	205	458	339.18	38.76
8	77,979	215	496	344.53	40.70
Total	255,314	205	496	339.04	38.52

Figure 3.1.2.4
Scale Scores: Litr 6-8 S403 Online



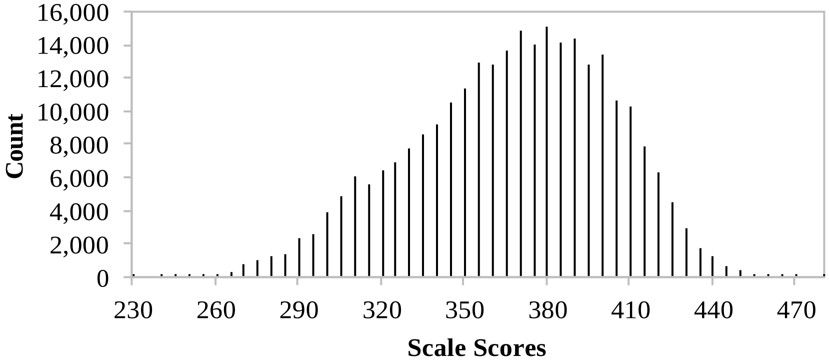
3.1.2.5 Grades 9–12

Table 3.1.2.5

Scale Score Descriptive Statistics: Litr 9-12 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	89,312	241	482	360.30	37.62
10	73,549	241	484	369.11	35.07
11	64,874	251	473	375.36	34.67
12	48,237	233	482	375.84	33.88
Total	275,972	233	484	368.90	36.21

Figure 3.1.2.5
Scale Scores: Litr 9-12 S403 Online



3.1.3 Comprehension

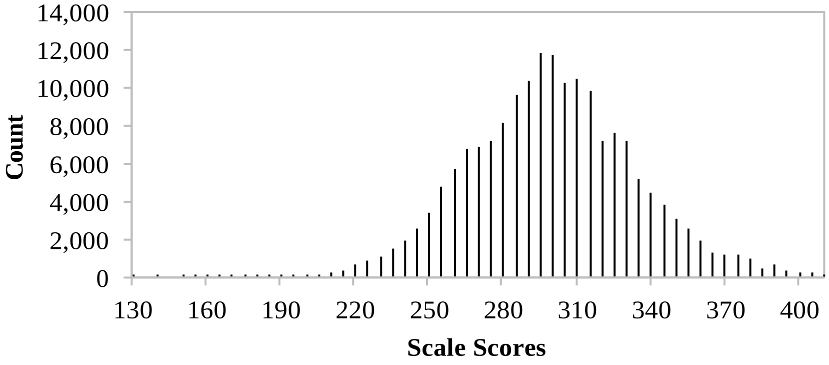
3.1.3.1 *Grade 1*

Table 3.1.3.1

Scale Score Descriptive Statistics: Cphn 1 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	175,698	131	413	301.87	33.37
Total	175,698	131	413	301.87	33.37

Figure 3.1.3.1
Scale Scores: Cphn 1 S403 Online

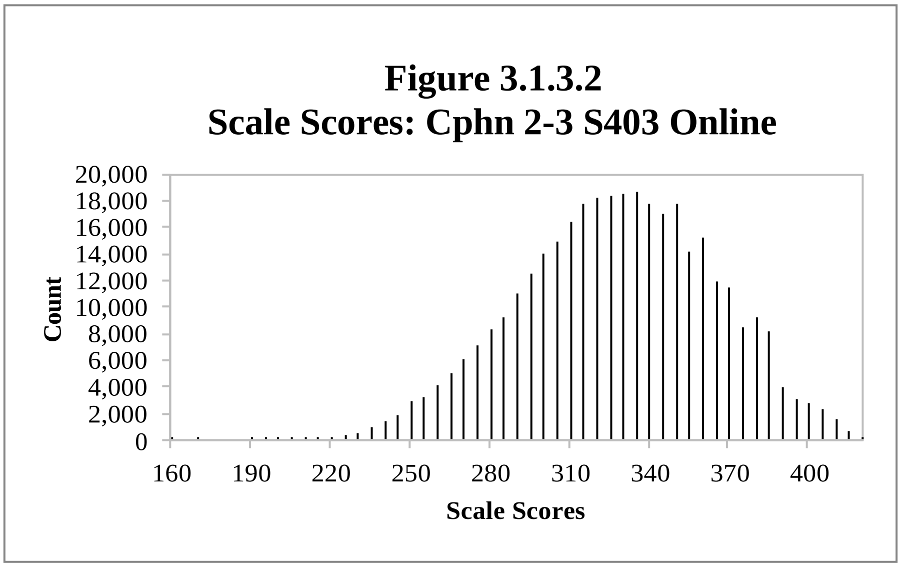


3.1.3.2 *Grades 2–3*

Table 3.1.3.2

Scale Score Descriptive Statistics: Cphn 2-3 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	178,466	172	423	320.68	33.52
3	179,511	163	423	338.35	36.92
Total	357,977	163	423	329.54	36.36



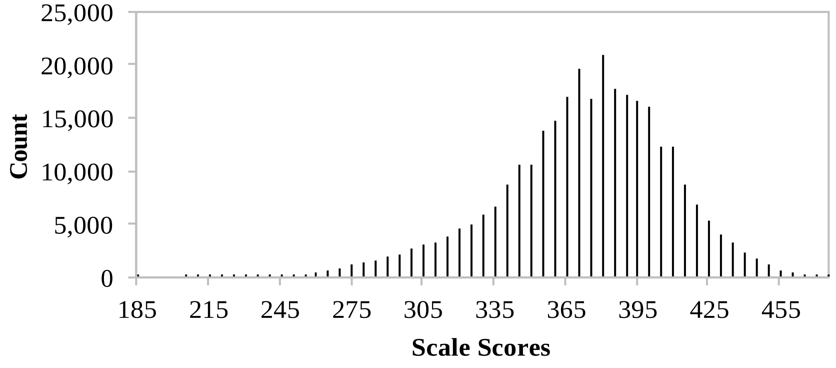
3.1.3.3 *Grades 4–5*

Table 3.1.3.3

Scale Score Descriptive Statistics: Cphn 4-5 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	168,857	188	475	372.11	33.80
5	136,488	207	475	378.25	36.97
Total	305,345	188	475	374.85	35.39

Figure 3.1.3.3
Scale Scores: Cphn 4-5 S403 Online



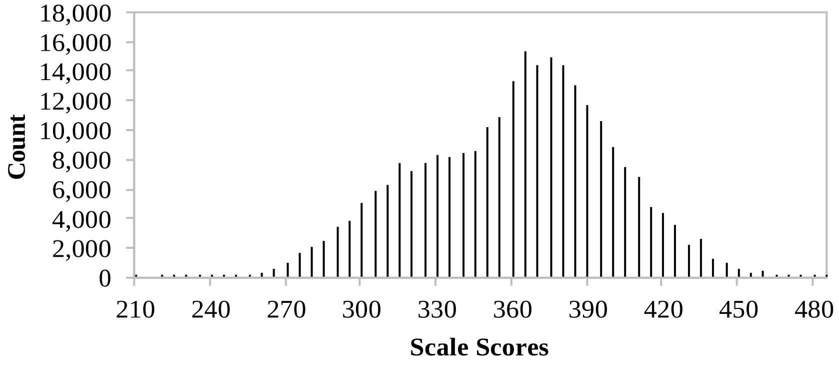
3.1.3.4 Grades 6–8

Table 3.1.3.4

Scale Score Descriptive Statistics: Cphn 6-8 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	96,582	224	487	358.12	34.79
7	86,076	212	487	363.59	38.84
8	80,182	214	487	369.70	41.89
Total	262,840	212	487	363.44	38.68

Figure 3.1.3.4
Scale Scores: Cphn 6-8 S403 Online



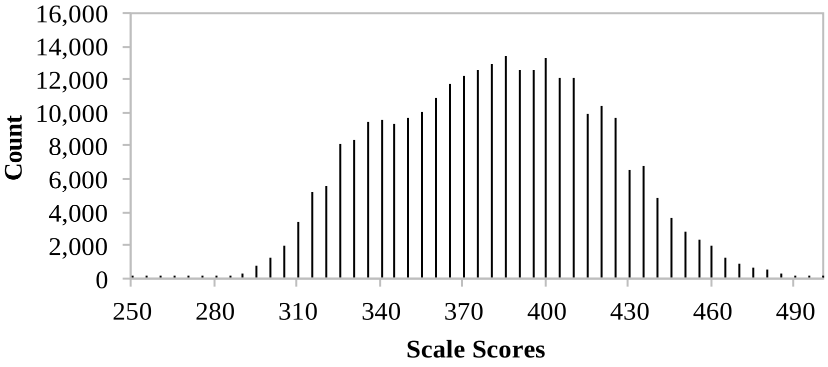
3.1.3.5 *Grades 9–12*

Table 3.1.3.5

Scale Score Descriptive Statistics: Cphn 9-12 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	91,152	258	500	375.56	37.98
10	75,172	254	500	382.56	38.41
11	66,590	265	500	388.55	38.80
12	49,226	268	500	388.97	37.80
Total	282,140	254	500	382.83	38.67

Figure 3.1.3.5
Scale Scores: Cphn 9-12 S403 Online



3.1.4 Overall

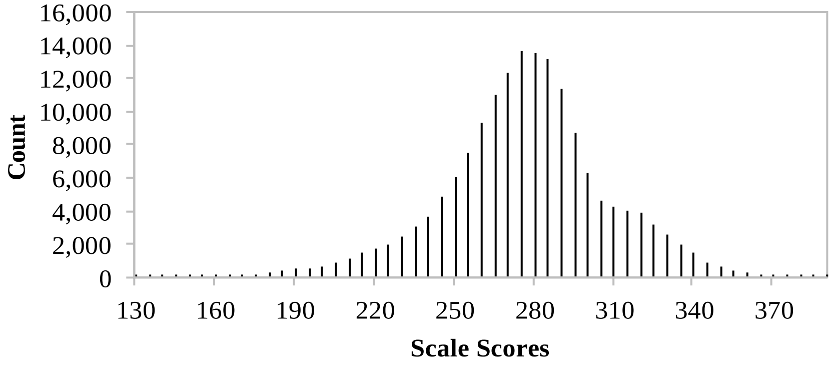
3.1.4.1 *Grade 1*

Table 3.1.4.1

Scale Score Descriptive Statistics: Over 1 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
1	165,022	131	391	278.65	30.30
Total	165,022	131	391	278.65	30.30

Figure 3.1.4.1
Scale Scores: Over 1 S403 Online



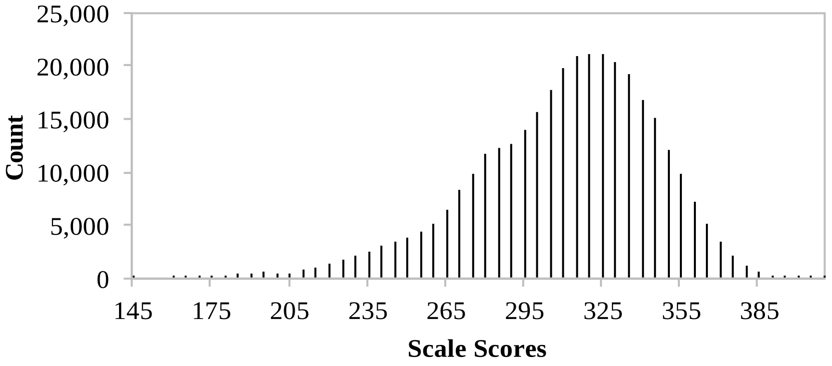
3.1.4.2 Grades 2–3

Table 3.1.4.2

Scale Score Descriptive Statistics: Over 2-3 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
2	167,720	162	402	302.95	32.41
3	169,474	146	411	320.90	34.83
Total	337,194	146	411	311.97	34.83

Figure 3.1.4.2
Scale Scores: Over 2-3 S403 Online



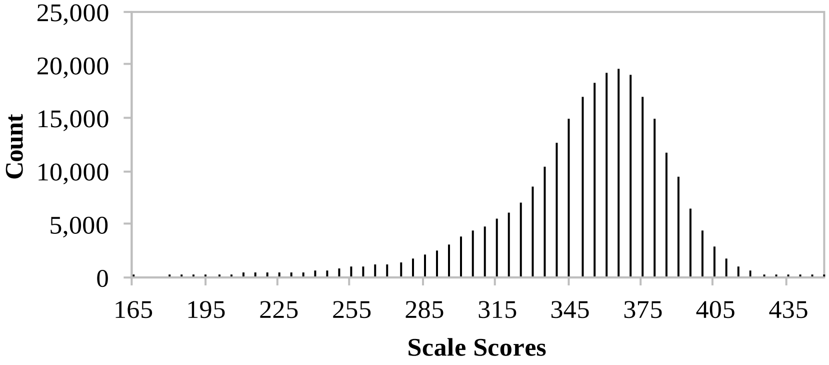
3.1.4.3 *Grades 4–5*

Table 3.1.4.3

Scale Score Descriptive Statistics: Over 4-5 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
4	144,258	168	445	349.74	33.47
5	118,083	188	450	356.26	36.04
Total	262,341	168	450	352.67	34.80

Figure 3.1.4.3
Scale Scores: Over 4-5 S403 Online



3.1.4.4 *Grades 6–8*

Table 3.1.4.4

Scale Score Descriptive Statistics: Over 6-8 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
6	83,413	208	476	339.68	34.62
7	74,786	207	456	343.68	38.58
8	70,081	209	470	348.29	41.54
Total	228,280	207	476	343.63	38.31

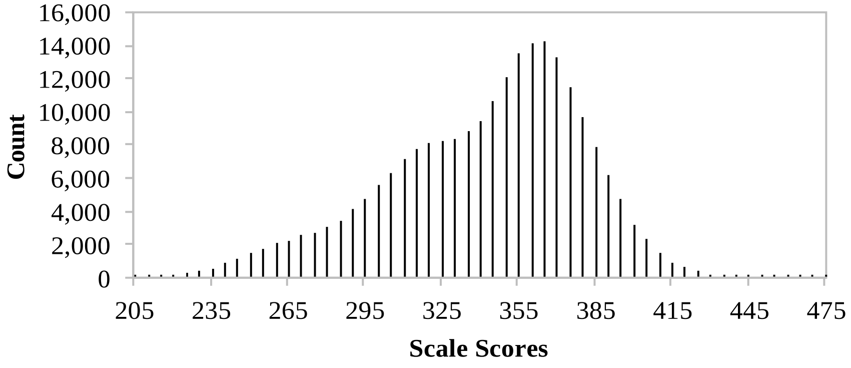


Figure 3.1.4.4
Scale Scores: Over 6-8 S403 Online

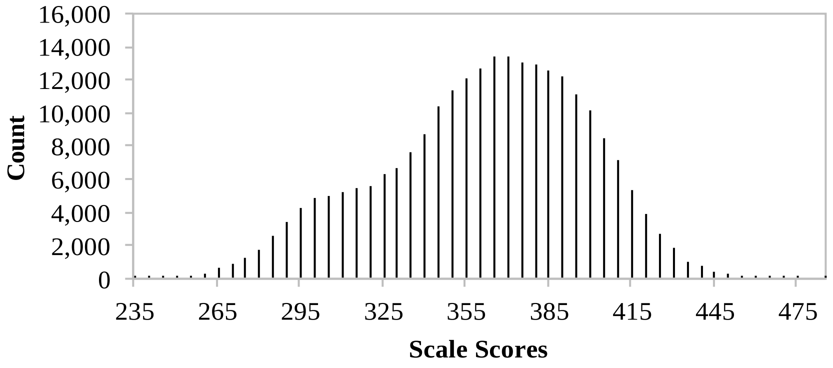
3.1.4.5 *Grades 9–12*

Table 3.1.4.5

Scale Score Descriptive Statistics: Over 9-12 S403 Online

Grade	No. of Students	Min.	Max.	Mean	Std. Dev.
9	80,043	236	477	355.79	37.84
10	66,038	236	473	363.60	36.14
11	57,988	248	473	369.27	36.10
12	43,523	241	488	369.58	35.12
Total	247,592	236	488	363.45	36.97

Figure 3.1.4.5
Scale Scores: Over 9-12 S403 Online



3.2 Proficiency Level Distribution for Composites

Figures and tables in this section provide information on the proficiency level distribution for each of the composites for each grade-level cluster.

In each figure, the horizontal axis shows the six WIDA proficiency levels. The vertical axis shows the percentage of students. Each bar shows the percentage of students who were placed into each proficiency level in the domain being tested on this test form.

The tables in this section present, by grade and by total for the grade-level cluster:

- The WIDA proficiency level designation (1–6)
- The number of students (count) whose performance on the test form placed them into that proficiency level in the domain being tested
- The percentage of students, out of the total number of students taking the form, who were placed into that proficiency level in the domain being tested

3.2.1 Oral

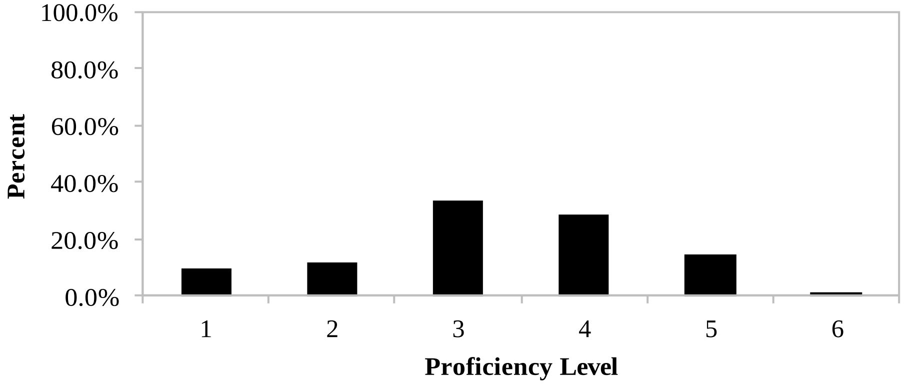
3.2.1.1 Grade 1

Table 3.2.1.1

Proficiency Level Distribution: Oral 1 S403 Online

Level	Grade 1		Total	
	Count	**Percent**	**Count**	**Percent**
1	16,268	9.58%	16,268	9.58%
2	20,630	12.14%	20,630	12.14%
3	56,868	33.47%	56,868	33.47%
4	48,739	28.69%	48,739	28.69%
5	24,819	14.61%	24,819	14.61%
6	2,567	1.51%	2,567	1.51%
Total	169,891	100.00%	169,891	100.00%

Figure 3.2.1.1
Proficiency Level: Oral 1 S403 Online



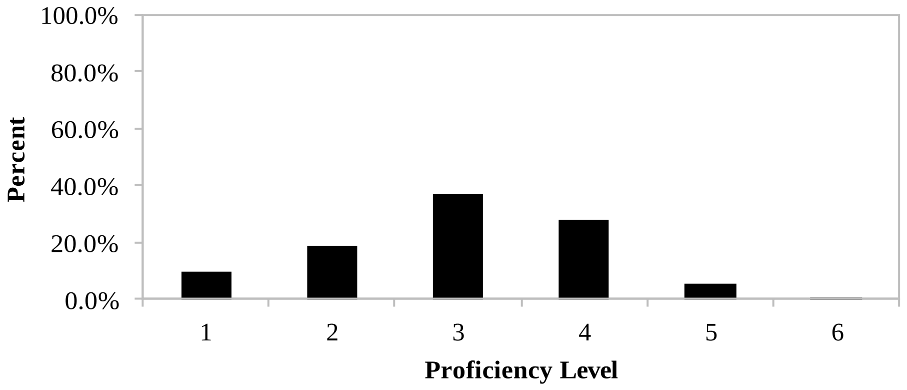
3.2.1.2 Grades 2–3

Table 3.2.1.2

Proficiency Level Distribution: Oral 2-3 S403 Online

	Grade 2		Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	17,801	10.21%	16,131	9.16%	33,932	9.68%
2	37,002	21.22%	29,877	16.97%	66,879	19.09%
3	65,597	37.63%	65,455	37.18%	131,052	37.40%
4	43,223	24.79%	55,613	31.59%	98,836	28.21%
5	10,297	5.91%	8,792	4.99%	19,089	5.45%
6	416	0.24%	163	0.09%	579	0.17%
Total	174,336	100.00%	176,031	100.00%	350,367	100.00%

Figure 3.2.1.2
Proficiency Level: Oral 2-3 S403 Online



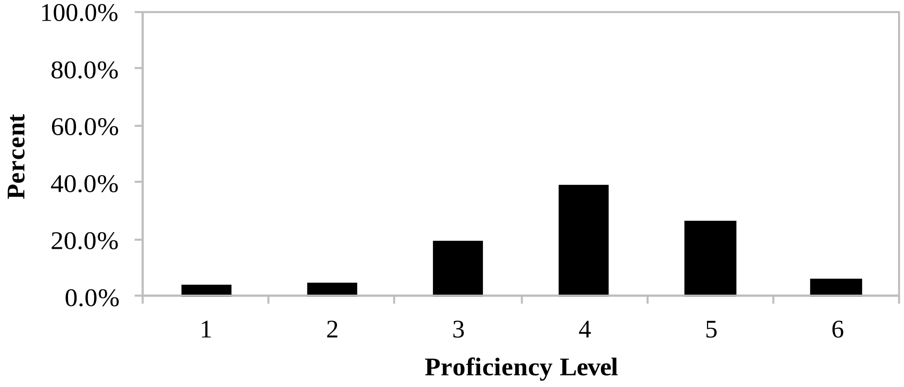
3.2.1.3 Grades 4–5

Table 3.2.1.3

Proficiency Level Distribution: Oral 4-5 S403 Online

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	5,431	3.27%	6,022	4.49%	11,453	3.82%
2	7,476	4.51%	7,033	5.25%	14,509	4.84%
3	31,198	18.80%	27,148	20.25%	58,346	19.45%
4	63,240	38.11%	53,649	40.02%	116,889	38.96%
5	47,417	28.57%	32,897	24.54%	80,314	26.77%
6	11,186	6.74%	7,319	5.46%	18,505	6.17%
Total	165,948	100.00%	134,068	100.00%	300,016	100.00%

Figure 3.2.1.3
Proficiency Level: Oral 4-5 S403 Online



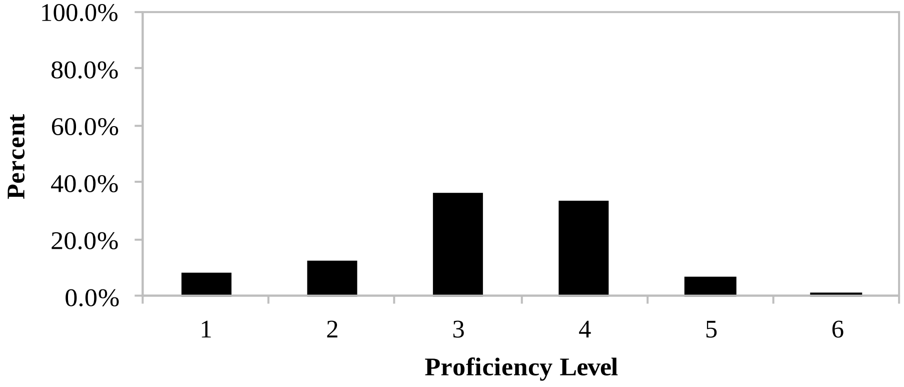
3.2.1.4 Grades 6–8

Table 3.2.1.4

Proficiency Level Distribution: Oral 6-8 S403 Online

Level	Grade 6		Grade 7		Grade 8		Total	
	Count	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	5,443	5.73%	7,387	8.75%	8,781	11.12%	21,611	8.37%
2	10,010	10.54%	11,217	13.29%	11,635	14.74%	32,862	12.72%
3	36,443	38.37%	31,350	37.14%	26,959	34.14%	94,752	36.68%
4	35,147	37.01%	27,528	32.62%	25,001	31.66%	87,676	33.94%
5	7,150	7.53%	5,924	7.02%	5,652	7.16%	18,726	7.25%
6	780	0.82%	996	1.18%	933	1.18%	2,709	1.05%
Total	94,973	100.00%	84,402	100.00%	78,961	100.00%	258,336	100.00%

Figure 3.2.1.4
Proficiency Level: Oral 6-8 S403 Online



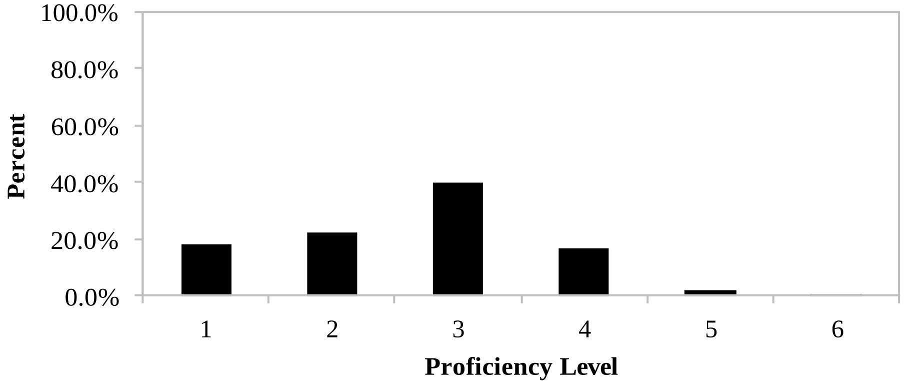
3.2.1.5 *Grades 9–12*

Table 3.2.1.5

Proficiency Level Distribution: Oral 9-12 S403 Online

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	15,570	17.11%	13,268	17.70%	12,537	18.94%	10,326	21.01%	51,701	18.38%
2	21,438	23.56%	16,281	21.73%	14,164	21.40%	10,207	20.77%	62,090	22.07%
3	35,073	38.54%	30,334	40.48%	26,569	40.15%	20,906	42.54%	112,882	40.13%
4	16,559	18.20%	13,051	17.42%	11,222	16.96%	6,911	14.06%	47,743	16.97%
5	2,123	2.33%	1,798	2.40%	1,535	2.32%	721	1.47%	6,177	2.20%
6	242	0.27%	209	0.28%	154	0.23%	71	0.14%	676	0.24%
Total	91,005	100.00%	74,941	100.00%	66,181	100.00%	49,142	100.00%	281,269	100.00%

Figure 3.2.1.5
Proficiency Level: Oral 9-12 S403 Online



3.2.2 Literacy

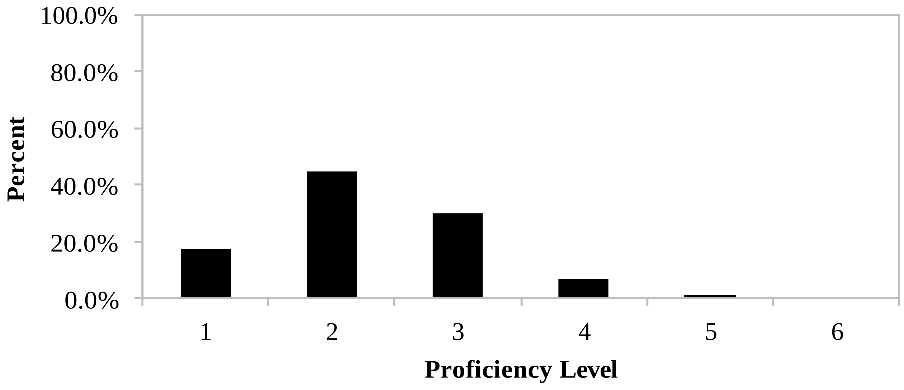
3.2.2.1 Grade 1

Table 3.2.2.1

Proficiency Level Distribution: Litr 1 S403 Online

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	32,238	17.58%	32,238	17.58%
2	82,333	44.90%	82,333	44.90%
3	54,653	29.81%	54,653	29.81%
4	12,156	6.63%	12,156	6.63%
5	1,815	0.99%	1,815	0.99%
6	166	0.09%	166	0.09%
Total	183,361	100.00%	183,361	100.00%

Figure 3.2.2.1
Proficiency Level: Litr 1 S403 Online



3.2.2.2 *Grades 2–3*

Table 3.2.2.2

Proficiency Level Distribution: Litr 2-3 S403 Online

	Grade 2		**Grade 3**		**Total**	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	14,033	7.53%	13,474	7.21%	27,507	7.37%
2	40,724	21.86%	29,061	15.55%	69,785	18.70%
3	93,705	50.29%	83,174	44.50%	176,879	47.39%
4	33,312	17.88%	51,125	27.35%	84,437	22.62%
5	4,130	2.22%	8,972	4.80%	13,102	3.51%
6	420	0.23%	1,100	0.59%	1,520	0.41%
Total	186,324	100.00%	186,906	100.00%	373,230	100.00%

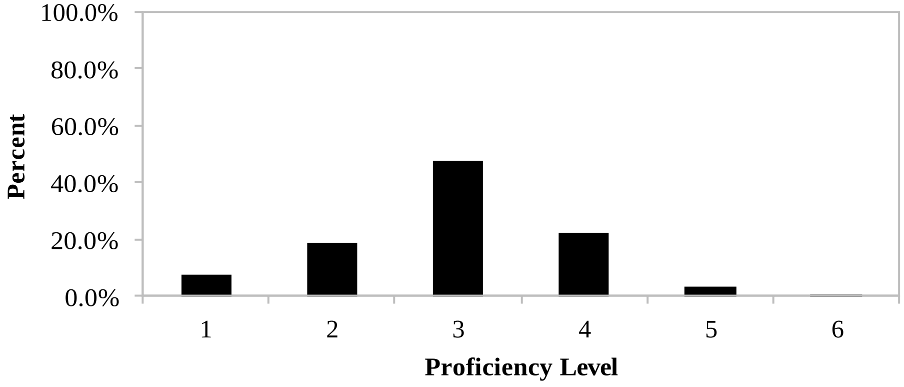


Figure 3.2.2.2
Proficiency Level: Litr 2-3 S403 Online

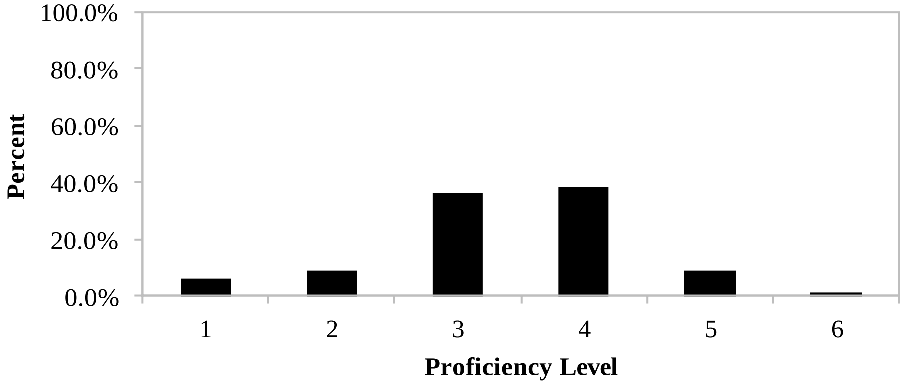
3.2.2.3 Grades 4–5

Table 3.2.2.3

Proficiency Level Distribution: Litr 4-5 S403 Online

	Grade 4		**Grade 5**		**Total**	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	8,774	5.47%	8,556	6.55%	17,330	5.96%
2	13,902	8.67%	12,237	9.37%	26,139	8.98%
3	60,909	37.98%	45,285	34.67%	106,194	36.49%
4	61,580	38.39%	50,507	38.67%	112,087	38.52%
5	13,384	8.34%	12,362	9.46%	25,746	8.85%
6	1,837	1.15%	1,673	1.28%	3,510	1.21%
Total	160,386	100.00%	130,620	100.00%	291,006	100.00%

Figure 3.2.2.3
Proficiency Level: Litr 4-5 S403 Online



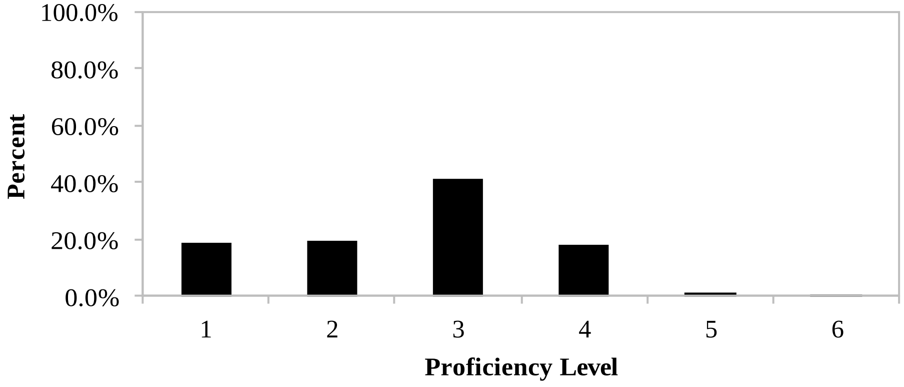
3.2.2.4 Grades 6–8

Table 3.2.2.4

Proficiency Level Distribution: Litr 6-8 S403 Online

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	16,261	17.41%	15,858	18.89%	16,202	20.78%	48,321	18.93%
2	17,134	18.35%	17,725	21.11%	15,647	20.07%	50,506	19.78%
3	42,989	46.04%	33,661	40.09%	28,573	36.64%	105,223	41.21%
4	16,026	17.16%	15,275	18.19%	15,873	20.36%	47,174	18.48%
5	901	0.96%	1,338	1.59%	1,596	2.05%	3,835	1.50%
6	67	0.07%	100	0.12%	88	0.11%	255	0.10%
Total	93,378	100.00%	83,957	100.00%	77,979	100.00%	255,314	100.00%

Figure 3.2.2.4
Proficiency Level: Litr 6-8 S403 Online



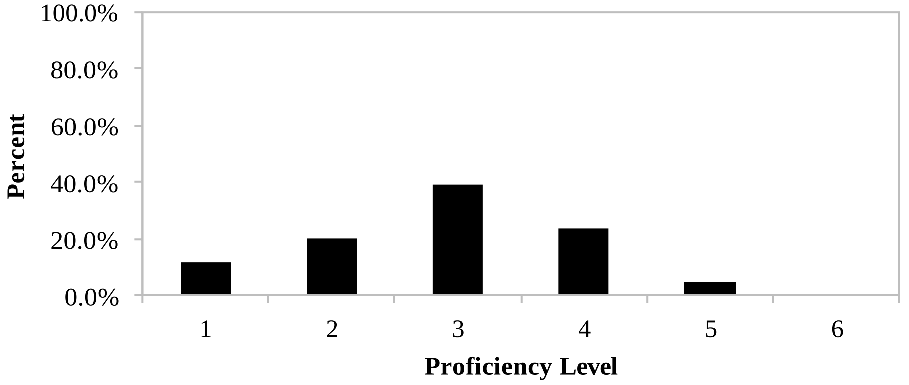
3.2.2.5 *Grades 9–12*

Table 3.2.2.5

Proficiency Level Distribution: Litr 9-12 S403 Online

	Grade 9		Grade 10		Grade 11		Grade 12		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	13,130	14.70%	7,728	10.51%	6,654	10.26%	6,021	12.48%	33,533	12.15%
2	16,653	18.65%	14,445	19.64%	13,284	20.48%	11,211	23.24%	55,593	20.14%
3	33,649	37.68%	29,139	39.62%	25,120	38.72%	19,544	40.52%	107,452	38.94%
4	21,104	23.63%	18,328	24.92%	16,266	25.07%	9,918	20.56%	65,616	23.78%
5	4,517	5.06%	3,749	5.10%	3,479	5.36%	1,530	3.17%	13,275	4.81%
6	259	0.29%	160	0.22%	71	0.11%	13	0.03%	503	0.18%
Total	89,312	100.00%	73,549	100.00%	64,874	100.00%	48,237	100.00%	275,972	100.00%

Figure 3.2.2.5
Proficiency Level: Litr 9-12 S403 Online



3.2.3 Comprehension

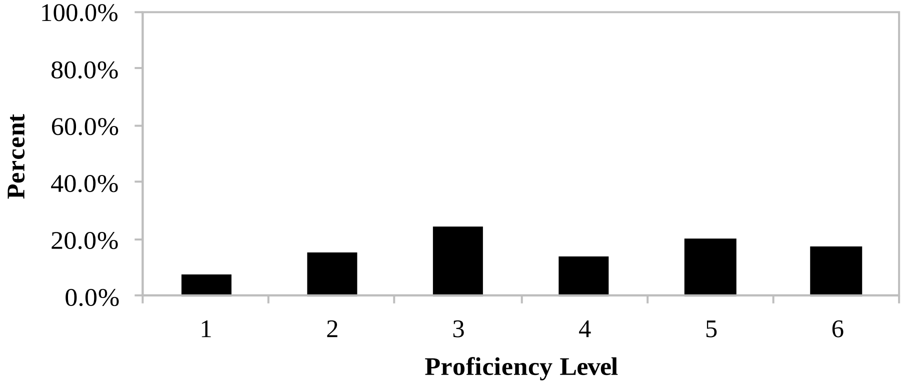
3.2.3.1 Grade 1

Table 3.2.3.1

Proficiency Level Distribution: Cphn 1 S403 Online

	Grade 1		Total	
Level	**Count**	**Percent**	**Count**	**Percent**
1	13,495	7.68%	13,495	7.68%
2	27,464	15.63%	27,464	15.63%
3	42,973	24.46%	42,973	24.46%
4	24,810	14.12%	24,810	14.12%
5	35,904	20.44%	35,904	20.44%
6	31,052	17.67%	31,052	17.67%
Total	175,698	100.00%	175,698	100.00%

Figure 3.2.3.1
Proficiency Level: Cphn 1 S403 Online



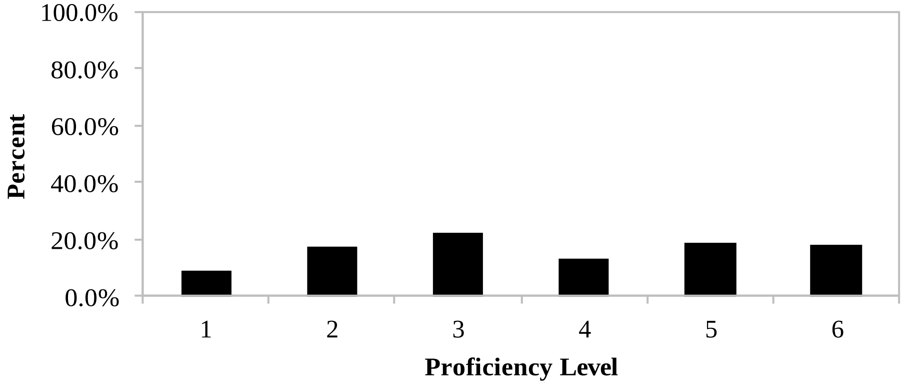
3.2.3.2 Grades 2–3

Table 3.2.3.2

Proficiency Level Distribution: Cphn 2-3 S403 Online

	Grade 2		Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	13,969	7.83%	17,148	9.55%	31,117	8.69%
2	33,380	18.70%	29,380	16.37%	62,760	17.53%
3	43,050	24.12%	38,238	21.30%	81,288	22.71%
4	25,596	14.34%	22,863	12.74%	48,459	13.54%
5	33,885	18.99%	34,283	19.10%	68,168	19.04%
6	28,586	16.02%	37,599	20.95%	66,185	18.49%
Total	178,466	100.00%	179,511	100.00%	357,977	100.00%

Figure 3.2.3.2
Proficiency Level: Cphn 2-3 S403 Online



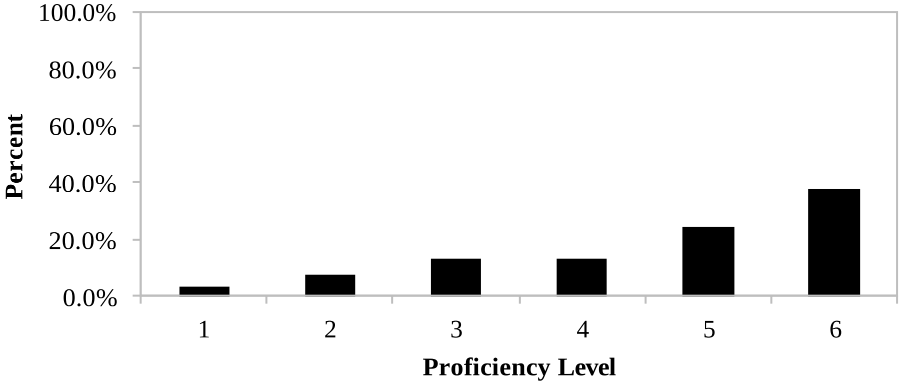
3.2.3.3 Grades 4–5

Table 3.2.3.3

Proficiency Level Distribution: Cphn 4-5 S403 Online

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	4,667	2.76%	6,795	4.98%	11,462	3.75%
2	12,384	7.33%	11,573	8.48%	23,957	7.85%
3	21,932	12.99%	18,538	13.58%	40,470	13.25%
4	21,539	12.76%	17,965	13.16%	39,504	12.94%
5	42,432	25.13%	32,201	23.59%	74,633	24.44%
6	65,903	39.03%	49,416	36.21%	115,319	37.77%
Total	168,857	100.00%	136,488	100.00%	305,345	100.00%

Figure 3.2.3.3
Proficiency Level: Cphn 4-5 S403 Online



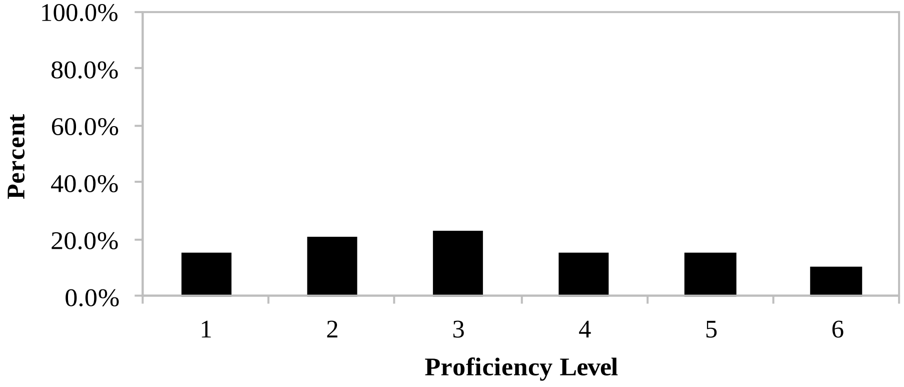
3.2.3.4 *Grades 6–8*

Table 3.2.3.4

Proficiency Level Distribution: Cphn 6-8 S403 Online

	Grade 6		**Grade 7**		**Grade 8**		**Total**	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	11,983	12.41%	13,690	15.90%	14,137	17.63%	39,810	15.15%
2	20,786	21.52%	18,150	21.09%	15,737	19.63%	54,673	20.80%
3	23,988	24.84%	19,709	22.90%	16,719	20.85%	60,416	22.99%
4	15,617	16.17%	12,848	14.93%	11,119	13.87%	39,584	15.06%
5	16,304	16.88%	12,035	13.98%	12,002	14.97%	40,341	15.35%
6	7,904	8.18%	9,644	11.20%	10,468	13.06%	28,016	10.66%
Total	96,582	100.00%	86,076	100.00%	80,182	100.00%	262,840	100.00%

Figure 3.2.3.4
Proficiency Level: Cphn 6-8 S403 Online



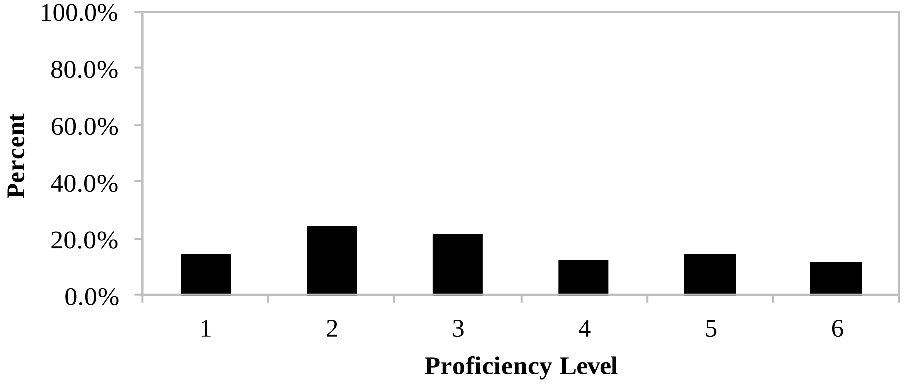
3.2.3.5 *Grades 9–12*

Table 3.2.3.5

Proficiency Level Distribution: Cphn 9-12 S403 Online

	Grade 9		**Grade 10**		**Grade 11**		**Grade 12**		**Total**	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	13,056	14.32%	10,940	14.55%	9,819	14.75%	8,040	16.33%	41,855	14.83%
2	24,046	26.38%	17,835	23.73%	15,689	23.56%	12,270	24.93%	69,840	24.75%
3	19,428	21.31%	16,506	21.96%	13,609	20.44%	10,649	21.63%	60,192	21.33%
4	10,980	12.05%	9,408	12.52%	7,962	11.96%	6,485	13.17%	34,835	12.35%
5	12,822	14.07%	10,921	14.53%	10,681	16.04%	6,933	14.08%	41,357	14.66%
6	10,820	11.87%	9,562	12.72%	8,830	13.26%	4,849	9.85%	34,061	12.07%
Total	91,152	100.00%	75,172	100.00%	66,590	100.00%	49,226	100.00%	282,140	100.00%

Figure 3.2.3.5
Proficiency Level: Cphn 9-12 S403 Online



3.2.4 Overall

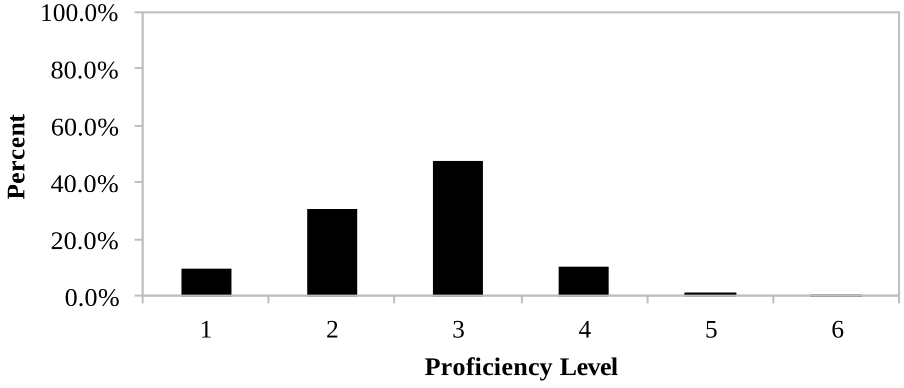
3.2.4.1 Grade 1

Table 3.2.4.1

Proficiency Level Distribution: Over 1 S403 Online

Level	Grade 1		Total	
	Count	**Percent**	**Count**	**Percent**
1	16,588	10.05%	16,588	10.05%
2	50,813	30.79%	50,813	30.79%
3	78,148	47.36%	78,148	47.36%
4	16,688	10.11%	16,688	10.11%
5	2,640	1.60%	2,640	1.60%
6	145	0.09%	145	0.09%
Total	165,022	100.00%	165,022	100.00%

Figure 3.2.4.1
Proficiency Level: Over 1 S403 Online



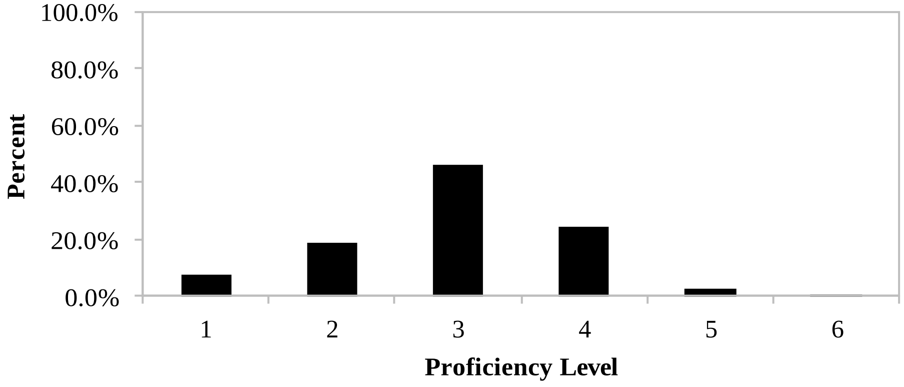
3.2.4.2 *Grades 2–3*

Table 3.2.4.2

Proficiency Level Distribution: Over 2-3 S403 Online

	Grade 2		Grade 3		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	12,831	7.65%	12,735	7.51%	25,566	7.58%
2	37,285	22.23%	25,795	15.22%	63,080	18.71%
3	80,984	48.29%	75,922	44.80%	156,906	46.53%
4	33,045	19.70%	48,942	28.88%	81,987	24.31%
5	3,443	2.05%	5,954	3.51%	9,397	2.79%
6	132	0.08%	126	0.07%	258	0.08%
Total	167,720	100.00%	169,474	100.00%	337,194	100.00%

Figure 3.2.4.2
Proficiency Level: Over 2-3 S403 Online



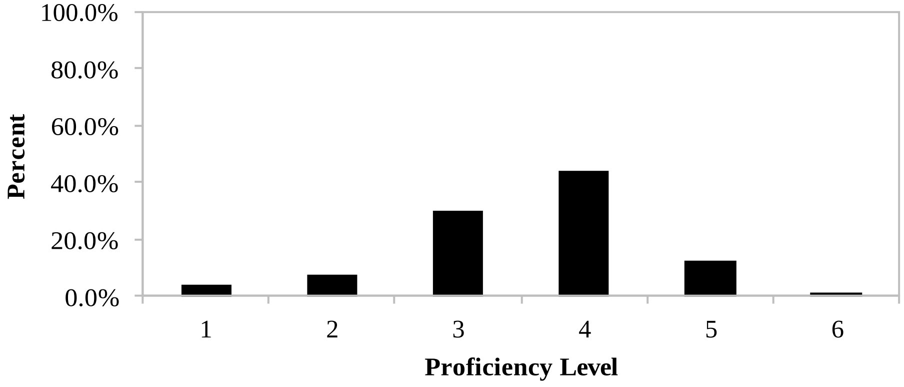
3.2.4.3 Grades 4–5

Table 3.2.4.3

Proficiency Level Distribution: Over 4-5 S403 Online

	Grade 4		Grade 5		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	5,651	3.92%	6,049	5.12%	11,700	4.46%
2	9,959	6.90%	9,363	7.93%	19,322	7.37%
3	44,590	30.91%	34,192	28.96%	78,782	30.03%
4	64,411	44.65%	52,006	44.04%	116,417	44.38%
5	17,767	12.32%	15,087	12.78%	32,854	12.52%
6	1,880	1.30%	1,386	1.17%	3,266	1.24%
Total	144,258	100.00%	118,083	100.00%	262,341	100.00%

Figure 3.2.4.3
Proficiency Level: Over 4-5 S403 Online



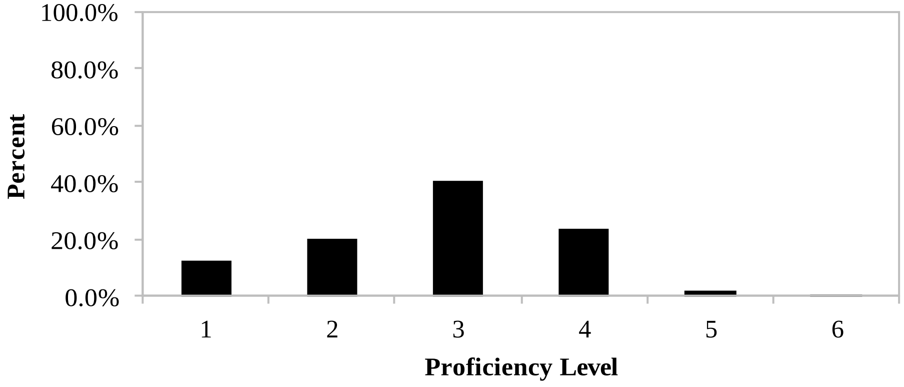
3.2.4.4 Grades 6–8

Table 3.2.4.4

Proficiency Level Distribution: Over 6-8 S403 Online

	Grade 6		Grade 7		Grade 8		Total	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	8,081	9.69%	9,910	13.25%	10,926	15.59%	28,917	12.67%
2	16,653	19.96%	15,734	21.04%	14,156	20.20%	46,543	20.39%
3	37,614	45.09%	29,784	39.83%	25,740	36.73%	93,138	40.80%
4	19,892	23.85%	17,643	23.59%	17,384	24.81%	54,919	24.06%
5	1,103	1.32%	1,621	2.17%	1,782	2.54%	4,506	1.97%
6	70	0.08%	94	0.13%	93	0.13%	257	0.11%
Total	83,413	100.00%	74,786	100.00%	70,081	100.00%	228,280	100.00%

Figure 3.2.4.4
Proficiency Level: Over 6-8 S403 Online



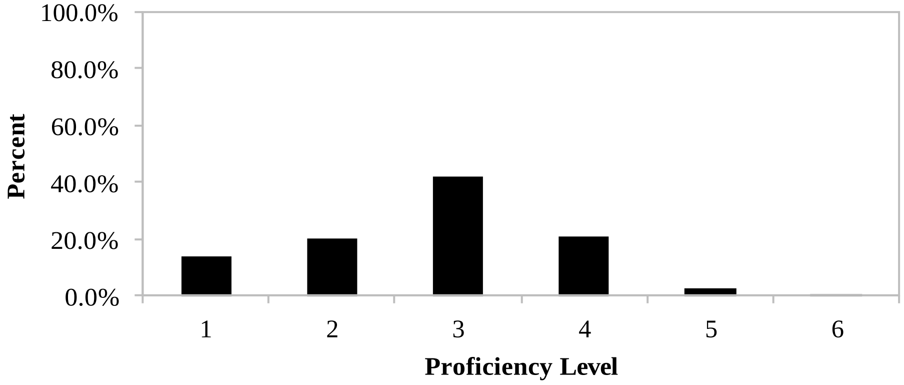
3.2.4.5 *Grades 9–12*

Table 3.2.4.5

Proficiency Level Distribution: Over 9-12 S403 Online

	Grade 9		**Grade 10**		**Grade 11**		**Grade 12**		**Total**	
Level	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
1	12,074	15.08%	8,554	12.95%	7,584	13.08%	6,495	14.92%	34,707	14.02%
2	15,329	19.15%	12,867	19.48%	11,627	20.05%	10,037	23.06%	49,860	20.14%
3	33,203	41.48%	28,125	42.59%	24,379	42.04%	18,764	43.11%	104,471	42.19%
4	16,888	21.10%	14,626	22.15%	12,686	21.88%	7,498	17.23%	51,698	20.88%
5	2,452	3.06%	1,812	2.74%	1,670	2.88%	722	1.66%	6,656	2.69%
6	97	0.12%	54	0.08%	42	0.07%	7	0.02%	200	0.08%
Total	80,043	100.00%	66,038	100.00%	57,988	100.00%	43,523	100.00%	247,592	100.00%

Figure 3.2.4.5
Proficiency Level: Over 9-12 S403 Online



3.3 Reliability

To estimate the score reliability of the composite scores, a stratified Cronbach's alpha coefficient (e.g., Rudner, 2001; Kamata, Turhan, & Darandari, 2003; Kane & Case, 2004) 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 = weight of component j
o_j^2 = variance of component j
σ_c^2 = variance of composite
ρ_j = reliability coefficient of component j

Tables are provided below to express the stratified Cronbach's alpha for each of the composites. The first table for each composite provides stratified Cronbach's alpha for all test takers. The second table for each composite provides the same information for the population of female students and the population of male students. The third table provides information by ethnicity, for Hispanic and non-Hispanic test takers, and the fourth table provides information for the population of students who have an individualized education plan (IEP).

Each table is divided by grade-level cluster. Tables first include the input values used to compute Cronbach's alpha. The table lists the number of components for each composite and their weight. (See also the beginning of this chapter for an overview of how the composites are computed.)

For the Listening and Reading domain components, the reliability coefficient is the Rasch separation reliability coefficient, provided in Section 2.10.

For Writing and Speaking domain components, which have multiple test forms for each grade-level cluster, a single reliability coefficient for the grade-level cluster is derived. To produce this single value, values for Cronbach's alpha for each of the tiers in the grade-level cluster (provided in Section 2.10) are weighted by the number of students who were administered the tier form, and a weighted average is expressed in the tables.

For each relevant domain component, the variance of the scale score is also provided. The variance of the composite scale score is also provided. The variances of domains and composites are computed for students who had valid results in all four domains.

Finally, the computed stratified Cronbach's alpha value for the composite is provided, by grade-level cluster.

The stratified Cronbach's alpha, presented in the tables in this section, was also used to produce the *Accuracy and Consistency* classification tables of the composites (Section 3.4).

3.3.1 Oral

Table 3.3.1.1

Reliability of Composite: Oral S403 Online

Cluster	Component	Weight	Variance	Reliability
1	Listening	0.50	2974.123	0.850
	Speaking	0.50	2712.492	0.832
	Oral		2109.764	0.893
2-3	Listening	0.50	3295.516	0.840
	Speaking	0.50	2285.756	0.791
	Oral		2152.724	0.883
4-5	Listening	0.50	2624.456	0.810
	Speaking	0.50	2311.342	0.816
	Oral		1929.018	0.880
6-8	Listening	0.50	2367.896	0.850
	Speaking	0.50	2942.663	0.820
	Oral		2095.367	0.895
9-12	Listening	0.50	2630.452	0.860
	Speaking	0.50	2980.093	0.844
	Oral		2167.276	0.904

Table 3.3.1.2

Reliability of Composite: Oral S403 Online by Gender

			Female		Male	
Cluster	Component	Weight	Variance	Reliability	Variance	Reliability
1	Listening	0.50	2827.260	0.840	3057.022	0.860
	Speaking	0.50	2664.132	0.834	2672.046	0.828
	Oral		2010.926	0.889	2134.236	0.896
2-3	Listening	0.50	3113.919	0.830	3427.582	0.840
	Speaking	0.50	2244.973	0.789	2245.847	0.788
	Oral		2052.442	0.878	2195.324	0.883
4-5	Listening	0.50	2452.009	0.800	2724.672	0.810
	Speaking	0.50	2247.115	0.814	2325.015	0.818
	Oral		1814.950	0.875	1984.829	0.881
6-8	Listening	0.50	2254.865	0.850	2427.926	0.860
	Speaking	0.50	2943.614	0.823	2895.799	0.818
	Oral		2048.234	0.895	2095.872	0.897
9-12	Listening	0.50	2530.143	0.860	2690.832	0.860
	Speaking	0.50	2901.633	0.837	3003.107	0.849
	Oral		2088.307	0.901	2201.447	0.906

Table 3.3.1.3

Reliability of Composite: Oral S403 Online by Ethnicity

			Hispanic		Other	
Cluster	Component	Weight	Variance	Reliability	Variance	Reliability
1	Listening	0.50	2882.609	0.850	3006.747	0.850
	Speaking	0.50	2706.897	0.834	2580.631	0.823
	Oral		2066.285	0.893	2052.961	0.889
2-3	Listening	0.50	3217.327	0.840	3267.025	0.830
	Speaking	0.50	2296.300	0.796	2134.472	0.774
	Oral		2118.970	0.884	2063.091	0.874
4-5	Listening	0.50	2479.652	0.800	2548.110	0.800
	Speaking	0.50	2224.972	0.816	2160.747	0.808
	Oral		1813.545	0.875	1818.337	0.873
6-8	Listening	0.50	2228.071	0.850	2378.218	0.850
	Speaking	0.50	2876.977	0.822	2554.752	0.802
	Oral		1988.997	0.894	1922.695	0.888
9-12	Listening	0.50	2574.848	0.860	2606.657	0.860
	Speaking	0.50	3012.001	0.851	2565.447	0.813
	Oral		2135.088	0.905	1994.886	0.894

Table 3.3.1.4

Reliability of Composite: Oral S403 Online by IEP status

Cluster	Component	Weight	Variance	Reliability
1	Listening	0.50	3446.907	0.890
	Speaking	0.50	2981.567	0.832
	Oral		2423.299	0.909
2-3	Listening	0.50	3436.539	0.870
	Speaking	0.50	2604.046	0.794
	Oral		2372.396	0.896
4-5	Listening	0.50	2289.802	0.810
	Speaking	0.50	2063.239	0.815
	Oral		1596.003	0.872
6-8	Listening	0.50	1823.462	0.830
	Speaking	0.50	2544.711	0.817
	Oral		1596.423	0.878
9-12	Listening	0.50	2053.781	0.830
	Speaking	0.50	2822.043	0.860
	Oral		1752.972	0.894

3.3.2 Literacy

Table 3.3.2.1

Reliability of Composite: Litr S403 Online

Cluster	Component	Weight	Variance	Reliability
1	Reading	0.50	1115.802	0.890
	Writing	0.50	1176.138	0.869
	Literacy		864.454	0.920
2-3	Reading	0.50	1126.997	0.880
	Writing	0.50	1639.481	0.908
	Literacy		1133.555	0.937
4-5	Reading	0.50	1143.278	0.890
	Writing	0.50	1741.556	0.914
	Literacy		1170.649	0.941
6-8	Reading	0.50	1534.276	0.910
	Writing	0.50	1851.229	0.894
	Literacy		1486.355	0.944
9-12	Reading	0.50	1424.066	0.910
	Writing	0.50	1774.853	0.903
	Literacy		1303.621	0.942

Table 3.3.2.2

Reliability of Composite: Litr S403 Online by Gender

			Female		Male	
Cluster	Component	Weight	Variance	Reliability	Variance	Reliability
1	Reading	0.50	1110.842	0.890	1112.741	0.890
	Writing	0.50	1093.523	0.863	1210.066	0.871
	Literacy		842.268	0.919	864.615	0.919
2-3	Reading	0.50	1103.924	0.880	1138.013	0.880
	Writing	0.50	1574.199	0.902	1627.394	0.909
	Literacy		1103.379	0.935	1127.416	0.937
4-5	Reading	0.50	1076.266	0.880	1188.426	0.890
	Writing	0.50	1610.724	0.905	1778.977	0.917
	Literacy		1092.514	0.935	1202.119	0.942
6-8	Reading	0.50	1461.439	0.910	1567.026	0.920
	Writing	0.50	1757.358	0.884	1851.998	0.898
	Literacy		1410.234	0.941	1500.519	0.948
9-12	Reading	0.50	1380.662	0.910	1429.940	0.910
	Writing	0.50	1635.265	0.895	1806.132	0.905
	Literacy		1228.445	0.940	1312.387	0.943

Table 3.3.2.3

Reliability of Composite: Litr S403 Online by Ethnicity

Cluster	Component	Weight	Hispanic		Other	
			Variance	Reliability	Variance	Reliability
1	Reading	0.50	963.914	0.870	1292.193	0.900
	Writing	0.50	1127.099	0.869	1143.819	0.864
	Literacy		754.372	0.910	962.501	0.926
2-3	Reading	0.50	1055.523	0.880	1209.920	0.880
	Writing	0.50	1650.264	0.908	1483.800	0.904
	Literacy		1100.444	0.937	1112.968	0.935
4-5	Reading	0.50	1059.421	0.880	1265.346	0.890
	Writing	0.50	1688.381	0.913	1513.377	0.909
	Literacy		1106.833	0.938	1140.136	0.939
6-8	Reading	0.50	1430.122	0.910	1657.833	0.920
	Writing	0.50	1800.056	0.894	1668.782	0.881
	Literacy		1414.232	0.944	1456.688	0.943
9-12	Reading	0.50	1362.089	0.910	1430.754	0.910
	Writing	0.50	1752.592	0.902	1561.495	0.894
	Literacy		1259.401	0.941	1210.989	0.939

Table 3.3.2.4

Reliability of Composite: Litr S403 Online by IEP status

Cluster	Component	Weight	Variance	Reliability
1	Reading	0.50	846.560	0.850
	Writing	0.50	1492.207	0.882
	Literacy		783.304	0.903
2-3	Reading	0.50	943.809	0.860
	Writing	0.50	1720.396	0.906
	Literacy		1041.603	0.930
4-5	Reading	0.50	1001.307	0.870
	Writing	0.50	1566.409	0.909
	Literacy		1013.339	0.933
6-8	Reading	0.50	1186.531	0.890
	Writing	0.50	1389.011	0.868
	Literacy		1095.603	0.928
9-12	Reading	0.50	961.703	0.870
	Writing	0.50	1381.965	0.886
	Literacy		848.712	0.917

3.3.3 Comprehension

Table 3.3.3.1

Reliability of Composite: Cphn S403 Online

Cluster	Component	Weight	Variance	Reliability
1	Listening	0.30	2974.123	0.850
	Reading	0.70	1115.802	0.890
	Comprehension		1117.960	0.910
2-3	Listening	0.30	3295.516	0.840
	Reading	0.70	1126.997	0.880
	Comprehension		1332.680	0.915
4-5	Listening	0.30	2624.456	0.810
	Reading	0.70	1143.278	0.890
	Comprehension		1277.643	0.917
6-8	Listening	0.30	2367.896	0.850
	Reading	0.70	1534.276	0.910
	Comprehension		1512.233	0.934
9-12	Listening	0.30	2630.452	0.860
	Reading	0.70	1424.066	0.910
	Comprehension		1495.504	0.936

Table 3.3.3.2

Reliability of Composite: Cphn S403 Online by Gender

			Female		Male	
Cluster	Component	Weight	Variance	Reliability	Variance	Reliability
1	Listening	0.30	2827.260	0.840	3057.022	0.860
	Reading	0.70	1110.842	0.890	1112.741	0.890
	Comprehension		1097.243	0.908	1121.609	0.912
2-3	Listening	0.30	3113.919	0.830	3427.582	0.840
	Reading	0.70	1103.924	0.880	1138.013	0.880
	Comprehension		1289.639	0.913	1358.254	0.914
4-5	Listening	0.30	2452.009	0.800	2724.672	0.810
	Reading	0.70	1076.266	0.880	1188.426	0.890
	Comprehension		1197.539	0.910	1328.557	0.917
6-8	Listening	0.30	2254.865	0.850	2427.926	0.860
	Reading	0.70	1461.439	0.910	1567.026	0.920
	Comprehension		1448.671	0.934	1540.992	0.940
9-12	Listening	0.30	2530.143	0.860	2690.832	0.860
	Reading	0.70	1380.662	0.910	1429.940	0.910
	Comprehension		1450.367	0.936	1508.094	0.936

Table 3.3.3.3

Reliability of Composite: Cphn S403 Online by Ethnicity

			Hispanic		**Other**	
Cluster	**Component**	**Weight**	**Variance**	**Reliability**	**Variance**	**Reliability**
1	Listening	0.30	2882.609	0.850	3006.747	0.850
	Reading	0.70	963.914	0.870	1292.193	0.900
	Comprehension		972.825	0.897	1289.474	0.919
2-3	Listening	0.30	3217.327	0.840	3267.025	0.830
	Reading	0.70	1055.523	0.880	1209.920	0.880
	Comprehension		1251.984	0.913	1409.042	0.914
4-5	Listening	0.30	2479.652	0.800	2548.110	0.800
	Reading	0.70	1059.421	0.880	1265.346	0.890
	Comprehension		1180.790	0.909	1361.030	0.916
6-8	Listening	0.30	2228.071	0.850	2378.218	0.850
	Reading	0.70	1430.122	0.910	1657.833	0.920
	Comprehension		1396.758	0.933	1614.473	0.940
9-12	Listening	0.30	2574.848	0.860	2606.657	0.860
	Reading	0.70	1362.089	0.910	1430.754	0.910
	Comprehension		1432.848	0.935	1499.543	0.936

Table 3.3.3.4

Reliability of Composite: Cphn S403 Online by IEP status

Cluster	**Component**	**Weight**	**Variance**	**Reliability**
1	Listening	0.30	3446.907	0.890
	Reading	0.70	846.560	0.850
	Comprehension		922.482	0.896
2-3	Listening	0.30	3436.539	0.870
	Reading	0.70	943.809	0.860
	Comprehension		1144.237	0.908
4-5	Listening	0.30	2289.802	0.810
	Reading	0.70	1001.307	0.870
	Comprehension		1048.379	0.902
6-8	Listening	0.30	1823.462	0.830
	Reading	0.70	1186.531	0.890
	Comprehension		1088.674	0.916
9-12	Listening	0.30	2053.781	0.830
	Reading	0.70	961.703	0.870
	Comprehension		983.063	0.906

3.3.4 Overall

Table 3.3.4.1

Reliability of Composite: Over S403 Online

Cluster	Component	Weight	Variance	Reliability
1	Listening	0.15	2974.123	0.850
	Reading	0.35	1115.802	0.890
	Speaking	0.15	2712.492	0.832
	Writing	0.35	1176.138	0.869
	Overall Composite		918.310	0.941
2-3	Listening	0.15	3295.516	0.840
	Reading	0.35	1126.997	0.880
	Speaking	0.15	2285.756	0.791
	Writing	0.35	1639.481	0.908
	Overall Composite		1212.877	0.952
4-5	Listening	0.15	2624.456	0.810
	Reading	0.35	1143.278	0.890
	Speaking	0.15	2311.342	0.816
	Writing	0.35	1741.556	0.914
	Overall Composite		1211.162	0.955
6-8	Listening	0.15	2367.896	0.850
	Reading	0.35	1534.276	0.910
	Speaking	0.15	2942.663	0.820
	Writing	0.35	1851.229	0.894
	Overall Composite		1467.804	0.959
9-12	Listening	0.15	2630.452	0.860
	Reading	0.35	1424.066	0.910
	Speaking	0.15	2980.093	0.844
	Writing	0.35	1774.853	0.903
	Overall Composite		1366.954	0.959

Table 3.3.4.2

Reliability of Composite: Over S403 Online by Gender

Cluster	Component	Weight	Female		Male	
			Variance	Reliability	Variance	Reliability
1	Listening	0.15	2827.260	0.840	3057.022	0.860
	Reading	0.35	1110.842	0.890	1112.741	0.890
	Speaking	0.15	2664.132	0.834	2672.046	0.828
	Writing	0.35	1093.523	0.863	1210.066	0.871
	Overall Composite		881.240	0.939	921.315	0.941
2-3	Listening	0.15	3113.919	0.830	3427.582	0.840
	Reading	0.35	1103.924	0.880	1138.013	0.880
	Speaking	0.15	2244.973	0.789	2245.847	0.788
	Writing	0.35	1574.199	0.902	1627.394	0.909
	Overall Composite		1167.193	0.951	1217.261	0.952
4-5	Listening	0.15	2452.009	0.800	2724.672	0.810
	Reading	0.35	1076.266	0.880	1188.426	0.890
	Speaking	0.15	2247.115	0.814	2325.015	0.818
	Writing	0.35	1610.724	0.905	1778.977	0.917
	Overall Composite		1128.342	0.951	1247.546	0.956
6-8	Listening	0.15	2254.865	0.850	2427.926	0.860
	Reading	0.35	1461.439	0.910	1567.026	0.920
	Speaking	0.15	2943.614	0.823	2895.799	0.818
	Writing	0.35	1757.358	0.884	1851.998	0.898
	Overall Composite		1412.417	0.957	1473.199	0.961
9-12	Listening	0.15	2530.143	0.860	2690.832	0.860
	Reading	0.35	1380.662	0.910	1429.940	0.910
	Speaking	0.15	2901.633	0.837	3003.107	0.849
	Writing	0.35	1635.265	0.895	1806.132	0.905
	Overall Composite		1300.070	0.958	1378.293	0.960

Table 3.3.4.3

Reliability of Composite: Over S403 Online by Ethnicity

			Hispanic		Other	
Cluster	**Component**	**Weight**	**Variance**	**Reliability**	**Variance**	**Reliability**
1	Listening	0.15	2882.609	0.850	3006.747	0.850
	Reading	0.35	963.914	0.870	1292.193	0.900
	Speaking	0.15	2706.897	0.834	2580.631	0.823
	Writing	0.35	1127.099	0.869	1143.819	0.864
	Overall Composite		823.422	0.935	986.143	0.944
2-3	Listening	0.15	3217.327	0.840	3267.025	0.830
	Reading	0.35	1055.523	0.880	1209.920	0.880
	Speaking	0.15	2296.300	0.796	2134.472	0.774
	Writing	0.35	1650.264	0.908	1483.800	0.904
	Overall Composite		1176.243	0.952	1182.019	0.950
4-5	Listening	0.15	2479.652	0.800	2548.110	0.800
	Reading	0.35	1059.421	0.880	1265.346	0.890
	Speaking	0.15	2224.972	0.816	2160.747	0.808
	Writing	0.35	1688.381	0.913	1513.377	0.909
	Overall Composite		1134.210	0.953	1159.883	0.953
6-8	Listening	0.15	2228.071	0.850	2378.218	0.850
	Reading	0.35	1430.122	0.910	1657.833	0.920
	Speaking	0.15	2876.977	0.822	2554.752	0.802
	Writing	0.35	1800.056	0.894	1668.782	0.881
	Overall Composite		1383.269	0.958	1407.438	0.957
9-12	Listening	0.15	2574.848	0.860	2606.657	0.860
	Reading	0.35	1362.089	0.910	1430.754	0.910
	Speaking	0.15	3012.001	0.851	2565.447	0.813
	Writing	0.35	1752.592	0.902	1561.495	0.894
	Overall Composite		1321.560	0.959	1260.077	0.956

Table 3.3.4.4

Reliability of Composite: Over S403 Online by IEP status

Cluster	Component	Weight	Variance	Reliability
1	Listening	0.15	3446.907	0.890
	Reading	0.35	846.560	0.850
	Speaking	0.15	2981.567	0.832
	Writing	0.35	1492.207	0.882
	Overall Composite		873.346	0.935
2-3	Listening	0.15	3436.539	0.870
	Reading	0.35	943.809	0.860
	Speaking	0.15	2604.046	0.794
	Writing	0.35	1720.396	0.906
	Overall Composite		1148.034	0.949
4-5	Listening	0.15	2289.802	0.810
	Reading	0.35	1001.307	0.870
	Speaking	0.15	2063.239	0.815
	Writing	0.35	1566.409	0.909
	Overall Composite		975.019	0.947
6-8	Listening	0.15	1823.462	0.830
	Reading	0.35	1186.531	0.890
	Speaking	0.15	2544.711	0.817
	Writing	0.35	1389.011	0.868
	Overall Composite		1011.906	0.945
9-12	Listening	0.15	2053.781	0.830
	Reading	0.35	961.703	0.870
	Speaking	0.15	2822.043	0.860
	Writing	0.35	1381.965	0.886
	Overall Composite		884.234	0.942

3.4 Accuracy and Consistency of Composites

Tables below present three sections of information related to the accuracy and consistency of placement into the WIDA language proficiency levels for each composite score. The first section provides overall indices related to the accuracy and consistency of classification, as well as Cohen’s kappa. The second section shows accuracy and consistency information conditional per proficiency level. The third section provides indices of classification accuracy, including the false-positives and false-negatives, 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 (e.g., determining which students have reached PL 5). Note that the consistency is generally higher at the cut points than for the proficiency levels. For practical purposes, the primary score used for such decisions is the overall composite score.

As noted above in Section 2.13, there may be cases where the number of test takers placed into the proficiency level is fewer than 200 and accuracy and consistency of classification conditional on that level either cannot be computed or the software produces estimates that are out of bounds. In these cases, ‘N/A’ has been placed in the table. In addition, there may be cases where due to the small percentage of test takers placed into the proficiency level and the range of observed scale scores, accuracy of classification conditional on that level cannot be estimated by the software program that is used. In such cases, a hyphen (-) has been placed in the table.

3.4.1 Oral

Table 3.4.1.1

Accuracy and Consistency of Classification Indices: Oral (Grade 1) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.669	0.560		0.424	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.880		0.790	
	2	0.608		0.476	
	3	0.753		0.648	
	4	0.599		0.503	
	5	0.581		0.462	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.970	0.010	0.020	0.958
	2/3	0.937	0.033	0.030	0.908
	3/4	0.884	0.046	0.069	0.839
	4/5	0.889	0.063	0.048	0.847
	5/6	0.985	0.015	0.000	0.982

Table 3.4.1.2

Accuracy and Consistency of Classification Indices: Oral (Grade 2) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.688	0.582		0.435	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.841		0.736	
	2	0.712		0.594	
	3	0.727		0.621	
	4	0.590		0.521	
	5	-		-	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.964	0.015	0.021	0.948
	2/3	0.913	0.040	0.048	0.876
	3/4	0.872	0.051	0.077	0.820
	4/5	0.939	0.061	0.000	0.919
	5/6	0.998	0.002	0.000	0.998

Table 3.4.1.3

Accuracy and Consistency of Classification Indices: Oral (Grade 3) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.686	0.579		0.415	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.846		0.749	
	2	0.704		0.575	
	3	0.714		0.581	
	4	0.625		0.569	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.970	0.014	0.016	0.957
	2/3	0.926	0.032	0.042	0.895
	3/4	0.840	0.051	0.109	0.777
	4/5	0.949	0.051	0.000	0.930
	5/6	N/A	N/A	N/A	N/A

Table 3.4.1.4

Accuracy and Consistency of Classification Indices: Oral (Grade 4) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.665	0.556		0.390	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.900		0.812	
	2	0.570		0.425	
	3	0.730		0.601	
	4	0.716		0.595	
	5	0.590		0.530	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.989	0.003	0.008	0.986
	2/3	0.973	0.015	0.012	0.959
	3/4	0.918	0.034	0.048	0.885
	4/5	0.851	0.052	0.097	0.791
	5/6	0.933	0.067	0.000	0.913

Table 3.4.1.5

Accuracy and Consistency of Classification Indices: Oral (Grade 5) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.638	0.531		0.357	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.893		0.804	
	2	0.544		0.400	
	3	0.730		0.599	
	4	0.689		0.568	
	5	0.530		0.467	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.986	0.004	0.009	0.981
	2/3	0.966	0.019	0.014	0.949
	3/4	0.911	0.033	0.056	0.877
	4/5	0.825	0.058	0.117	0.761
	5/6	0.945	0.055	0.000	0.929

Table 3.4.1.6

Accuracy and Consistency of Classification Indices: Oral (Grade 6) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.725	0.630		0.468	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.863		0.775	
	2	0.643		0.506	
	3	0.806		0.714	
	4	0.673		0.631	
	5	-		-	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.981	0.007	0.012	0.973
	2/3	0.945	0.030	0.025	0.921
	3/4	0.881	0.043	0.076	0.836
	4/5	0.917	0.083	0.000	0.897
	5/6	0.992	0.008	0.000	0.992

Table 3.4.1.7

Accuracy and Consistency of Classification Indices: Oral (Grade 7) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.703	0.599		0.449	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.860		0.761	
	2	0.622		0.489	
	3	0.769		0.672	
	4	0.659		0.598	
	5	0.466		0.316	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.971	0.011	0.018	0.958
	2/3	0.930	0.036	0.033	0.900
	3/4	0.881	0.046	0.073	0.835
	4/5	0.924	0.068	0.008	0.902
	5/6	0.988	0.012	0.000	0.988

Table 3.4.1.8

Accuracy and Consistency of Classification Indices: Oral (Grade 8) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.686	0.580		0.435	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.854		0.758	
	2	0.622		0.490	
	3	0.735		0.627	
	4	0.633		0.582	
	5	-		-	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.964	0.016	0.020	0.948
	2/3	0.924	0.037	0.039	0.892
	3/4	0.881	0.045	0.074	0.834
	4/5	0.917	0.083	0.000	0.895
	5/6	0.988	0.012	0.000	0.987

Table 3.4.1.9

Accuracy and Consistency of Classification Indices: Oral (Grade 9) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.743	0.647		0.520	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.864		0.779	
	2	0.682		0.572	
	3	0.771		0.690	
	4	0.672		0.579	
	5	0.552		0.310	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.947	0.022	0.031	0.925
	2/3	0.907	0.048	0.046	0.869
	3/4	0.915	0.042	0.043	0.879
	4/5	0.975	0.025	0.000	0.969
	5/6	0.997	0.003	0.000	0.997

Table 3.4.1.10

Accuracy and Consistency of Classification Indices: Oral (Grade 10) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.741	0.645		0.515	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.859		0.774	
	2	0.652		0.536	
	3	0.783		0.706	
	4	0.661		0.566	
	5	0.545		0.325	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.945	0.024	0.032	0.921
	2/3	0.906	0.048	0.046	0.868
	3/4	0.916	0.042	0.043	0.880
	4/5	0.974	0.026	0.000	0.969
	5/6	0.997	0.003	0.000	0.997

Table 3.4.1.11

Accuracy and Consistency of Classification Indices: Oral (Grade 11) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.736	0.640		0.508	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.856		0.771	
	2	0.642		0.524	
	3	0.775		0.697	
	4	0.651		0.554	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.942	0.026	0.032	0.916
	2/3	0.905	0.047	0.048	0.866
	3/4	0.915	0.042	0.043	0.878
	4/5	0.975	0.026	0.000	0.969
	5/6	N/A	N/A	N/A	N/A

Table 3.4.1.12

Accuracy and Consistency of Classification Indices: Oral (Grade 12) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.744	0.650		0.512	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.864		0.784	
	2	0.625		0.505	
	3	0.788		0.715	
	4	0.633		0.519	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.939	0.028	0.033	0.913
	2/3	0.904	0.048	0.049	0.864
	3/4	0.917	0.041	0.042	0.880
	4/5	0.984	0.016	0.000	0.980
	5/6	N/A	N/A	N/A	N/A

3.4.2 Literacy

Table 3.4.2.1

Accuracy and Consistency of Classification Indices: Litr (Grade 1) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.810	0.731		0.605	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.796		0.694	
	2	0.837		0.777	
	3	0.786		0.709	
	4	0.783		0.671	
	5	0.853		0.685	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.931	0.037	0.032	0.900
	2/3	0.912	0.039	0.049	0.875
	3/4	0.973	0.019	0.008	0.963
	4/5	0.994	0.005	0.001	0.993
	5/6	N/A	N/A	N/A	N/A

Table 3.4.2.2

Accuracy and Consistency of Classification Indices: Litr (Grade 2) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.823	0.755		0.631	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.829		0.742	
	2	0.794		0.705	
	3	0.865		0.822	
	4	0.747		0.667	
	5	0.710		0.493	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.973	0.013	0.014	0.963
	2/3	0.933	0.032	0.036	0.907
	3/4	0.934	0.034	0.032	0.908
	4/5	0.980	0.018	0.001	0.978
	5/6	0.998	0.002	0.000	0.999

Table 3.4.2.3

Accuracy and Consistency of Classification Indices: Litr (Grade 3) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.802	0.724		0.604	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.864		0.781	
	2	0.754		0.651	
	3	0.848		0.793	
	4	0.769		0.695	
	5	0.606		0.451	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.977	0.009	0.013	0.968
	2/3	0.946	0.026	0.029	0.923
	3/4	0.919	0.040	0.041	0.886
	4/5	0.964	0.027	0.010	0.950
	5/6	0.994	0.006	0.000	0.994

Table 3.4.2.4

Accuracy and Consistency of Classification Indices: Litr (Grade 4) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.786	0.712		0.582	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.853		0.793	
	2	0.707		0.584	
	3	0.858		0.797	
	4	0.765		0.715	
	5	0.543		0.410	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.983	0.008	0.009	0.978
	2/3	0.963	0.018	0.019	0.949
	3/4	0.919	0.035	0.046	0.888
	4/5	0.925	0.057	0.018	0.903
	5/6	0.989	0.011	0.000	0.990

Table 3.4.2.5

Accuracy and Consistency of Classification Indices: Litr (Grade 5) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.771	0.692		0.563	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.879		0.814	
	2	0.715		0.596	
	3	0.844		0.775	
	4	0.754		0.701	
	5	0.514		0.393	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.983	0.008	0.010	0.976
	2/3	0.963	0.018	0.019	0.948
	3/4	0.921	0.034	0.046	0.889
	4/5	0.912	0.062	0.026	0.885
	5/6	0.987	0.013	0.000	0.987

Table 3.4.2.6

Accuracy and Consistency of Classification Indices: Litr (Grade 6) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.821	0.750		0.643	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.918		0.868	
	2	0.729		0.627	
	3	0.863		0.811	
	4	0.732		0.646	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.964	0.014	0.022	0.950
	2/3	0.941	0.031	0.028	0.916
	3/4	0.926	0.034	0.040	0.896
	4/5	0.990	0.010	0.000	0.988
	5/6	N/A	N/A	N/A	N/A

Table 3.4.2.7

Accuracy and Consistency of Classification Indices: Litr (Grade 7) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.810	0.734		0.634	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.904		0.849	
	2	0.740		0.643	
	3	0.838		0.778	
	4	0.745		0.668	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.958	0.017	0.024	0.941
	2/3	0.936	0.032	0.032	0.909
	3/4	0.933	0.033	0.034	0.905
	4/5	0.983	0.017	0.000	0.977
	5/6	N/A	N/A	N/A	N/A

Table 3.4.2.8

Accuracy and Consistency of Classification Indices: Litr (Grade 8) S403 Online

<table>
<tr><th rowspan="2">Overall Indices</th><th>Accuracy</th><th colspan="2">Consistency</th><th colspan="2">Kappa (k)</th></tr>
<tr><td>0.800</td><td colspan="2">0.720</td><td colspan="2">0.622</td></tr>
<tr><th rowspan="7">Conditional on Level</th><th>Level</th><th colspan="2">Accuracy</th><th colspan="2">Consistency</th></tr>
<tr><td>1</td><td colspan="2">0.911</td><td colspan="2">0.860</td></tr>
<tr><td>2</td><td colspan="2">0.733</td><td colspan="2">0.631</td></tr>
<tr><td>3</td><td colspan="2">0.815</td><td colspan="2">0.744</td></tr>
<tr><td>4</td><td colspan="2">0.739</td><td colspan="2">0.669</td></tr>
<tr><td>5</td><td colspan="2">-</td><td colspan="2">-</td></tr>
<tr><td>6</td><td colspan="2">N/A</td><td colspan="2">N/A</td></tr>
<tr><th rowspan="7">Indices at Proficiency Level Cut Points</th><th rowspan="2">Proficiency Level Cut Point</th><th colspan="3">Accuracy</th><th rowspan="2">Consistency</th></tr>
<tr><th>Accuracy</th><th>False Positives</th><th>False Negatives</th></tr>
<tr><td>1/2</td><td>0.958</td><td>0.018</td><td>0.024</td><td>0.941</td></tr>
<tr><td>2/3</td><td>0.937</td><td>0.031</td><td>0.032</td><td>0.911</td></tr>
<tr><td>3/4</td><td>0.927</td><td>0.036</td><td>0.038</td><td>0.897</td></tr>
<tr><td>4/5</td><td>0.978</td><td>0.022</td><td>0.000</td><td>0.970</td></tr>
<tr><td>5/6</td><td>N/A</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
</table>

Table 3.4.2.9

Accuracy and Consistency of Classification Indices: Litr (Grade 9) S403 Online

<table>
<tr><th rowspan="2">Overall Indices</th><th>Accuracy</th><th colspan="2">Consistency</th><th colspan="2">Kappa (k)</th></tr>
<tr><td>0.796</td><td colspan="2">0.714</td><td colspan="2">0.617</td></tr>
<tr><th rowspan="7">Conditional on Level</th><th>Level</th><th colspan="2">Accuracy</th><th colspan="2">Consistency</th></tr>
<tr><td>1</td><td colspan="2">0.903</td><td colspan="2">0.840</td></tr>
<tr><td>2</td><td colspan="2">0.725</td><td colspan="2">0.626</td></tr>
<tr><td>3</td><td colspan="2">0.824</td><td colspan="2">0.759</td></tr>
<tr><td>4</td><td colspan="2">0.776</td><td colspan="2">0.693</td></tr>
<tr><td>5</td><td colspan="2">0.673</td><td colspan="2">0.518</td></tr>
<tr><td>6</td><td colspan="2">-</td><td colspan="2">-</td></tr>
<tr><th rowspan="7">Indices at Proficiency Level Cut Points</th><th rowspan="2">Proficiency Level Cut Point</th><th colspan="3">Accuracy</th><th rowspan="2">Consistency</th></tr>
<tr><th>Accuracy</th><th>False Positives</th><th>False Negatives</th></tr>
<tr><td>1/2</td><td>0.963</td><td>0.013</td><td>0.023</td><td>0.949</td></tr>
<tr><td>2/3</td><td>0.939</td><td>0.031</td><td>0.030</td><td>0.914</td></tr>
<tr><td>3/4</td><td>0.929</td><td>0.037</td><td>0.034</td><td>0.900</td></tr>
<tr><td>4/5</td><td>0.968</td><td>0.020</td><td>0.012</td><td>0.954</td></tr>
<tr><td>5/6</td><td>0.997</td><td>0.003</td><td>0.000</td><td>0.997</td></tr>
</table>

Table 3.4.2.10

Accuracy and Consistency of Classification Indices: Litr (Grade 10) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.808	0.729		0.629	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.882		0.803	
	2	0.763		0.671	
	3	0.827		0.767	
	4	0.796		0.718	
	5	0.753		0.592	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.969	0.012	0.019	0.957
	2/3	0.940	0.029	0.031	0.915
	3/4	0.929	0.039	0.033	0.900
	4/5	0.970	0.018	0.011	0.957
	5/6	N/A	N/A	N/A	N/A

Table 3.4.2.11

Accuracy and Consistency of Classification Indices: Litr (Grade 11) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.806	0.727		0.627	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.873		0.790	
	2	0.770		0.680	
	3	0.820		0.758	
	4	0.799		0.721	
	5	0.762		0.603	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.969	0.012	0.019	0.956
	2/3	0.938	0.029	0.033	0.913
	3/4	0.929	0.039	0.032	0.900
	4/5	0.970	0.018	0.011	0.957
	5/6	N/A	N/A	N/A	N/A

Table 3.4.2.12

Accuracy and Consistency of Classification Indices: Litr (Grade 12) S403 Online

Overall Indices	Accuracy	Consistency	Kappa (k)
	0.811	0.734	0.632

Conditional on Level	Level	Accuracy	Consistency
	1	0.873	0.794
	2	0.776	0.689
	3	0.827	0.769
	4	0.795	0.711
	5	0.740	0.547
	6	N/A	N/A

Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.964	0.015	0.021	0.949
	2/3	0.933	0.031	0.036	0.905
	3/4	0.935	0.036	0.028	0.909
	4/5	0.980	0.014	0.006	0.971
	5/6	N/A	N/A	N/A	N/A

3.4.3 Comprehension

Table 3.4.3.1

Accuracy and Consistency of Classification Indices: Cphn (Grade 1) S403 Online

Overall Indices	Accuracy	Consistency	Kappa (k)
	0.644	0.539	0.435

Conditional on Level	Level	Accuracy	Consistency
	1	0.762	0.602
	2	0.604	0.486
	3	0.628	0.520
	4	0.422	0.324
	5	0.635	0.518
	6	0.850	0.759

Indices at Proficiency Level Cut Points	Proficiency Level Cut Point	Accuracy			Consistency
		Accuracy	False Positives	False Negatives	
	1/2	0.957	0.015	0.028	0.939
	2/3	0.920	0.040	0.041	0.886
	3/4	0.898	0.055	0.046	0.859
	4/5	0.906	0.049	0.045	0.869
	5/6	0.943	0.031	0.026	0.918

Table 3.4.3.2

Accuracy and Consistency of Classification Indices: Cphn (Grade 2) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.664	0.559		0.461	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.814		0.680	
	2	0.691		0.582	
	3	0.648		0.539	
	4	0.450		0.345	
	5	0.632		0.514	
	6	0.847		0.749	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.964	0.013	0.023	0.949
	2/3	0.923	0.037	0.041	0.891
	3/4	0.907	0.047	0.046	0.869
	4/5	0.914	0.047	0.039	0.879
	5/6	0.944	0.033	0.023	0.920

Table 3.4.3.3

Accuracy and Consistency of Classification Indices: Cphn (Grade 3) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.656	0.555		0.459	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.836		0.721	
	2	0.656		0.539	
	3	0.612		0.501	
	4	0.404		0.307	
	5	0.590		0.476	
	6	0.861		0.772	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.962	0.014	0.024	0.946
	2/3	0.926	0.034	0.039	0.896
	3/4	0.906	0.047	0.046	0.869
	4/5	0.910	0.047	0.043	0.873
	5/6	0.932	0.041	0.027	0.904

Table 3.4.3.4

Accuracy and Consistency of Classification Indices: Cphn (Grade 4) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.714	0.625		0.497	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.860		0.735	
	2	0.698		0.571	
	3	0.619		0.489	
	4	0.459		0.349	
	5	0.637		0.529	
	6	0.873		0.819	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.988	0.003	0.009	0.984
	2/3	0.967	0.013	0.020	0.952
	3/4	0.934	0.032	0.034	0.905
	4/5	0.910	0.044	0.046	0.874
	5/6	0.904	0.046	0.050	0.865

Table 3.4.3.5

Accuracy and Consistency of Classification Indices: Cphn (Grade 5) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.691	0.601		0.479	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.876		0.772	
	2	0.654		0.525	
	3	0.585		0.457	
	4	0.443		0.338	
	5	0.605		0.496	
	6	0.860		0.801	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.982	0.005	0.013	0.975
	2/3	0.961	0.018	0.022	0.943
	3/4	0.925	0.038	0.036	0.894
	4/5	0.903	0.046	0.050	0.867
	5/6	0.903	0.045	0.052	0.864

Table 3.4.3.6

Accuracy and Consistency of Classification Indices: Cphn (Grade 6) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.697	0.597		0.505	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.890		0.816	
	2	0.747		0.653	
	3	0.682		0.577	
	4	0.528		0.417	
	5	0.649		0.533	
	6	0.774		0.635	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.965	0.012	0.023	0.951
	2/3	0.932	0.035	0.033	0.903
	3/4	0.912	0.048	0.040	0.878
	4/5	0.922	0.040	0.039	0.890
	5/6	0.960	0.023	0.017	0.942

Table 3.4.3.7

Accuracy and Consistency of Classification Indices: Cphn (Grade 7) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.689	0.590		0.502	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.891		0.820	
	2	0.705		0.604	
	3	0.654		0.546	
	4	0.513		0.402	
	5	0.593		0.471	
	6	0.831		0.720	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.955	0.016	0.029	0.938
	2/3	0.927	0.039	0.034	0.897
	3/4	0.914	0.048	0.039	0.880
	4/5	0.927	0.038	0.035	0.897
	5/6	0.957	0.025	0.018	0.939

Table 3.4.3.8

Accuracy and Consistency of Classification Indices: Cphn (Grade 8) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.682	0.584		0.497	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.893		0.826	
	2	0.684		0.576	
	3	0.628		0.516	
	4	0.480		0.373	
	5	0.588		0.470	
	6	0.831		0.725	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.953	0.018	0.029	0.935
	2/3	0.928	0.038	0.034	0.898
	3/4	0.914	0.046	0.040	0.881
	4/5	0.924	0.039	0.037	0.893
	5/6	0.951	0.029	0.021	0.929

Table 3.4.3.9

Accuracy and Consistency of Classification Indices: Cphn (Grade 9) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.707	0.610		0.522	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.858		0.770	
	2	0.758		0.673	
	3	0.656		0.544	
	4	0.482		0.371	
	5	0.635		0.515	
	6	0.845		0.750	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.952	0.019	0.030	0.932
	2/3	0.927	0.037	0.036	0.897
	3/4	0.927	0.039	0.034	0.897
	4/5	0.936	0.036	0.028	0.910
	5/6	0.959	0.023	0.017	0.942

Table 3.4.3.10

Accuracy and Consistency of Classification Indices: Cphn (Grade 10) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.703	0.605		0.519	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.868		0.784	
	2	0.738		0.646	
	3	0.664		0.554	
	4	0.482		0.372	
	5	0.627		0.508	
	6	0.842		0.747	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.953	0.018	0.029	0.934
	2/3	0.928	0.036	0.036	0.899
	3/4	0.925	0.041	0.034	0.894
	4/5	0.932	0.038	0.030	0.905
	5/6	0.956	0.024	0.019	0.938

Table 3.4.3.11

Accuracy and Consistency of Classification Indices: Cphn (Grade 11) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.698	0.601		0.515	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.869		0.786	
	2	0.740		0.647	
	3	0.637		0.524	
	4	0.455		0.349	
	5	0.643		0.526	
	6	0.839		0.743	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.953	0.018	0.029	0.935
	2/3	0.929	0.035	0.036	0.899
	3/4	0.924	0.042	0.035	0.893
	4/5	0.928	0.040	0.032	0.900
	5/6	0.954	0.025	0.021	0.934

Table 3.4.3.12

Accuracy and Consistency of Classification Indices: Cphn (Grade 12) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.703	0.605		0.516	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.872		0.792	
	2	0.740		0.649	
	3	0.647		0.537	
	4	0.505		0.393	
	5	0.644		0.523	
	6	0.834		0.725	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.950	0.020	0.030	0.930
	2/3	0.925	0.037	0.038	0.894
	3/4	0.923	0.043	0.034	0.893
	4/5	0.934	0.037	0.029	0.908
	5/6	0.964	0.021	0.015	0.947

3.4.4 Overall

Table 3.4.4.1

Accuracy and Consistency of Classification Indices: Over (Grade 1) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.841	0.776		0.662	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.841		0.746	
	2	0.819		0.749	
	3	0.872		0.830	
	4	0.756		0.653	
	5	0.889		0.742	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			
		Accuracy	**False Positives**	**False Negatives**	**Consistency**
	1/2	0.966	0.016	0.018	0.951
	2/3	0.924	0.038	0.039	0.892
	3/4	0.960	0.023	0.017	0.943
	4/5	0.991	0.008	0.001	0.989
	5/6	N/A	N/A	N/A	N/A

Table 3.4.4.2

Accuracy and Consistency of Classification Indices: Over (Grade 2) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.839	0.779		0.673	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.874		0.798	
	2	0.830		0.756	
	3	0.878		0.835	
	4	0.753		0.690	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.979	0.009	0.011	0.970
	2/3	0.945	0.026	0.029	0.922
	3/4	0.937	0.030	0.034	0.910
	4/5	0.979	0.021	0.000	0.978
	5/6	N/A	N/A	N/A	N/A

Table 3.4.4.3

Accuracy and Consistency of Classification Indices: Over (Grade 3) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.824	0.763		0.654	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.895		0.831	
	2	0.784		0.693	
	3	0.876		0.825	
	4	0.763		0.717	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.982	0.008	0.011	0.974
	2/3	0.954	0.023	0.023	0.935
	3/4	0.924	0.032	0.044	0.893
	4/5	0.964	0.036	0.000	0.961
	5/6	N/A	N/A	N/A	N/A

Table 3.4.4.4

Accuracy and Consistency of Classification Indices: Over (Grade 4) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.799	0.734		0.610	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.805		0.840	
	2	0.754		0.650	
	3	0.874		0.816	
	4	0.815		0.755	
	5	0.580		0.482	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.986	0.008	0.005	0.987
	2/3	0.972	0.017	0.011	0.966
	3/4	0.933	0.031	0.036	0.912
	4/5	0.906	0.055	0.040	0.875
	5/6	0.987	0.013	0.000	0.992

Table 3.4.4.5

Accuracy and Consistency of Classification Indices: Over (Grade 5) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)	
	0.777	0.711		0.583	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.848		0.845	
	2	0.762		0.656	
	3	0.863		0.798	
	4	0.773		0.729	
	5	0.544		0.434	
	6	-		-	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.986	0.008	0.006	0.984
	2/3	0.971	0.016	0.013	0.962
	3/4	0.935	0.028	0.036	0.913
	4/5	0.884	0.076	0.040	0.856
	5/6	0.988	0.012	0.000	0.992

Table 3.4.4.6

Accuracy and Consistency of Classification Indices: Over (Grade 6) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.848	0.791		0.697	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.899		0.841	
	2	0.831		0.758	
	3	0.875		0.826	
	4	0.800		0.743	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.979	0.010	0.011	0.970
	2/3	0.952	0.022	0.025	0.933
	3/4	0.931	0.031	0.038	0.903
	4/5	0.986	0.014	0.000	0.985
	5/6	N/A	N/A	N/A	N/A

Table 3.4.4.7

Accuracy and Consistency of Classification Indices: Over (Grade 7) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.832	0.770		0.681	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.899		0.841	
	2	0.806		0.725	
	3	0.855		0.801	
	4	0.788		0.740	
	5	-		-	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.971	0.013	0.016	0.959
	2/3	0.947	0.025	0.028	0.924
	3/4	0.938	0.030	0.032	0.912
	4/5	0.977	0.023	0.000	0.974
	5/6	N/A	N/A	N/A	N/A

Table 3.4.4.8

Accuracy and Consistency of Classification Indices: Over (Grade 8) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)
	0.823	0.756		0.669
Conditional on Level	**Level**	**Accuracy**		**Consistency**
	1	0.906		0.853
	2	0.791		0.705
	3	0.839		0.777
	4	0.782		0.736
	5	-		-
	6	N/A		N/A
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.969	0.014	0.017	0.955
	2/3	0.946	0.025	0.028	0.924
	3/4	0.935	0.031	0.034	0.909
	4/5	0.973	0.027	0.000	0.968
	5/6	N/A	N/A	N/A	N/A

Table 3.4.4.9

Accuracy and Consistency of Classification Indices: Over (Grade 9) S403 Online

Overall Indices	Accuracy	Consistency		Kappa (k)
	0.839	0.774		0.689
Conditional on Level	**Level**	**Accuracy**		**Consistency**
	1	0.922		0.875
	2	0.768		0.682
	3	0.873		0.825
	4	0.801		0.731
	5	0.744		0.555
	6	N/A		N/A
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.970	0.011	0.019	0.958
	2/3	0.948	0.028	0.024	0.926
	3/4	0.943	0.028	0.029	0.919
	4/5	0.979	0.015	0.006	0.971
	5/6	N/A	N/A	N/A	N/A

Table 3.4.4.10

Accuracy and Consistency of Classification Indices: Over (Grade 10) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.845	0.782		0.695	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.912		0.858	
	2	0.780		0.697	
	3	0.873		0.827	
	4	0.825		0.760	
	5	0.759		0.575	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.971	0.011	0.018	0.960
	2/3	0.948	0.027	0.025	0.926
	3/4	0.944	0.029	0.027	0.920
	4/5	0.983	0.012	0.005	0.976
	5/6	N/A	N/A	N/A	N/A

Table 3.4.4.11

Accuracy and Consistency of Classification Indices: Over (Grade 11) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.841	0.777		0.690	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.905		0.848	
	2	0.787		0.705	
	3	0.869		0.823	
	4	0.813		0.749	
	5	0.753		0.556	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.971	0.012	0.017	0.959
	2/3	0.947	0.027	0.026	0.925
	3/4	0.944	0.029	0.028	0.920
	4/5	0.980	0.015	0.005	0.973
	5/6	N/A	N/A	N/A	N/A

Table 3.4.4.12

Accuracy and Consistency of Classification Indices: Over (Grade 12) S403 Online

Overall Indices	**Accuracy**	**Consistency**		**Kappa (k)**	
	0.845	0.785		0.697	
Conditional on Level	**Level**	**Accuracy**		**Consistency**	
	1	0.906		0.851	
	2	0.802		0.724	
	3	0.874		0.829	
	4	0.791		0.725	
	5	0.770		0.505	
	6	N/A		N/A	
Indices at Proficiency Level Cut Points	**Proficiency Level Cut Point**	**Accuracy**			**Consistency**
		Accuracy	**False Positives**	**False Negatives**	
	1/2	0.968	0.014	0.018	0.955
	2/3	0.943	0.029	0.028	0.919
	3/4	0.949	0.026	0.025	0.927
	4/5	0.986	0.013	0.001	0.983
	5/6	N/A	N/A	N/A	N/A

3.5 CSEM for Composites

Conditional standard errors of measurement (CSEMs) for the four ACCESS composites provide test users a benchmark of how free the composite scale score is from measurement errors at the various points of the composites. Due to the differential weights applied to different ACCESS domains (see the introduction to this chapter for the weighting conventions), the CSEMs were estimated using a procedure based on item response theory (IRT; Lord, 1980) and developed by Price, Lurie, Raju, Wilkins, and Zhu (2006). Price et al. (2006) extended the work by Lord (1980) and Kolen, Hanson, and Brennan (1992) in estimating the CSEM of a composite consisting of subtests. The basic premise of this procedure is that the student-level CSEM for a weighted composite can be estimated empirically using the IRT-based CSEMs for each student on the subtests and the weights associated with the subtests. This method was used to estimate the CSEM for ACCESS composites by treating the ACCESS domains as subtests.

A three-step process was used in deriving the CSEM for ACCESS composites. The derivation was conducted by grade and composite to obtain a unique CSEM for each composite score by grade by composite. Since this procedure replies on empirical student data, which are subject to year-to-year fluctuation, all population student data from the previous ACCESS series were used in the derivation to obtain more stable estimates than using only data from a single series.

Step 1. Since ACCESS domains were calibrated separately, measurement errors associated with each of the ACCESS domains, as expressed in the conditional errors of measurement, are independent of each other. Therefore, the CSEM for a student with composite score x, SEM_x, can be estimated using the equation derived by Price et al. (2006):

$$SEM_x = \sqrt{W_1^2 SEM_1^2 + W_2^2 SEM_2^2 + W_3^2 SEM_3^2 + \cdots + W_k^2 SEM_k^2}$$

Where SEM_i^2 is the student's IRT-based score error variance or student's squared CSEM in ACCESS domain i and W_i is the weight applied to domain i, for i=1,…,k.

Step 2. Due to the differential weights applied to different ACCESS domains, two students with the same sum of weighted domain score, or composite, may obtain different CSEMs; therefore, an additional step was taken to obtain a unique value for each composite score. Specifically, the expected value of the CSEM functions for a composite score was estimated using a regression approach, and this expected value was used as the reported CSEM for that composite score.

Step 3. A linear smoothing procedure was applied to derive the CSEMs for composite scores that were not observed in the data.

The figures in this section show graphically the CSEMs for various composite scores by grade level. Figures show the relationship between the students' composite scores on the horizontal axis and conditional measurement errors on the vertical axis. Each point in the figures represents a student in the dataset, expressing both the student's CSEM and that student's scale score for the given composite score. Values for students who received the lowest possible scores on any ACCESS domains are not plotted, as the conditional measurement errors for these students cannot be computed accurately. For grade-level clusters with multiple grades, different colors are used in the figures to represent students in different grades.

Five vertical lines in the figure indicate the five ACCESS cut scores for the highest grade in the grade-level cluster for the test form, dividing the figure into six sections for each of the WIDA proficiency levels (1–6) for the composites.

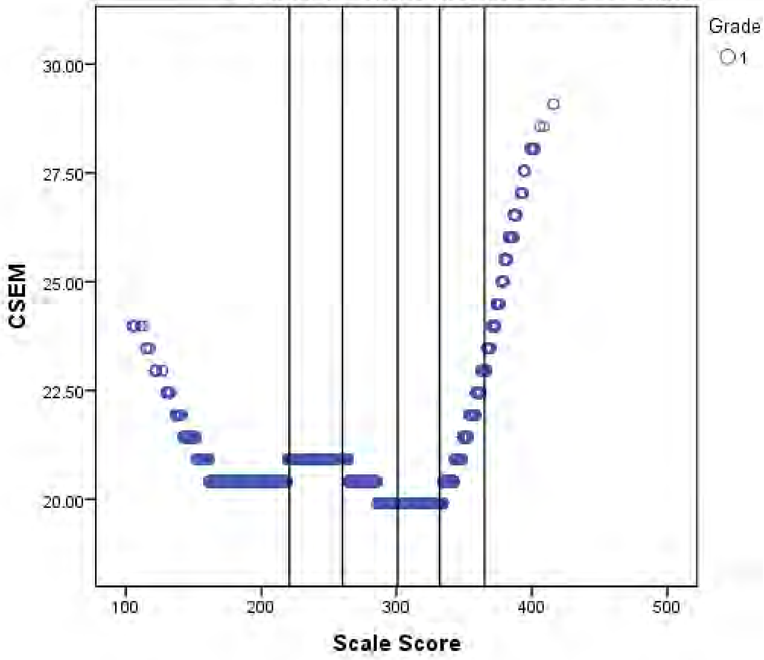
Low CSEM values indicate less measurement error or more accuracy in measurement. The general trend in these figures shows that the CSEMs are lower and fairly constant in the middle of the score range and higher and more variable for extreme low and high scores, as expected.

3.5.1 Oral

3.5.1.1 *Grade 1*

Figure 3.5.1.1

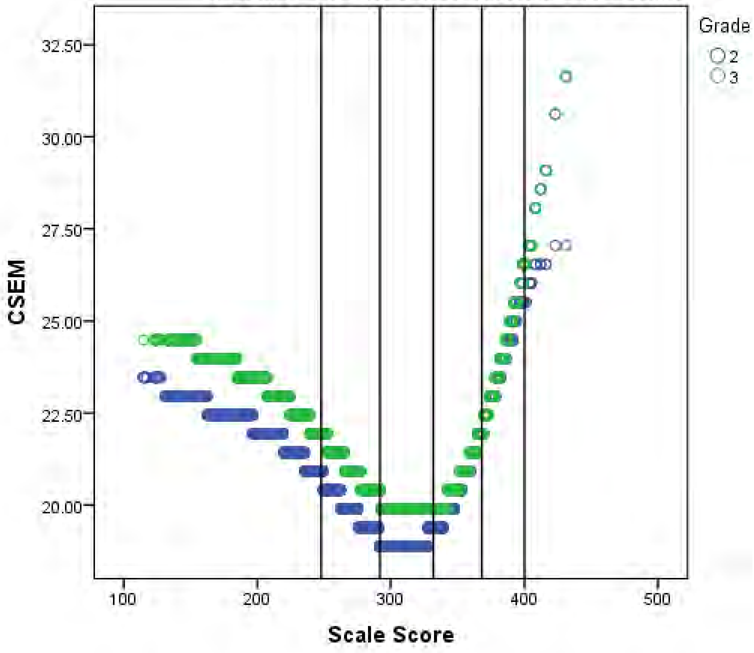
CSEM for Composite: Oral 1 S403 Online



3.5.1.2 *Grades 2–3*

Figure 3.5.1.2

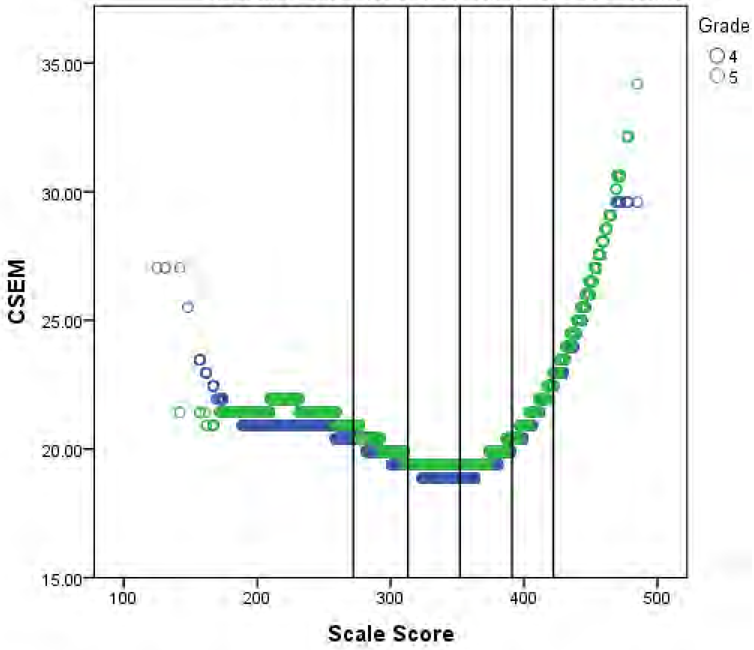
CSEM for Composite: Oral 2-3 S403 Online



3.5.1.3 *Grades 4–5*

Figure 3.5.1.3

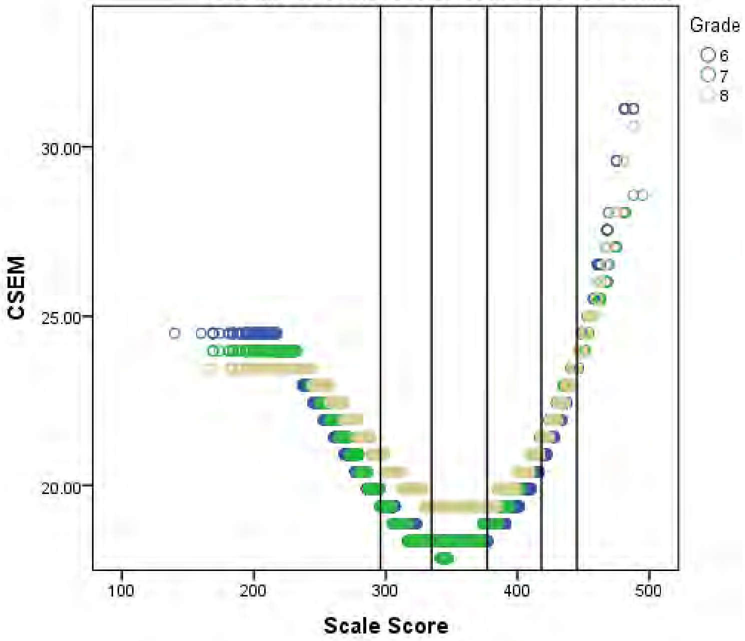
CSEM for Composite: Oral 4-5 S403 Online



3.5.1.4 *Grades 6–8*

Figure 3.5.1.4

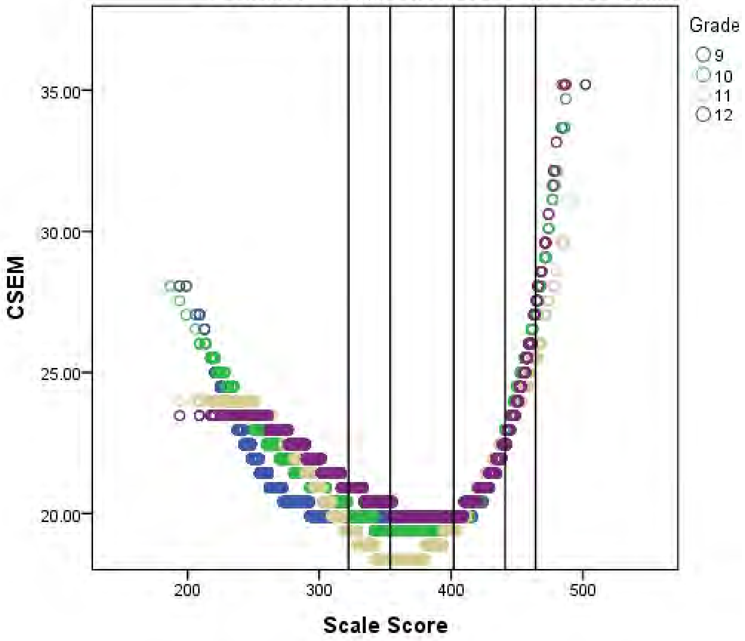
CSEM for Composite: Oral 6-8 S403 Online



3.5.1.5 Grades 9–12

Figure 3.5.1.5

CSEM for Composite: Oral 9-12 S403 Online



3.5.2 Literacy

3.5.2.1 *Grade 1*

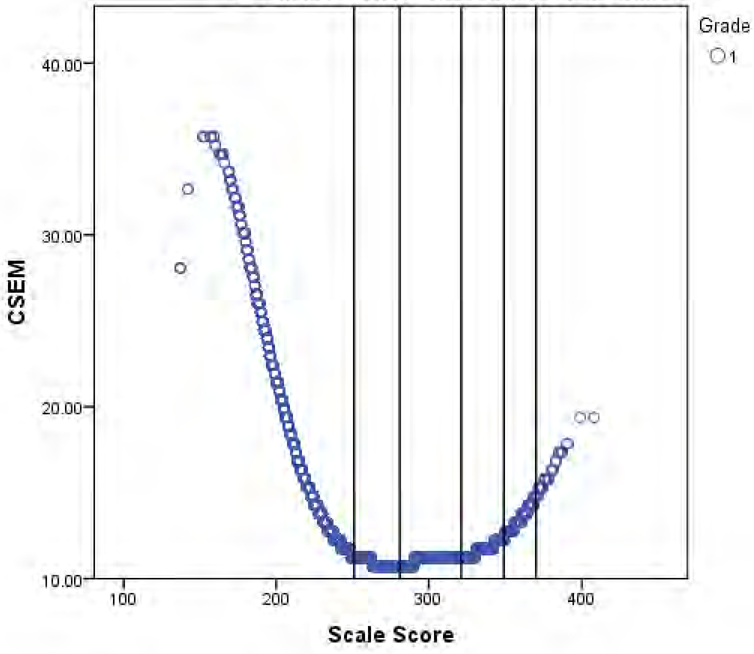


Figure 3.5.2.1

CSEM for Composite: Litr 1 S403 Online

3.5.2.2 *Grades 2–3*

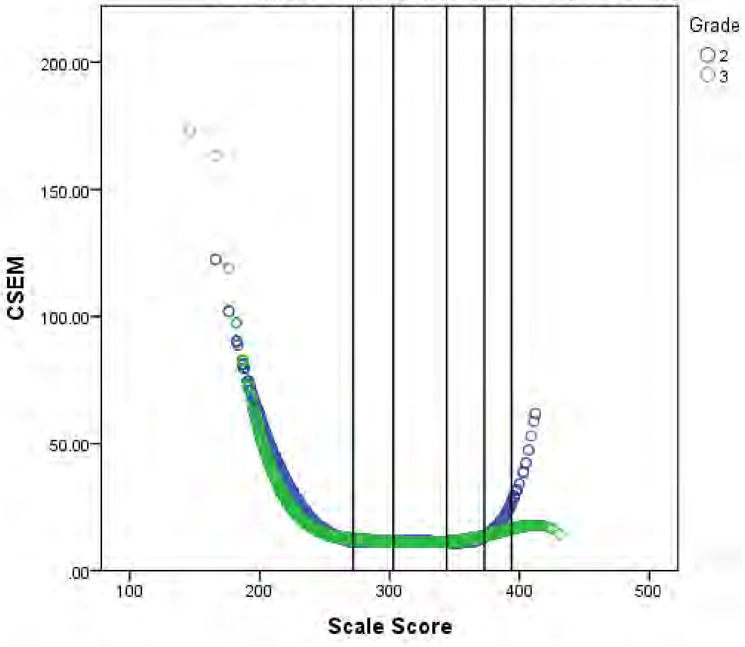


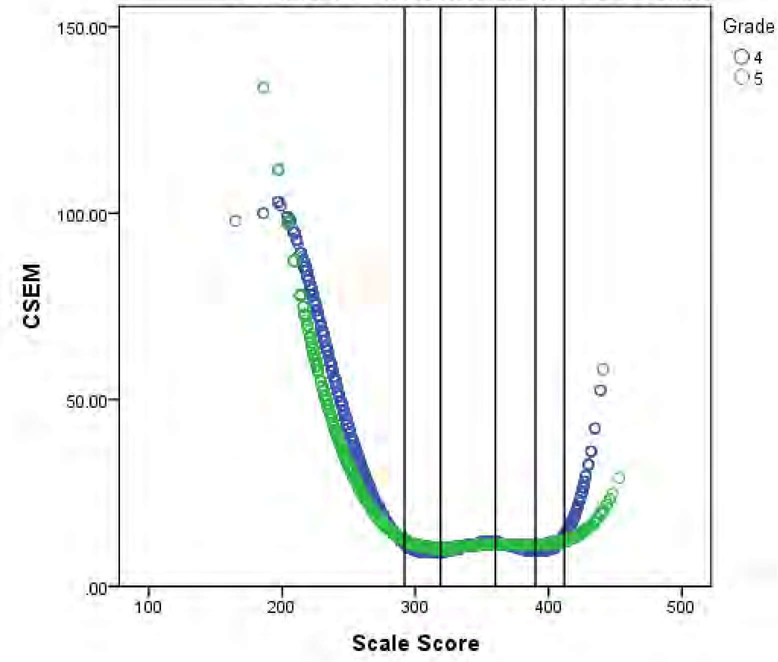
Figure 3.5.2.2

CSEM for Composite: Litr 2-3 S403 Online

3.5.2.3 Grades 4–5

Figure 3.5.2.3

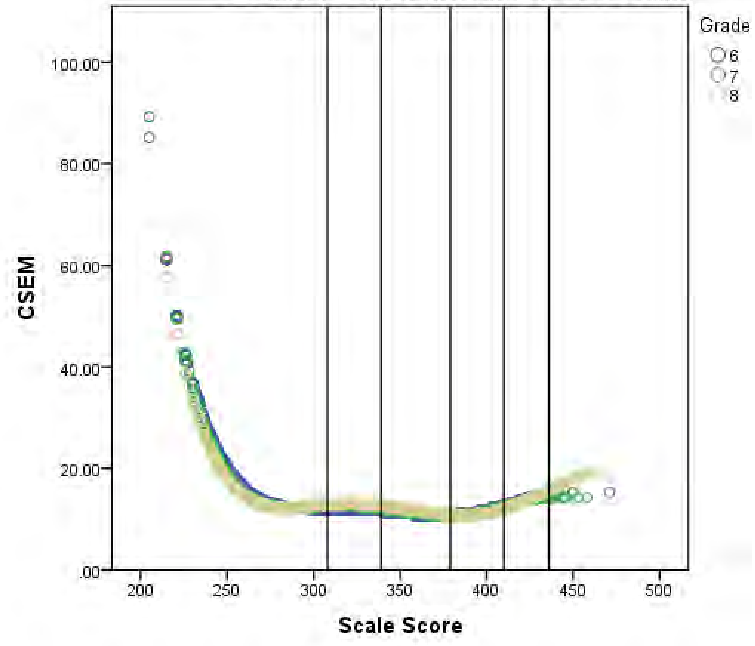
CSEM for Composite: Litr 4-5 S403 Online



3.5.2.4 Grades 6–8

Figure 3.5.2.4

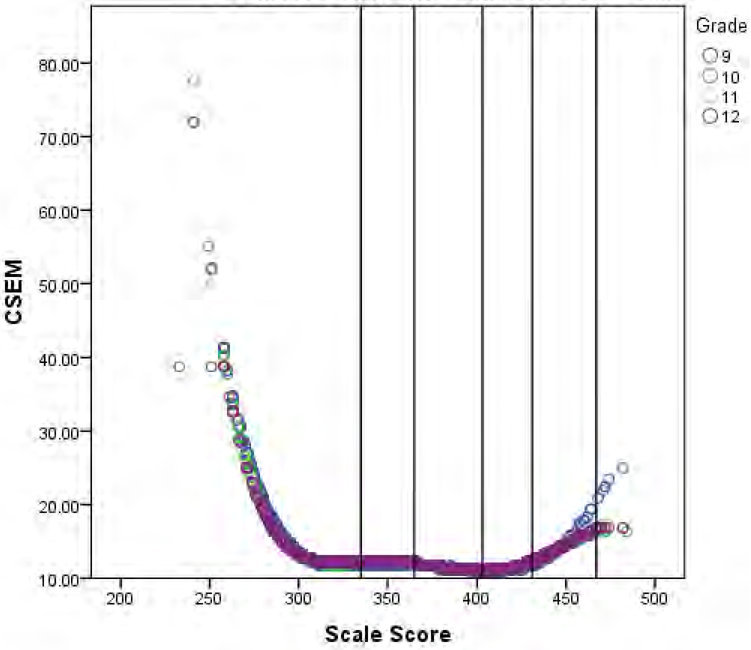
CSEM for Composite: Litr 6-8 S403 Online



3.5.2.5 *Grades 9–12*

Figure 3.5.2.5

CSEM for Composite: Litr 9-12 S403 Online

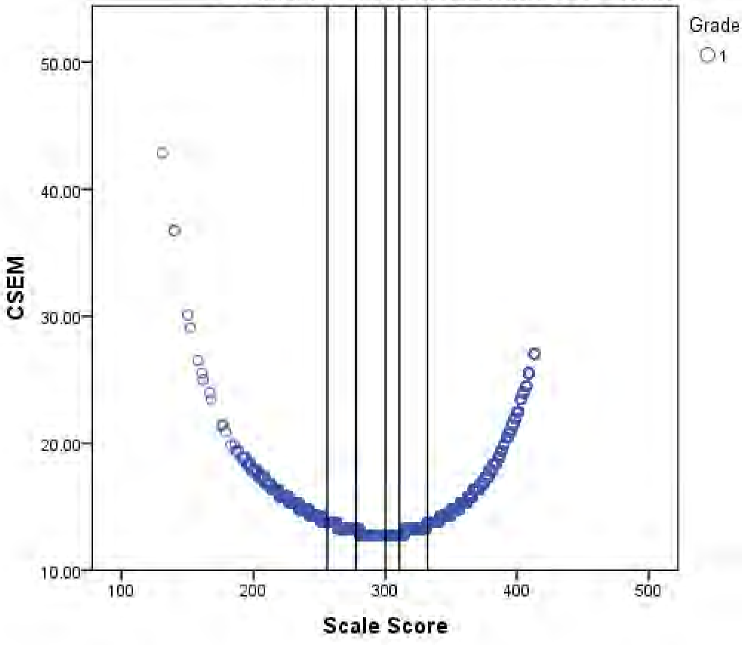


3.5.3 Comprehension

3.5.3.1 Grade 1

Figure 3.5.3.1

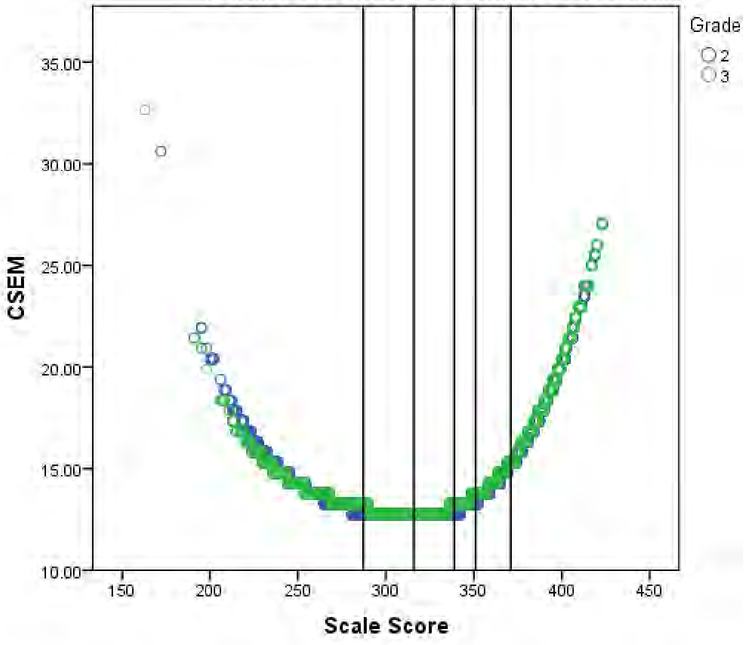
CSEM for Composite: Cphn 1 S403 Online



3.5.3.2 Grades 2–3

Figure 3.5.3.2

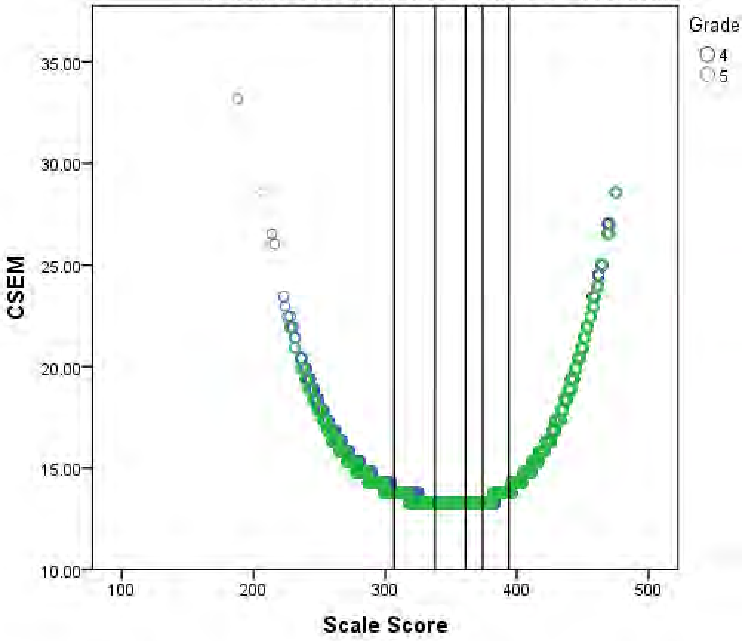
CSEM for Composite: Cphn 2-3 S403 Online



3.5.3.3 Grades 4–5

Figure 3.5.3.3

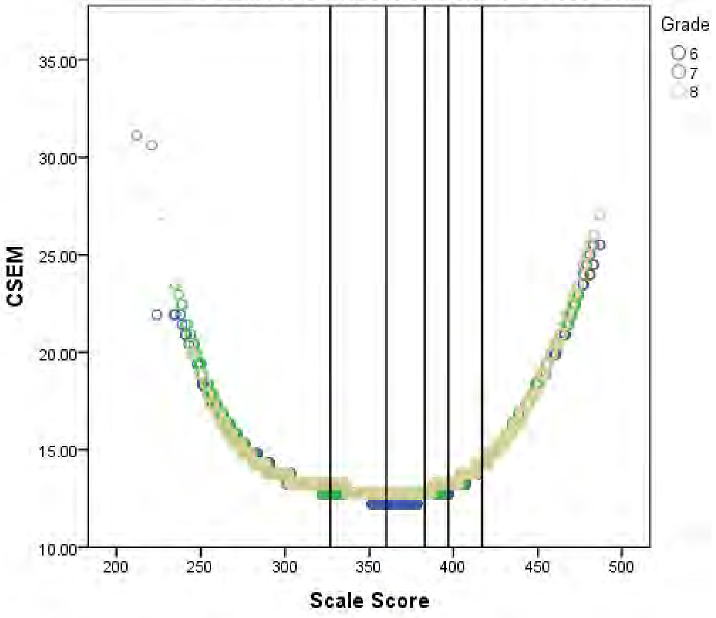
CSEM for Composite: Cphn 4-5 S403 Online



3.5.3.4 Grades 6–8

Figure 3.5.3.4

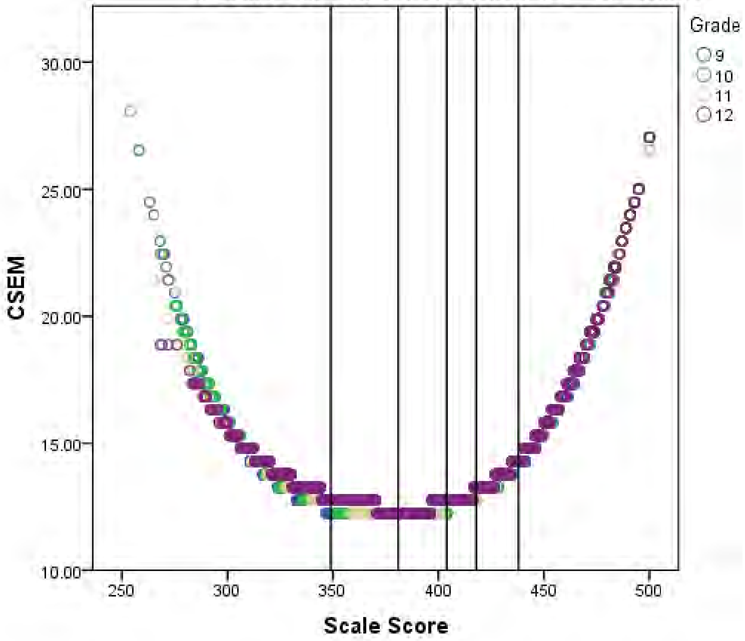
CSEM for Composite: Cphn 6-8 S403 Online



3.5.3.5 Grades 9–12

Figure 3.5.3.5

CSEM for Composite: Cphn 9-12 S403 Online

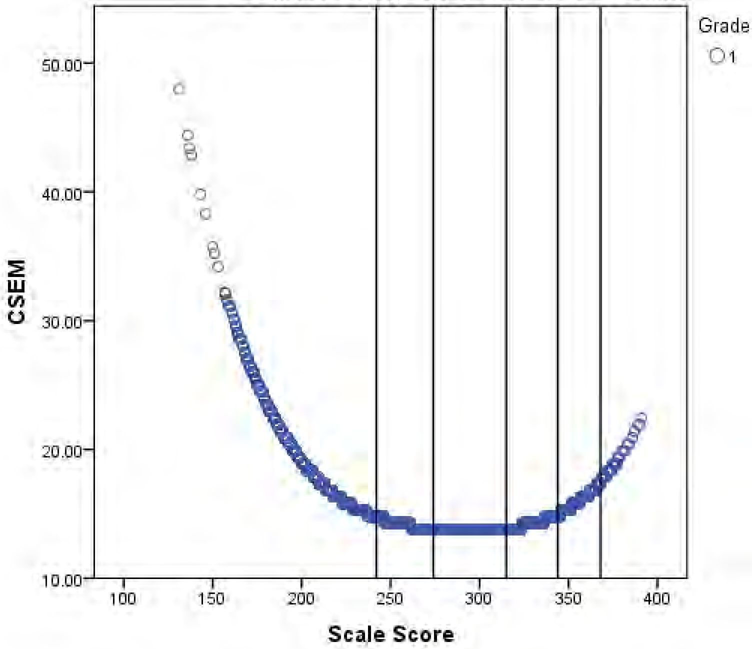


3.5.4 Overall

3.5.4.1 Grade 1

Figure 3.5.4.1

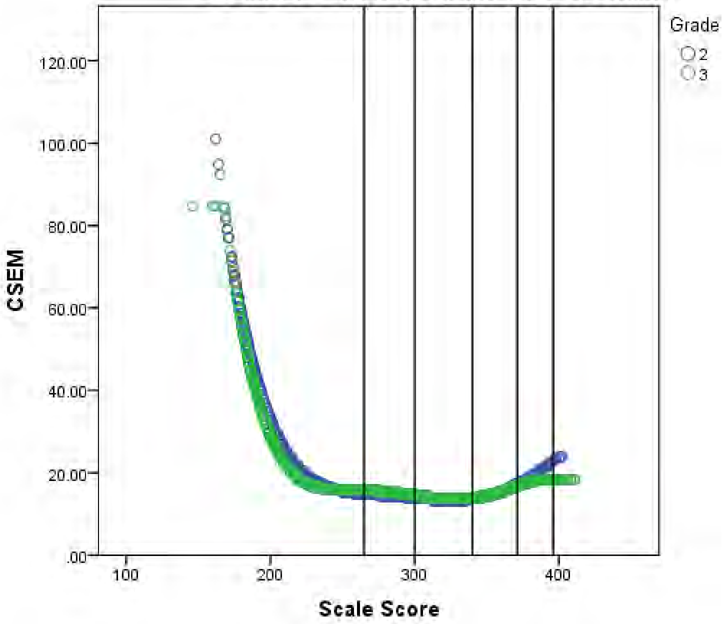
CSEM for Composite: Over 1 S403 Online



3.5.4.2 Grades 2–3

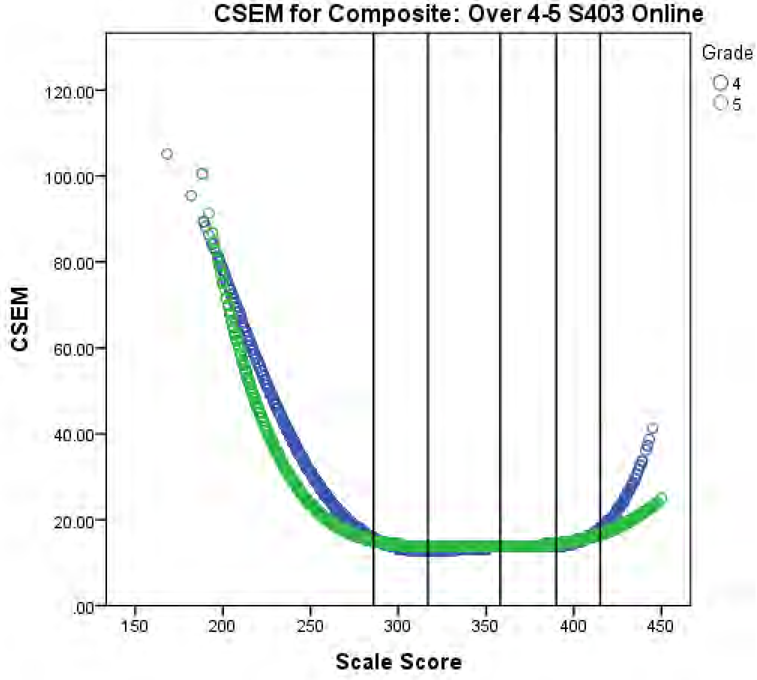
Figure 3.5.4.2

CSEM for Composite: Over 2-3 S403 Online



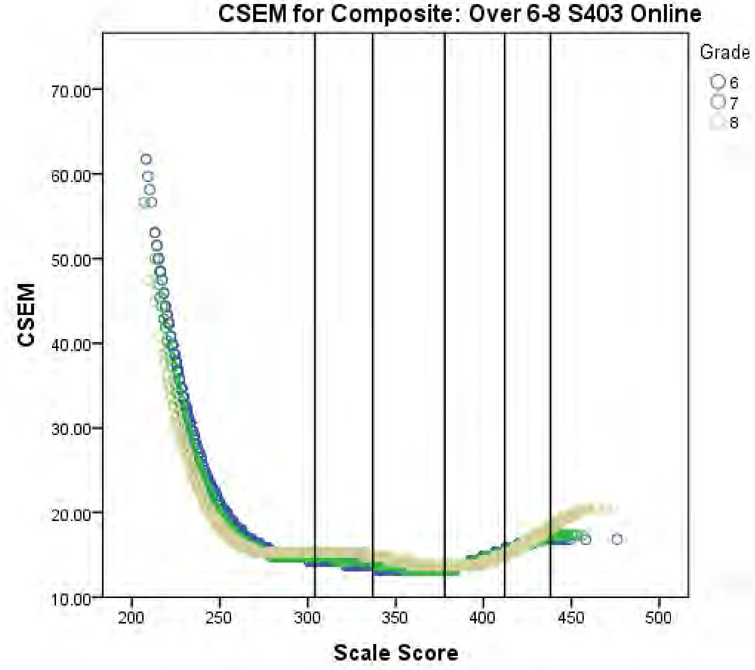
3.5.4.3 Grades 4–5

Figure 3.5.4.3



3.5.4.4 Grades 6–8

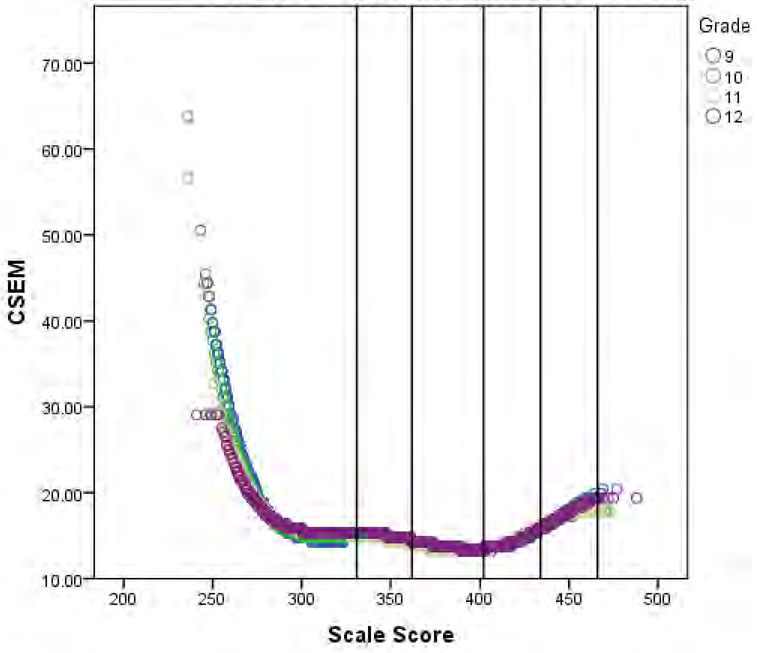
Figure 3.5.4.4



3.5.4.5 Grades 9–12

Figure 3.5.4.5

CSEM for Composite: Over 9-12 S403 Online



4. **Annual Updates of Validity Evidence**

According to the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 2014), validity is the degree to which all the accumulated evidence supports the intended interpretation of test scores for the proposed use.

Particular interpretations for specified uses begin by specifying the construct the test is intended to measure. Validity is the degree to which all the accumulated evidence supports the intended interpretation of the test scores for the proposed use. Rather than referring to distinct types of validity, the *Standards* refer to types of validity evidence.

According to the *Standards* the evidence can be based on 1) test content, 2) response processes, 3) internal structure, and 4) relations to other variables.

4.1. Standards

4.1.1. Test content

Important validity evidence can be obtained from an analysis of the relationship between the content of a test and the construct it is intended to measure. Test content refers to the themes, wording, and format of the items, tasks, or questions on a test. Administration and scoring may also be relevant to content-based evidence. Evidence based on test content can include logical or empirical analyses of the adequacy with which the test content represents the content domain and of the relevance of the content domain to the proposed interpretation of test scores. Evidence based on test content can also come from expert judgement of the relationship between parts of the test and content.

4.1.2. Response processes

Theoretical and empirical analyses of the response processes of test takers can provide evidence concerning the fit between the construct and the detailed nature of the performance or response actually engaged in by test takers. Evidence based on response processes generally comes from analysis of individual responses. Evidence of response processes can contribute to answering questions about differences in meaning or interpretation of test scores across relevant sub-groups of test takers. Studies of response processes are not limited to the test taker. Assessment often relies on observers or judges to record and/or evaluate test takers’ performances or products.

4.1.3. Internal structure

Analyses of the internal structure of a test can indicate the degree to which the relationships among the test items and test components conform to the construct on which the proposed test score interpretations are based. The conceptual framework for a test may imply a single dimension of behavior, or it may posit several components that are each expected to be homogeneous.

4.1.4. Relations to other structure

In many cases, the intended interpretation for a given use implies that the construct should be related to some other variables, and as a result, analysis of the relationship of the scores to variables external to the test provides another important source of validity evidence. Evidence about relations to other variables is also used to investigate questions of differential prediction for subgroups. In the test-criterion relationship the fundamental question is, how accurately do test scores predict criterion performance? Historically, two designs, often called predictive and concurrent, have been differentiated for evaluating test-criterion relationships. A predictive study indicates the strength of the relationship between test scores and criterion scores that are obtained at a later time. A concurrent study obtains test scores and criterion information at about the same time.

4.2. Annual validity studies

4.2.1. ACCESS for ELLs 2.0 Series 400 Construct Validity Study

Grant, R. & Kraninger, N. (2019). *ACCESS for ELLs 2.0 Construct Validation Study - ACCESS for ELLs 2.0, Series 400 with 2015-2016 Data*. WIDA Technical Report TR-2019-1 (unpublished).

A research study was conducted to examine the construct validity of the internal structures of the paper and online versions of ACCESS for the grade-level cluster forms within Grades 1-12 through construct validity techniques at the item level. The research questions were:

1. Are there 4 distinct, orthogonally related, first-order academic English language factors (Listening, Reading, Speaking, and Writing) underpinning English Learners' responses to Tiers A, B and C items within the grade-level cluster forms of the paper and online versions of the assessment?
2. If so, is there a 2nd-order, underlying factor (academic English language) orthogonally related to each first-order factor within each tier?

Confirmatory factor analyses were used to investigate the first research question and structural equation modeling techniques were used to investigate the second question. All 33 models had satisfactory model fit statistics. The conclusion was that the ACCESS for ELLs 2.0, Series 400 assessment (online and paper) has an internal structure consistent with four academic English language domains (Listening, Reading, Writing and Speaking) and a second order underlying language factor (academic English language). This suggests that it is appropriate to report separate English language domain scores and an overall composite score for the paper and online versions of the test.

4.2.2. WIDA Screener Online and WIDA ACCESS for ELLs: Examining the Relationship between Student Scores on Two Assessments

MacGregor, D., and Sahakyan, N. (in preparation). *WIDA Screener Online and WIDA ACCESS for ELLs: Examining the Relationship Between Student Scores on Two Assessments*. WIDA Technical Brief.

WIDA conducted a comparability study between WIDA Online Screener (hereafter Screener) and WIDA ACCESS (hereafter ACCESS). The purpose of the study was to understand to what extent scores on Screener predict scores on ACCESS, and how that relationship is potentially affected by several covariates, such as grade level, time between tests, and IEP status.

The results of this analysis show that Screener scores are highly correlated with and strongly predictive of ACCESS scores controlling for individual-level and institutional factors. Controlling for all of the above-mentioned factors, a one-point increase in the Screener Overall Composite Scale Score is associated with about a 0.70 point increase in the ACCESS Overall Composite Scale. Students with special needs score about 8 scale score points lower on ACCESS, compared to non-IEP students. Hispanic ELs score about 5 points below non-Hispanics, while female ELs score about 2.3 scale score points higher compared to their male counterparts. Finally, parameter estimates indicate that in addition to individual-level covariates, there are statistically significant grade, school, district and state effects; however, the predictive relationship between the two tests is not altered when these factors are accounted for.

These results provide evidence for the predictive validity of Screener for schools in the WIDA Consortium. The stable and strong predictive relationship between Screener and ACCESS Overall composite scores suggests that schools can confidently use the WIDA Screener Overall composite score when making placement decisions for ELLs. In addition, they provide validity evidence for the use of Screener as an identification tool for students in the WIDA Consortium.

4.2.3. 2018 ACCESS for ELLs 2.0 Speaking Study

Bishop, K., Read, S., Gocer-Sahin, S., and Akanda, M. (2019). *ACCESS for ELLs Speaking study*. WIDA Technical Report.

Although the amount of score difference varied year to year, in the domain of Speaking, paper scores have been consistently higher than online counterpart. The purpose of this study was to investigate potential differences in the test-taking and scoring procedures between the paper Speaking test and the online Speaking test. The results of this study are related to content validity since it explores possible reasons of differences among different administration and scoring procedures (AERA, APA & NCME, 2014) and evidence based on response processes.

In this study a WIDA researcher used the Speaking scores of ACCESS 401 population data. Data was collected from IL and FL. The WIDA researcher went to each school site to set up the speaking test sessions and met teachers to administer and score the test. The student's entire testing session and spoken responses were audio recorded by the WIDA researcher.

A quantitative and qualitative analysis was conducted. As the first step of the quantitative analysis, descriptive statistics such as raw score distributions, means across states, tiers, and raters were calculated. Second, to explore the rater effects of test administrative modes, Patz's (1996) hierarchical rater model (HRM) for polytomous Speaking rating data scored by multiple raters to scale examinees and items was applied to model aspects of consensus among raters, and to model individual rater severity and consistency effects (Patz, Junker, Johnson, and Mariano, 2002).

Qualitative analysis was conducted in four steps.

1) Interviews: After the WIDA researcher recorded students' Paper speaking test responses, we interviewed paper raters for their reasoning of scores. For online raters, the interview questions were sent to them and after their scorings were done, they typed their answers to the survey questionnaire document.
2) Categorizing interviews and observations: Based on interviews and the observations from administering tests and testing environment, the researcher categorized the interview comments into themes.
3) FL score comparison: one WIDA researcher, a rater trainer in FL scored students' responses as she was observing the paper testing. The WIDA researcher's scores and FL local raters' scores and online raters' score in FL were compared.
4) Score difference examination: two WIDA researchers examined paper and online ratings where there were score differences for 50 tasks to explore possible reasons of score differences.

The quantitative analysis results, HRM results, showed us that online raters are more consistent in assigning the same score to work of the same quality than paper raters. Tier B/C is more difficult than Tier A. Interestingly, although number of ideal scoring is three for paper raters for tier A, it is six for tier B/C. Similar to paper raters, in tier A, number of ideal scoring of online raters is seven, however it is eight for tier B/C. Although there isn't much change for online raters, in general it can be said that when test becomes more difficult, raters become more reliable in scoring.

Based on the local observations and rater reliability results, the WIDA Researcher has identified a need for additional clarification of the various aspects of the Speaking domain that should be provided in local and online trainings. It can be said that there is no one reason to give a high or low score. However, WIDA Researcher observed that, word choice was overlooked as sophistication. Even if sophistication was not present in the answer, online raters scored 3. Even though one of three criteria was missing, both raters still gave score 3 no matter which criterion is missing.

It was also observed that sometimes teacher prompting occurred, but they were within legitimate promoting options according to test administrator manual. Occasionally some prompting was very guided and specific, but student answers did not get more sophisticated. Teacher-prompted answers did not always lead to higher scores by paper raters. Some paper raters paid more attention to criteria outside of the rubric such as “complete sentences” and scored lowered. Interpretation of Nina model to the rubric description is an issue. For example, for P3 or P5 questions, Nina model answer has detailed explanations. It is not clear whether this explanation is referring to sophistication level in the rubric.

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, DC: American Psychological Association.

Brennan, R. (2004). *Linking with equivalent group or single group design (LEGS)* (Version 2.0) [Computer software]. Iowa City, IA: Center for Advanced Studies in Measurement and Assessment.

Center for Applied Linguistics. (2016). *ACCESS for ELLs® Series 400 Listening and Reading scale maintenance: Technical brief.* Washington, DC: Author.

Center for Applied Linguistics. (2017). *ACCESS for ELLs® 2.0 Speaking and Writing score scale reconstruction: Technical brief.* Washington, DC: Author.

Cook, H. G., & MacGregor, D. (2017). *The ACCESS for ELLs 2.0 2016 Standard setting study* (Technical Report). Madison, WI: Board of Regents of the University of Wisconsin System.

Elementary and Secondary Education Act of 1965, amended 2015. 20 USC §6301-8961.

Gottlieb, M. (2004). *English Language Proficiency Standards for English language learners in Kindergarten through Grade 12: Framework for large-scale state and classroom assessment.* Madison, WI: WIDA Consortium.

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

Kenyon, D. M. (2006). *Development and field test of ACCESS for ELLs®* (WIDA Consortium Technical Report No. 1). Washington, DC: Center for Applied Linguistics.

Kenyon, D. M., Ryu, J. R., & MacGregor, D. (2013). *Setting grade level cut scores for ACCESS for ELLs®* (WIDA Consortium Technical Report No. 4). Washington, DC: Center for Applied Linguistics.

Kolen, M. J., Hanson, B.A., & Brennan, R. L. (1992). Conditional standard errors of measurement. *Journal of Educational Measurement, 29*, 285–307.

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. (1999). Relating Cronbach and Rasch reliabilities. *Rasch Measurement Transactions, 13*(2), 696. Retrieved from http://www.rasch.org/rmt/rmt132i.htm

Linacre, J. M. (2002, Autumn). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions, 16*(2), 878. Retrieved from http://www.rasch.org/rmt/rmt162f.htm

Linacre, J. M. (2006). Winsteps Rasch analysis (Version 3.60.1) [Computer software]. Retrieved from http://www.winsteps.com

Linacre, J. M. (n.d.). *Displacement measures*. Retrieved from http://www.winsteps.com/winman/displacement.htm

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

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NO: Lawrence Erlbaum Associates.

Mantel, N., & Haenszel, W. (1959). Statistical aspect of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, *22,* 719–748.

Meyer, J. P. (2018). jMetrik [Computer software]. Retrieved from http://itemanalysis.com/jmetrik-download/

Patz, R. J. (1996). *Markov chain Monte Carlo methods for item response theory models with applications for the National Assessment of Educational Progress*. Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA.

Patz, R. J., Junker, B. W., Johnson, M. S., & Mariano, L. T. (2002). The hierarchical rater model for rated test items and its application to large-scale educational assessment data. *Journal of Educational and Behavioral Statistics, 27*(4), 341–384. doi:10.3102/10769986027004341

Price, L. R., Lurie, A., Raju, N., Wilkins, C., & Zhu, J. (2006). Conditional standard errors of measurement for composite scores on the Wechsler Preschool and Primary Scale of Intelligence – Third Edition. *Psychological Reports, 98*(1), 237–252.

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

Thissen, D. (2000). Reliability and measurement precision. In H. Wainer, N. Dorans, D. Eignor, R. Flaugher, B. Green, R. Mislevy, L. Steinberg, & D. Thissen (Eds.), *Computerized adaptive testing: A primer* (2nd ed., pp. 159–184). Hillsdale, NJ: Lawrence Erlbaum Associates.

WIDA Consortium. (2007). *English Language Proficiency Standards and resource guide, 2007 edition, PreKindergarten through Grade 12*. Madison, WI: Board of Regents of the University of Wisconsin System.

WIDA Consortium. (2012). *2012 amplification of the English Language Development Standards Kindergarten–Grade 12*. Madison, WI: 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 Technical Report 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). DIF statistics in test development. In P.W. Holland and H. Wainer (Eds.), *Differential item functioning* (pp. 337-347). Hillsdale, NJ: Erlbaum.

Zwick, R., & Bridgeman, B. (2014). Evaluating validity, fairness, and differential item functioning in multistage testing. In Y. Duanli, A. A. von Davier, & C. Lewis (Eds.), *Computer multistage testing: Theory and applications* (pp. 271–284). Hoboken, NJ: CRC Press.

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

Zwick, R., Thayer, D. T., & Wingersky, M. (1993). *A simulation study of methods for addressing differential item functioning in computer-adaptive tests* [ETS Research Report RR-93-11]. Princeton, NJ: Educational Testing Service. Retrieved from https://onlinelibrary.wiley.com/doi/epdf/10.1002/j.2333-8504.1993.tb01522.x

Acknowledgments

We would like to extend our appreciation to the many CAL and WIDA staff members who have supported this work, including the following:

From CAL:

Keira Ballantyne, Ph.D.
Tanya Bitterman, M.A.
Yage (Leah) Guo, Ph.D.
Bernadette Jerome, M.S.
Michele Kawood, M.S.Ed.
Justin Kelly, Ph.D.
Dorry M. Kenyon, Ph.D.
Tristan Kirkman, M.A.
Mohammed Louguit, Ph.D.
Erin Shaw-Meadow, M.Sc.
Samantha Musser, M.A.
Yu-Chia Wu, M.Ed.
Shu Jing Yen, Ph.D.
Xin Yu, M.A.

From WIDA:
Mohammad Akanda, Ph.D.
Kyoungwon Bishop, Ph.D.
Sakine Göçer Sahin, Ph.D.